Joe's Little Book of String[1]
Class Notes, Phys 230A, String Theory, Winter 2010

Last update: 3/14/12

When I wrote the Big Book of String, I had many goals. One was to make it very readable, so you would pick it up, be unable to put it down, and after staying up all night reading you would know string theory. But in spite of much effort, this didn't happen. The desire to be general and systematic pulled in the opposite direction. So these notes are intended to be the book I might have written, and I can leave many details to the Big Book.

In addition, the subject has moved on and broadened. Much of the current research does not need heavy world-sheet machinery such as BRST. Some subjects still depend on this, notably string field theory, the covariant treatment of the superstring, pure spinors, and the topological string, but most applications of gauge/gravity duality do not. These are notes for a ten week course, with the goal of presenting an introduction to the world-sheet approach to string theory which is sufficiently complete to prepare you for 230B, which will focus on gauge/gravity duality.

1 Overview

Our goal is to find the underlying laws of physics, from which all else descend. Right now we have gotten to the three particle interactions, the $SU(3)\times SU(2)\times U(1)$ Standard Model, plus general relativity, with the quarks, leptons, and Higgs. This is fairly simple, fairly beautiful, and explains almost everything, but it can't be the end: there are too many moving parts, and too many arbitrary choices. The experience in physics (e.g. Maxwell) is that this should be unified. In addition there are problems, most notably in extending general relativity into the quantum regime, that require a new framework.

We don't have many clues from nature, because the SM+GR works so well. My book mentions neutrino masses, now confirmed and well-measured. Since that time we also have dark matter and dark energy. But these don't point conclusively in any direction. Neutrino masses are only a small step beyond the SM, not a big surprise. Dark matter and dark energy so far are only seen through their astrophysical effects, we haven't gotten our hands on them in the laboratory. We are therefore strongly dependent on theoretical consistency, finding *any* complete and consistent theory that incorporates what we already know. Fortunately, this has been fruitful, and it is the direction we will follow at first.

The first problem that one hits with QM+GR is nonrenormalizability. Imagine gravitational scattering of electrons, comparing the one-gravity exchange A_1 with the two-graviton

[1] Not to be confused with Steve Gubser's Little Book of String. These notes are dedicated to DG, who found the Big Book very, very confusing.

exchange A_2. A_2 has an extra factor of Newton's constant G_N. We will work with units $\hbar = c = 1$, wherein mass $\sim$ energy $\sim$ inverse time $\sim$ inverse length. Then

$$G_N = M_P^{-2} = L_P^2\,, \tag{1.1}$$

where the Planck mass and Planck length are

$$M_P = 1.2 \times 10^{19}\ \mathrm{GeV}/c^2\,, \quad L_P = 1.6 \times 10^{-33}\ \mathrm{cm}\,. \tag{1.2}$$

To balance the units, we must have

$$\frac{A_2}{A_1} \sim G_N \int dE'\, E' \tag{1.3}$$

where E' is the energy scale of the virtual particles in the loop. Thus diverges quadratically in the UV, and the divergences get worse with each additional loop, pointing to an incompleteness of the theory. This is nonrenormalizability, but we can also think of it as 'spacetime foam,' the fluctuations of the metric growing without bound at distances below the Planck length.

Such divergences have been an important clue in the past. Fermi's weak interaction theory with a 4-fermion vertex has a coupling $G_F = (300\ \mathrm{GeV})^{-2}$, the same units as G_N and so the same problem. In position space, these divergences arise when all the interactions take place at the same spacetime point. Somehow, new physics must smear out the interaction. There is one big reason why this is hard: special relativity implies that if we smear in space we must also smear in time, and this gives problems with causality or other basic principles like conservation of probability. If not for this, it would be easy to find consistent theories, and consistency would give little guidance. But in the case of the weak interaction, following this clue led to the idea that the four-Fermi vertex should be resolved into exchange of a W-boson, and moreover that this should come from spontaneously broken gauge symmetry. Before there was any direct indication of these vector bosons, we knew that the $W^\pm$ and Z^0 should exist, and what their masses and couplings should be to high accuracy. This is not to minimize the importance of experiment, which of course played a big role, but to illustrate that in some circumstances theoretical consistency can lead us quite far.

Note that the divergence does not become large until we get down to very small distances of order L_P. Max Planck first did this dimensional analysis in 1899, the year before he published his black body formula. I don't believe that he ever commented on how incredibly short this length scale is, but it means that the unification of gravity and quantum mechanics takes place at a scale far beyond most of our observational tools. From a practical point of view this may not seem important, but the attempt to resolve the problem has taught us a lot.

Smearing out the gravitational interaction seems to require more than adding a few particles (pictorially, separating a four-point vertex into two three-point vertices is straightforward, but how do we split the three-point gravitational vertex?). String theory begins with the idea that we replace the point-like particles of quantum field theory with one-dimensional loops and strands. This is not an obvious idea, and languished in obscurity from 1974 (when proposed by Scherk, Schwarz, and Yoneya), until the 'First Superstring Revolution' in 1984 when the evidence for it reached a tipping point.[2]

Before going on, what about the alternate ideas that one hears about? Unfortunately, many of these, including most versions of loop quantum gravity, give up Lorentz invariance at an early stage. It's a bit like quantum field theory before Feynman, Schwinger, and Tomonaga. This makes the problem much easier, but it is never explained how one recovers all of the precise tests of Lorentz invariance, such as different species of particles having the same asymptotic velocity c; it seems that the importance of this is not appreciated. (I have recently reiterated the problem, in arXiv:1106.6346).

Another possibility is that the UV divergences are an artifact of perturbation theory, and disappear if the series is summed, so-called 'asymptotic safety.' This is a logical possibility, and there are some examples of nonrenormalizable theories in which it occurs. There is a lot of research going on claiming to find evidence for this in gravity, but I am skeptical. Another thing that must be checked is positivity of probability: one can make theories better behaved in the UV by replacing $1/p^2$ propagators with $1/(p^2 + p^4/\Lambda^2)$, but this has negative spectral weight.

If another solution to the UV problem were found, it would be interesting, but in the meantime string theory has gone on to provide solutions to some of the other seemingly intractable problems of quantum gravity, including black hole quantum mechanics. It has also provided fruitful new ideas to many other parts of physics, including particle physics, cosmology, and nuclear physics, and mathematics. Of course, string theory is not a finished theory, and in the past it has acquired important ideas from particle physics, cosmology, supergravity, and other approaches to quantum gravity, and it may do so again.

So we're going to quantize one dimensional objects in a Lorentz invariant way, and we're going to see that we automatically get gravity, and that the finite size of strings cuts off the loop integrals at short distance. We're also going to see that the quantization - surprise - requires more than four spacetime dimensions, which is a very old idea for unifying gravity with the other interactions. It also requires supersymmetry; we will first see the need for

[2]Following the discovery of asymptotic freedom, 1974 was a pretty busy year: lattice gauge theory, Hawking radiation, GUTs, monopole solutions, and the large-N limit of gauge theory were all discovered, as well as charm on the experimental side. So it's not too surprising that it took a while to get around to this other bizarre idea, though in retrospect there are deep connections with Hawking radiation, the large-N limit, and the rest.

this in the second lecture, and I hope to explain how it works by the end of the quarter.

Another attractive feature, which we will see early, is that the equations that define the theory are unique, they do not have all the free parameters, e.g. of the SM. But the flip side is that the number of solutions is extremely large. This is how physics usually works, all the diversity of nature is explained by a simple set of equations having a rich set of solutions, but it means that having unique equations does not lead to unique predictions. Actually, there are long-standing arguments, predating string theory, that in order to understand the vacuum energy we need a such a theory, we will see a bit about this as we go along.

The big open question is that we are still looking for the final form of the theory. In this quarter we will formulate string theory as depicted above, a sum over world-sheets that corresponds to fattened Feynman diagrams. But even in QFT the sum over Feynman diagrams does not converge. Thus it does not define the theory, and misses important physics at strong coupling. For QFT the first nonperturbative construction was given by Wilson in terms of the path integral plus the renormalization group. For string theory we are still looking for the defining principle. In fact, we have gone far beyond the sum over world-sheets, but this is still the best starting point.

Essentially, there is some quantum theory, and one of its weakly coupled limits is strings propagating in a classical spacetime. In quantum field theory, the weakly coupled limit is the same as the classical field limit, both are governed by a saddle point expansion of the path integral. So string theory is one limit of this quantum theory, but there are many others. First, all the different string theories (which I will try to tell you about by the end of the quarter) are different classical limits of this theory, but so also is supergravity in one additional dimension, and quantum mechanics of large matrices, and quantum field theory itself (via AdS/CFT, the subject of 230B).

A quantum theory having many classical limits might seem counterintuitive, but already in QFT, by holding different things fixed as we take $\hbar$ to zero, we get classical particles or classical fields as different limits. So wave-particle duality is just the most familiar example of this phenomenon of duality. Another remarkable example, which you will learn about in 230B, is that black holes and hot nuclear matter are also different classical limits of a single quantum object.

2 Classical strings in light-cone gauge

In this chapter and the next we want to get a first exposure to some of the physics, including the appearance of gravity and the need for a certain number of dimensions. In this chapter we find the classical action principle, the equations of motion and boundary conditions, the fixing of the world-sheet coordinates and the general classical solution. We will see that the classical motion separates into straight-line motion plus internal oscillations, with the mass of the string depending on the internal state of oscillation. In the next chapter we will quantize this. I will largely follow the classic paper Goddard, Goldstone, Rebbi, Thorn, Nucl.Phys.B56:109-135,1973 (GGRT); the text takes a more formal path integral approach based on the Polyakov action, but gets to the same physics.

Nambu-Goto action

For a nonrelativistic point particle we describe the motion in terms of $\vec{X}(t)$, with the action

$$S_{nonrel} = \frac{m}{2}\int dt\, \dot{X}^2 \tag{2.1}$$

In the relativistic case this becomes

$$S_{rel} = -m\int dt\,\sqrt{1-\dot{X}^2} = \int dt\left(-m + \frac{m}{2}\dot{X}^2 + \ldots\right) \tag{2.2}$$

The first term is the rest mass, which we leave out in the nonrelativistic case. The square root looks relativistic, but to make the Lorentz invariance manifest we can't use $\vec{X}(t)$ because it puts time on a different footing form the start. Rather, we introduce an arbitrary parameterization τ of the world-line $X^\mu(\tau)$, with $X^\mu = (t, \vec{X})$ on an equal footing. The action

$$S_{rel'} = -m\int d\tau\,\sqrt{-\partial_\tau X^\mu \partial_\tau X_\mu}\,, \tag{2.3}$$

where indices are lowered with $\eta_{\mu\nu} = diag(-1, 1, 1, \ldots)$, is manifestly invariant under Lorentz transformations and translations,

$$X^\mu \to \Lambda^\mu{}_\nu X^\nu + a^\mu\,. \tag{2.4}$$

(We use a spacelike-positive metric, and leave the number of spacetime dimensions D unspecified.) In addition, if we adopt two different parameterizations of the same path, the actions are the same. In other words, if the new parameter is some monotonic function of the old, $\tau'(\tau)$, and the paths $X^{\mu\prime}$ and X^μ are related by $X^{\mu\prime}(\tau') = X^\mu(\tau)$, then these functions trace the same path in spacetime and the actions are the same, the Jacobians cancelling. If we choose the parameterization $\tau = t$ then we recover the original action (2.2). We must choose between a noncovariant description, and a redundant one where different paths represent the same physical history.

For the string we use the same idea. The one dimensional string sweeps out a two dimensional world-sheet. We describe this with two coordinates, $X^\mu(\tau, \sigma)$. To write an invariant action we introduce first the combination

$$h_{ab} = \partial_a X^\mu \partial_b X_\mu\,, \tag{2.5}$$

where a, b run over τ, σ. Then the Nambu-Goto action

$$S_{NG} = -T \int d\tau d\sigma \sqrt{-\det h} \tag{2.6}$$

is manifestly Lorentz and translation invariant. It is also easily verified to be invariant under changing to a new parameterization $\tau'(\tau, \sigma), \sigma'(\tau, \sigma)$, the Jacobians canceling. The action $S_{rel'}$ (2.3) is proportional to the total proper time along the particle path, and the action S_{NG} is proportional to the total proper area. The constant T is the total rest mass per unit length, also equal to the string tension. For historic reasons, and partly for convenience, we work in terms of the *Regge slope*

$$\alpha' = 1/2\pi T\,. \tag{2.7}$$

Equations of motion and boundary conditions

To vary the action, recall a result from linear algebra for the variation of the determinant of any matrix,

$$\delta \det h = \delta h_{ab} h^{ab} \det h \tag{2.8}$$

where h^{ab} is the inverse matrix. Then

$$\begin{aligned} \delta S_{\rm NG} &= -\frac{1}{4\pi\alpha'} \int d\tau d\sigma\, \delta h_{ab} h^{ab} \sqrt{-\det h} \\ &= -\frac{1}{2\pi\alpha'} \int d\tau d\sigma\, \partial_a \delta X_\mu \partial_b X^\mu h^{ab} \sqrt{-\det h} \\ &= \frac{1}{2\pi\alpha'} \int d\tau d\sigma\, \delta X_\mu \partial_a \left(\partial_b X^\mu h^{ab} \sqrt{-\det h} \right) + \text{surface term}\,. \end{aligned} \tag{2.9}$$

The equation of motion is then

$$\partial_a \left(\partial_b X^\mu h^{ab} \sqrt{-\det h} \right) = 0\,. \tag{2.10}$$

We will be interested both in strings in the form of closed loops, and open strands. We will see in the next lecture that the closed string spectrum always contains a graviton, and that every string theory must have closed strings. Open strings are present in some string theories but not others, but have proven to be vital in giving a bridge to the various dual theories.

For closed strings it is convenient to let $0 \leq \sigma \leq 2\pi$ with periodic boundary conditions

$$X^\mu(\tau, \sigma + 2\pi) = X^\mu(\tau, \sigma)\,. \tag{2.11}$$

The surface terms in the σ direction then cancel. We will not worry now about possible surface terms at $\tau \to \pm\infty$; the sources that create and destroy strings (vertex operators) will be the subject of a later lecture. For open strings we let $0 \leq \sigma \leq \pi$, so

$$\delta S_{\rm NG} = \text{eq. of motion} - \frac{1}{2\pi\alpha'} \int d\tau \; \delta X_\mu \partial_b X^\mu h^{\sigma b} \sqrt{-\det h} \Big|_{\sigma=0}^{\sigma=\pi} . \tag{2.12}$$

Letting X^μ vary freely gives the boundary condition

$$h^{\sigma b} \partial_b X^\mu = 0 \,, \quad \sigma = 0, \pi \,. \tag{2.13}$$

This is Lorentz and translation invariant. In the next chapter we will mention other possibilities.

Conformal gauge

The equation of motion is quite nonlinear, but simplifies greatly in the right coordinate system. Assume that we can choose the parameterization τ, σ such that

$$\begin{aligned} \partial_\tau X^\mu \partial_\sigma X_\mu &= 0 \,, \\ \partial_\tau X^\mu \partial_\tau X_\mu &= -\partial_\sigma X^\mu \partial_\sigma X_\mu \,, \end{aligned} \tag{2.14}$$

(conformal gauge). This is two conditions on two coordinates, so is reasonable; we will revisit the meaning of these conditions below. The conditions (2.14) imply that

$$h_{ab} = \eta_{ab} \sqrt{-\det h} \,, \tag{2.15}$$

where $\eta_{ab} = diag(-1, 1)$. The equation of motion reduces to the massless wave equation

$$\partial_\tau^2 X^\mu = \partial_\sigma^2 X^\mu \,, \tag{2.16}$$

much like a (linearized) violin string. Also, the open string boundary condition becomes the Neumann condition $\partial_\sigma X^\mu = 0$ at $\sigma = 0, \pi$.

The general solution is a sum of left- and right-moving waves,

$$X^\mu(\tau, \sigma) = f^\mu(\tau + \sigma) + g^\mu(\tau - \sigma) \,. \tag{2.17}$$

We can then expand the σ dependence in a complete set. For the closed string this is $e^{in\sigma}$, giving

$$X^\mu(\tau, \sigma) = x^\mu + v^\mu \tau + i \sqrt{\frac{\alpha'}{2}} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{n} \left(\alpha_n^\mu e^{-in(\tau+\sigma)} + \tilde{\alpha}_n^\mu e^{-in(\tau-\sigma)} \right) . \tag{2.18}$$

This describes a string whose center of mass is moving in a straight line with velocity v^μ, and with internal excitations of amplitude $\alpha_n^\mu, \tilde{\alpha}_n^\mu$ depending on the direction of oscillation μ, harmonic n, and direction of motion along the string.

For the open string, the complete set with Neumann boundary conditions is $\cos n\sigma$. We can get to this simply by setting $\alpha_n^\mu = \tilde{\alpha}_n^\mu$ in the closed string expansion (the doubling trick),

$$X^\mu(\tau,\sigma) = x^\mu + v^\mu \tau + i\sqrt{2\alpha'} \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{1}{n} \alpha_n^\mu e^{-in\tau} \cos n\sigma \,. \tag{2.19}$$

Light-cone gauge

To properly count independent solutions we must look more closely at the gauge choice. We can rewrite it as

$$\begin{aligned} (\partial_\tau X^\mu + \partial_\sigma X^\mu)(\partial_\tau X_\mu + \partial_\sigma X_\mu) &= 0\,, \\ (\partial_\tau X^\mu - \partial_\sigma X^\mu)(\partial_\tau X_\mu - \partial_\sigma X_\mu) &= 0\,. \end{aligned} \tag{2.20}$$

What this says is that if we consider a curve on the world-sheet along which $\sigma = \tau+$constant, or one such that $\sigma = -\tau+$constant, the tangent vectors to these are null in the spacetime metric. In other words, we can trace out the null curves on any world-sheet, and assign a value of $\tau + \sigma$ to each left-moving curve and a value of $\tau - \sigma$ to each right-moving curve, and we are in conformal gauge.

But this leaves us the freedom to assign different values, so that $\tau' + \sigma'$ is any (monotonic) function of $\tau + \sigma$, and $\tau' - \sigma'$ is any monotonic function of $\tau - \sigma$. In other words, the general coordinate freedom is two functions of two variables, but after imposing conformal gauge we still have two functions each of one variable to specify.

There is a way to fully fix the coordinates that has some nice properties, but is noncovariant. We have to single out x^0 and x^1 coordinate directions, and define

$$x^\pm = \frac{x^0 \pm x^1}{\sqrt{2}}\,. \tag{2.21}$$

Also, denote the $D-2$ remaining transverse directions x^i, $i = 2, 3, \ldots, D-1$. Note that the metric is $ds^2 = -2dx^+dx^- + dx^i dx^i$, and that $x^\mu y_\mu = -x^+y^- - x^-y^+ + x^i y^i$. According to the equation of motion, $X^+(\tau,\sigma) = f^+(\tau+\sigma) + g^+(\tau-\sigma)$. Now, defining

$$\tau' + \sigma' = kf^+(\tau+\sigma)\,, \quad \tau' - \sigma' = kg^+(\tau-\sigma) \tag{2.22}$$

for some constant k, we get $X^+ = 2\tau'/k$. So we drop the prime and impose this as a final condition on our coordinates,

$$X^+ = 2\tau/k\,. \tag{2.23}$$

In other words, the world-sheet time is proportional to the spacetime x^+ coordinate. We are not free to get rid of the constant k (this would require rescaling σ along with τ and not agree with our chosen ranges for σ). The conformal gauge conditions then become

$$\partial_\tau X^i \partial_\sigma X^i = 2\partial_\sigma X^- / k \,, \tag{2.24}$$

$$\partial_\tau X^i \partial_\tau X^i + \partial_\sigma X^i \partial_\sigma X^i = 4\partial_\tau X^- / k \,. \tag{2.25}$$

Now, the oscillators $\alpha_n^+, \tilde{\alpha}_n^+$ are set to zero by the light cone condition (2.23), while the oscillators $\alpha_n^-, \tilde{\alpha}_n^-$ are determined by the conditions (2.24,2.25) in terms of the transverse $\alpha_n^i, \tilde{\alpha}_n^i$. This counting is as we expect: there are only $D-2$ physical oscillations, because oscillations tangent to the string world-sheet (timelike and longitudinal) are just oscillations of the coordinate system.

Mass shell condition

A final classical result that we need is the relation between the mass of a string and its internal state of vibration. We will do the closed string first. The momentum densities on the string are

$$P_\mu = \frac{\partial L}{\partial(\partial_\tau X^\mu)} = \frac{1}{2\pi\alpha'} \partial_\tau X_\mu \,. \tag{2.26}$$

The total momenta are then

$$p^\mu = \frac{1}{2\pi\alpha'} \int_0^{2\pi} \partial_\tau X^\mu = \frac{1}{\alpha'} v^\mu \,. \tag{2.27}$$

For $\mu = +$, the light cone condition (2.23) gives

$$p^+ = \frac{2}{\alpha' k} \,. \tag{2.28}$$

The condition (2.25) gives

$$p^- = \frac{\alpha' k}{4} p^i p^i + \frac{k}{2} \sum_{n=1}^{\infty} \sum_{i=2}^{D-1} (|\alpha_n^i|^2 + |\tilde{\alpha}_n^i|^2) \,. \tag{2.29}$$

Combining these we get

$$2p^+ p^- - p^i p^i = \frac{2}{\alpha'} \sum_{n=1}^{\infty} (|\alpha_n^i|^2 + |\tilde{\alpha}_n^i|^2) \,. \tag{2.30}$$

The left hand side is $-p^\mu p_\mu$, which is the mass-squared of the string, and so we obtain a covariant result for this, in terms of the internal state of excitation of the string. For the closed string there is one constraint in this state. Integrating Eq. (2.24) from 0 to 2π, the RHS vanishes by the periodicity of X^-. On the LHS this gives

$$\sum_{n=1}^{\infty} |\alpha_n^i|^2 = \sum_{n=1}^{\infty} |\tilde{\alpha}_n^i|^2 \,, \tag{2.31}$$

so the total excitation levels of the left-movers and right-movers are equal.

For the open string we end up with

$$M^2 = 2p^+p^- - p^ip^i = \frac{1}{\alpha'} \sum_{n=1}^{\infty} \sum_{i=2}^{D-1} |\alpha_n^i|^2 \,. \tag{2.32}$$

3 Quantization in light-cone gauge

The main thing we expect from quantization is that the excitation levels just discussed will be quantized, so that the possible masses are discrete. This is true, but it will come with some surprises.

Canonical quantization

We impose the equal-time canonical commutators[3]

$$\begin{aligned} [X^i(\tau,\sigma), X^j(\tau,\sigma')] &= [P^i(\tau,\sigma), P^j(\tau,\sigma')] = 0 \,, \\ [X^i(\tau,\sigma), P^j(\tau,\sigma')] &= i\delta^{ij}\delta(\sigma-\sigma') \,. \end{aligned} \tag{3.1}$$

Using the result (2.26) for the momentum density, inserting the mode expansions (2.18, 2.19) for the transverse components, and Fourier transforming to project out the individual modes, we get

$$\begin{aligned} [x^i(\tau), p^j(\tau)] &= i\delta^{ij} \,, \\ [\alpha_m^i, \alpha_n^j] = [\tilde{\alpha}_m^i, \tilde{\alpha}_n^j] &= m\delta^{ij}\delta_{m+n,0} \,. \end{aligned} \tag{3.2}$$

This is what we expect in free field theory, that each mode gives an oscillator algebra. We have $D-2$ free massless fields, labeled by i, on the 1+1 dimensional world-sheet. The normalization is perhaps not the usual: for $m > 0$, α_m^i (and $\tilde{\alpha}_m^i$ in the closed string) are

[3]This simple derivation sets off red flags for me, because one has used the equation of motion before quantizing, which in the path-integral is a no-no. So in the Big Book I showed that by being careful about the order of steps it all worked in the path integral. That was for the Polyakov form of the action, which we haven't seen yet, but similar steps can be applied to the Nambu-Goto action.

$\sqrt{m}$ times a standard lowering operator, and α^i_{-m} and $\tilde\alpha^i_{-m}$ are $\sqrt{m}$ times the corresponding raising operator. Also

$$\alpha^i_{-m}\alpha^i_m = mN^i_m \tag{3.3}$$

is m times the number operator for the mode of given i and m. The state of minimum excitation for the quantized string, which we'll label $|0,k\rangle$ is annihilated by all the lowering operators,

$$\alpha^i_n|0,k\rangle = 0 \text{ for all } n > 0 \text{ and all } i\,, \tag{3.4}$$

and the same for $\tilde\alpha^i_n$ in the closed string. Also, we must specify the momenta p^i and p^+, and then p^- is determined by Eq. (2.30, 2.32), whose translation into the quantum theory we discuss below. (Looking back through the notes, I seem to have used k^μ and p^μ interchangably).

We can then form excited states by acting in a general way with the raising operators. In the closed string there is the 'level-matching' constraint (2.31), that

$$\sum_{m,i} mN^i_m = \sum_{m,i} m\tilde N^i_m\,. \tag{3.5}$$

Zero point energy

To get the mass of these string states we must quantize the relation (2.30, 2.32). Here we will do this in a hand-waving way, and reproduce the same results more systematically later. It is a general result in free field theory that the operator ordering comes out averaged,

$$|\alpha^i_m|^2 \to \frac{1}{2}\left(\alpha^i_{-m}\alpha^i_m + \alpha^i_m\alpha^i_{-m}\right) = \alpha^i_{-m}\alpha^i_m + \frac{m}{2} = m\left(N^i_m + \frac{1}{2}\right)\,; \tag{3.6}$$

this is familiar for the zero point energy for a harmonic oscillator. Then

$$M^2 = \frac{1}{\alpha'}\sum_{m=1}^{\infty}\sum_{i=2}^{D-1} m\left(N^i_m + \frac{1}{2}\right) \tag{3.7}$$

for the open string, and correspondingly for the closed.

For the ground state, all $N^i_m = 0$ and

$$M_0^2 = \frac{D-2}{2\alpha'}\sum_{n=1}^{\infty} n\,. \tag{3.8}$$

The sum is badly divergent, so what has happened to this wonderful theory that we were supposed to have?

As an aside, this sum over zero point energies comes up in many interesting places. One is the Casimir energy for fields in a cavity; actually, this is precisely what are looking at here, the 'cavity' being the finite-size string. Another is in calculating the Higgs potential, which is affected by the zero point energies of the fields to which the Higgs couples; this is known as the Coleman-Weinberg potential. In all these cases there is some way to 'renormalize' the sum and get the correct finite answer. In fact, the answer for the string is only infinite because we have not preserved the symmetries carefully, and the symmetries determine the finite result. This is explained a bit around 1.3.34 of the text, but for now we will wave our hands, and later derive the same result without infinities.

Hand-wave #1: define the ζ function

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} \tag{3.9}$$

This converges for $s > 1$ and has a pole at $s = 1$. It can be continued around the pole, and $\zeta(-1) = -\frac{1}{12}$. So this is the value we assign to the $\sum_{n=1}^{\infty} n$, by analogy to dimensional regularization.

Hand-wave #2: regulate

$$\sum_{n=1}^{\infty} n e^{-\epsilon n} = \frac{e^{-\epsilon}}{(1 - e^{-\epsilon})^2} = \frac{1}{\epsilon^2} - \frac{1}{12} + O(\epsilon^2) \tag{3.10}$$

In the book I argue that symmetry requires that we drop exactly the $1/\epsilon^2$ term, giving the same answer.[4]

So

$$M_0^2 = \frac{2 - D}{24\alpha'} \tag{3.11}$$

for the open string, and 4 times this for the closed string.

The tachyon

It seems that we are out of the frying pan and into the fire. We have handwaved the divergence away, but now our ground state string has a *negative* mass-squared as long as $D > 2$. The potential energy density for a scalar field is $\frac{1}{2}M^2\phi^2$, so our would be vacuum is unstable. This is true: we don't get a stable vacuum if we just quantize strings. The missing ingredient is supersymmetry. In a supersymmetric theory, the Hamiltonian is the square

[4] I should emphasize that I am not appealing to some general theory of summing divergent series here, but rather this is what one gets from regulating the path integral along the lines of Pauli-Villars, and subtracting a counterterm because the regulator breaks conformal invariance.

of the supercharge, $H = Q^2$, and so non-negative (I'm leaving out nonrelevant subtleties here). This forbids the kind of unbounded below potential we've just found, and there is no tachyon.

However, rather than trying to teach you supersymmetry and string theory all at once, we will study this unstable theory, sort of by analytic continuation from positive mass-squared. The instability does not affect the tree amplitudes at all, and affects the loop amplitudes only in a rather subtle way.

As an aside, for open string tachyons we can in some cases understand the final state of the instability, often a state with closed strings only. When there are closed strings tachyons, then in general there is no final state, spacetime itself decays.

The photon and the critical dimension

For general open string states, we now have

$$M^2 = \frac{2-D}{24\alpha'} + \frac{1}{\alpha'}\sum_{n=1}^{\infty}\sum_{i=2}^{D-1} nN^i_m \,. \tag{3.12}$$

In the lowest excited states, we excite one of the $D-2$ oscillators with $m = 1$,

$$\alpha^i_{-1}|0,k\rangle\,, \quad M_1^2 = \frac{2-D}{24\alpha'} + \frac{1}{\alpha'} = \frac{26-D}{24\alpha'}\,. \tag{3.13}$$

Now, this state has a vector index. In the rest frame, there should be $D-1$ states, because the vector can point in any spatial direction. But we have only $D-2$ states, so Lorentz invariance has broken down. Again, our light-cone quantization singles out the 0- and 1-directions. Only the rotional invariance among the directions $2, \ldots, D-1$ is manifest, and so we need to check the rest, and it fails.

There is one out. If $M^2 = 0$, there is no rest frame, and rotations around the direction of motion generate only $D-2$ states (e.g. in $D = 4$ the Z^0 has three states but the photon two), so we're OK. But this means that Lorentz invariance can only hold if

$$M_1 = 0 \implies D = 26\,. \tag{3.14}$$

A rather shaky thread of logic has led us to the conclusion that the string can only be quantized in 26 dimensions. But the result is robust, and survives more thorough scrutiny. For example, GGRT calculate the algebra of the Lorentz generators, and it comes out right only in $D = 26$. We will understand it from a more systematic point of view later.

This sort of thing is not unprecedented. The Standard Model has six quarks and six leptons. If it had six of one and four of the other, the classical theory would have all the expected symmetries, but quantization would lead to an anomaly. In the quantum theory, one of Lorentz invariance, conservation of probability, or renormalizability would break down (you can choose which to give up).

Because our theory has gravity (as we will see shortly), the extra dimensions can be highly curved. Observation only tells us that there are four dimensions large enough to detect.

Finding a massless vector is an interesting result. General principles (Weinberg) imply that it must couple like a gauge boson, and we will find this when we study the amplitudes.

With $D = 26$, the tachyon mass is $M_0^2 = -1/\alpha'$. The lowest states of positive mass-squared are obtained by acting with an $m = -2$ operator or two $m = -1$ operators,

$$\alpha^i_{-2}|0, k\rangle \ , \ \alpha^i_{-1}\alpha^j_{-1}|0, k\rangle \ , \quad M_2^2 = \frac{1}{\alpha'} \, . \tag{3.15}$$

One can verify that these do form complete multiplets of $SO(25)$. Let the index I run from 1 to 25, and consider a set of states that transform as a symmetric traceless matrix M_{IJ}. The components M_{ij} transform like $\alpha^i_{-1}\alpha^j_{-1}|0, k\rangle$, the components $M_{1i} = M_{i1}$ transform like $\alpha^i_{-2}|0, k\rangle$, and the component M_{11} can be expressed in terms of the others by the traceless condition. The same holds at all higher levels — this is implied by the result of GGRT. The higher levels are spaced by multiples of $1/\alpha'$.

Graviton, dilaton, axion

The analysis of the closed string spectrum is similar, with two sets of oscillators. The ground state is

$$\alpha^i_n|0, k\rangle = \tilde{\alpha}^i_n|0, k\rangle = 0 \text{ for all } n > 0 \text{ and all } i \, , \quad M_0^2 = \frac{2 - D}{6\alpha'} \, . \tag{3.16}$$

The first excited levels are

$$\alpha^i_{-1}\tilde{\alpha}^j_{-1}|0, k\rangle \ , \quad M_1^2 = \frac{26 - D}{6\alpha'} \, . \tag{3.17}$$

Notice that we have to excite both a right- and left-moving oscillator, by the condition (3.5). The same argument as before requires these state to be massless, so again $D = 26$.

The states (3.17) actually decompose into several different particles. Contract with a polarization tensor, $e_{ij}\alpha^i_{-1}\tilde{\alpha}^j_{-1}|0, k\rangle$. The state with $e_{ij} = \delta_{ij}$ is invariant under rotations; it is the massless scalar dilaton. These states where e_{ij} is symmetric and traceless also mix only among themselves; this is the spin of the graviton, and again general principles and direct calculation lead to the interactions of general relativity. Finally, the states where e_{ij} is antisymmetric mix only among themselves.

The massless string states play a big role in the theory: if we don't have enough energy to excite the massive strings, or equivalently if we are interested in length scales much greater than $\sqrt{\alpha'}$ then they are all we see. Further, we can describe them by an effective field theory. The graviton is described by a metric field $G_{\mu\nu}$, the dilaton by a scalar field Φ, the antisymmetric tensor by an antisymmetric tensor field (2-form) $B_{\mu\nu}$ (which has a gauge invariance, which we will discuss later). Also, the open string vector is described by a gauge field A_μ.

You don't usually encounter a two-form in QFT or GR. In D dimensions, with transverse indices taking $D-2$ values, the symmetric traceless tensor has $\frac{1}{2}(D-1)(D-2)-1$ states, which is indeed two in $D=4$. An antisymmetric tensor has $\frac{1}{2}(D-2)(D-3)$ states, which is just one in $D=4$. In fact, in four dimensions a two-form gauge field can be rewritten as a massless scalar. Actually, this two-form has the right properties to be a dark matter candidate, the axion, though not one that is easily seen at the LHC or via direct detection.[5]

The first massive level, $M_2^2 = 4/\alpha'$, is the tensor product of two copies of the open string spectrum (3.15), one from the left-movers and one from the right-. This product form repeats at all higher levels as well.

D-branes, etc.

We will see later on that all consistent string theories have closed strings (and the graviton) but that some have open strings of the above sort and others do not. But there are still more possibilities.

Coming back to the surface term in the equation of motion, $\delta X^\mu \partial_\sigma X_\mu = 0$, another way to satisfy it would be to constrain the variations by imposing $X^\mu = 0$ at the endpoints, a Dirichlet boundary condition. So the string endpoints are stuck at the origin in space and time. We could even mix boundary conditions, for example taking Neumann for some values of μ and Dirichlet for others. Let us consider the case $\mu = 0, 1, \ldots, p$ being Dirichlet, and the rest Neumann, we can think of the string endpoints as stuck on a flat p-dimensional hyperplane in space. The mode expansion is then

$$\begin{aligned} X^\mu(\tau,\sigma) &= x^\mu + 2\alpha' p^\mu \tau + i\sqrt{2\alpha'} \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{1}{n} \alpha_n^\mu e^{-in\tau} \cos n\sigma\,, \quad \mu \leq p\,, \\ &= \sqrt{2\alpha'} \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{1}{n} \alpha_n^\mu e^{-in\tau} \sin n\sigma\,, \quad \mu \geq p+1\,. \end{aligned} \tag{3.18}$$

[5]The dark matter particle is massive. The 'axion' is massless to the order that we are working, but symmetry breaking effects will give it a mass.

Notice that there is no term corresponding to linear motion in the Dirichlet directions, because the endpoints are fixed.

This boundary condition breaks Lorentz invariance and translation invariance. This sounds bad, but it's not! For spacetime itself, it starts flat, but once we have the gravitons we know that its geometry is allowed to fluctuate. It is the same for these hyperplanes, but where the gravitons come from closed strings, the hyperplane fluctuations come from the open strings that are stuck to them. At the massless level, the states $\alpha^i_{-1}|0,k\rangle$ for $i \leq p$ have polarizations parallel to the plane and can be thought of as a gauge field living on it. But what of the states with $i \geq p+1$? Their polarization is perpendicular to the plane, and it turns out that they correspond to oscillations of the plane. Thus it is a dynamical object, a Dirichlet p-membranes, or Dp-branes. Translation invariance, and Lorentz invariance, are restored, because different locations and orientations for the D-brane are just different states in the theory.

We will explain later, in the context of T-duality, why such D-branes must actually appear in the theory. As another comment, if you put a lot of D-branes on top of one another (and I will have to explain later how to quantify 'a lot' in terms of the string coupling constant) then they warp spacetime and develop a horizon, and we call the result a black brane. By varying the string coupling adiabatically we can go back and forth between the D-brane and black brane descriptions: this is the trick behind the Strominger-Vafa entropy counting, and ultimately it leads to AdS/CFT duality.

For uniformity, we can think of a Neumann condition as a string ending on a D25-brane. That is, the brane is extended in all 25 dimensions of space, as well as time, so the endpoints can be anywhere. But now suppose we have multiple D25-branes on top on one another. The open strings then have an index for each end, which indicates which brane it ends on, for example

$$|0,k,I,J\rangle \tag{3.19}$$

This is known as a Chan-Paton index. (In the book I use i, j rather than I, J for these — not to be confused with the use of i, j for the transverse dimensions, I just ran out of letters.) So each open string state has two extra indices labeling which D25-branes it starts and ends on: we get n^2 copies of the open string spectrum. In particular, the massless vector

$$\alpha^i_{-1}|0,k,I,J\rangle \tag{3.20}$$

has the interactions of a non-Abelian $U(n)$ gauge field. Dp-branes for $p < 25$ also have gauge fields living on them. It is suggestive that this might be the origin of the $SU(3) \times SU(2) \times U(1)$ gauge symmetry of nature, though to make a realistic model things get a little more complicated.

[Illustrations: D-branes of various dimensions, and strings stretched between them]

Another generalization is unoriented strings. For the strings thus far we can imagine an arrow that gives them a direction. For an *unoriented* string we must coherently superpose the string and reverse string in every amplitude, for example

$$\alpha^i_{-1}\tilde{\alpha}^j_{-1}|0,k\rangle + \alpha^j_{-1}\tilde{\alpha}^i_{-1}|0,k\rangle \tag{3.21}$$

and this has the effect of projecting onto states that are *even* under the world-sheet parity operation $\sigma \to -\sigma$. For the closed string the graviton, dilaton, and tachyon survive the projection but the two-form does not. We will not mention the unoriented strings much, we don't need the added complication, but we will encounter them again when we list the five consistent string theories in the final lecture.

4 The Polyakov path integral

If we sum over all world-sheets weighted by e^{iS}, we automatically generate interactions - see the pictures and discussion in Ch. 3.1. In fact, this is the only consistent way to introduce interactions, and the result is unique. If we include only world-sheets without boundaries, we get only closed strings. If we allow world-sheets with boundary, we get both closed and open strings. The closed strings are always there.

[Illustrations: open string tree level scattering, and its time slicing; same for closed string; open strings exchanging a closed string. The latter graph is topologically a long cylinder, with on ingoing and one outgoing string attached to each end. In the limit that the cylinder is short, it is like an annulus (one loop open string graph), so is automatically part of the quantum theory.]

The Polyakov action

The square root form of the Nambu-Goto action is messier than anything you saw in QFT, and it is difficult to use in a path integral. So we will first find a nicer action that gives the same classical theory. Let's first return to the point particle action,

$$S_{rel'} = -m\int d\tau\sqrt{-\partial_\tau X^\mu \partial_\tau X_\mu}\,, \tag{4.1}$$

which also has the annoying square root. Consider instead

$$S_{rel''} = \frac{1}{2}\int d\tau\left(\eta^{-1}\partial_\tau X^\mu\partial_\tau X_\mu - \eta m^2\right)\,, \tag{4.2}$$

where we have introduced a new field $\eta(\tau)$ that has no time derivative. The variation of action with respect to η is

$$\delta S_{rel''} = \frac{1}{2} \int d\tau \delta\eta \left(-\eta^{-2} \partial_\tau X^\mu \partial_\tau X_\mu - m^2 \right) . \tag{4.3}$$

For this to be stationary for arbitrary $\delta\eta(\tau)$, we get

$$\eta^2 = -\partial_\tau X^\mu \partial_\tau X_\mu / m^2 . \tag{4.4}$$

Solving this (take the positive root) and inserting back into the action gives the previous $S_{rel'}$, so we then get the original equations of motion when we vary X^μ. So this has the same classical solutions; this trick of introducing an extra field without derivatives to get a simpler action is common.

And the final action is simpler. First, it's quadratic in X^μ, so the path integral is gaussian just like in free field theory. But what about η? Notice that η and $d\tau$ appear only in the combination $\eta d\tau$. If we make a coordinate transformation, this combination is invariant, $\eta d\tau = \eta' d\tau'$. So we can simply choose $\eta' = 1$ everywhere, and this specifies τ'. So we go to the coordinate system $\eta = 1$ and then the action is just a massless Klein-Gordon action in one dimension, and we have a gaussian path integral which is easy. (There is also a determinant for the gauge fixing, of a type that we will deal with later for the string.)

The field η has a simple interpretation. Suppose we introduced a metric $\gamma_{\tau\tau}$ along the string. (I use γ so as to save g for later.) Then $\gamma'_{\tau\tau} d\tau'^2 = \gamma_{\tau\tau} d\tau^2$ is how a metric transforms. So η is basically $\sqrt{-\gamma_{\tau\tau}}$, where the minus sign is included because of my convention that timelike metrics are negative. In other words, it's like a vierbein, but there's just one so it's an einbein.

Based on this example we introduce an independent metric field $\gamma_{ab}(\tau, \sigma)$ on the string world-sheet. This has the same indices as h_{ab}, but that was constructed entirely from the X^μ's. Our action is

$$S_P = -\frac{1}{4\pi\alpha'} \int d\tau d\sigma \sqrt{-\det\gamma}\, \gamma^{ab} \partial_a X^\mu \partial_b X_\mu = -\frac{1}{4\pi\alpha'} \int d\tau d\sigma \sqrt{-\det\gamma}\, \gamma^{ab} h_{ab} , \tag{4.5}$$

where the P is short for Brink-di Vecchia-Howe-Deser-Zumino. You might have also expected a term $\mu\sqrt{-\det\gamma}$, by analogy with the second term in $S_{rel''}$ but this actually wouldn't work, you would get a sick field equation for γ_{ab} as you can check. Now, varying γ_{ab},

$$\delta S_P = -\frac{1}{4\pi\alpha'} \int d\tau d\sigma \sqrt{-\det\gamma}\, (\delta\gamma^{ab} h_{ab} - \tfrac{1}{2} \delta\gamma^{ab} \gamma_{ab} \gamma^{cd} h_{cd}) . \tag{4.6}$$

The first term is from the variation of the inverse metric and the second from the variation of the determinant, where I've used $\mathrm{Tr}(\gamma^{-1}\delta\gamma) = -\mathrm{Tr}(\delta\gamma^{-1}\gamma)$, which follows because $\mathrm{Tr}(\gamma^{-1}\gamma)$ is a constant so its variation is zero.

So our field equation is

$$h_{ab} = \tfrac{1}{2}\gamma^{cd}h_{cd}\gamma_{ab}\,. \tag{4.7}$$

The determinant of this is

$$\det h = \tfrac{1}{4}(\gamma^{cd}h_{cd})^2 \det\gamma\,, \tag{4.8}$$

the exponent 2 because these are 2×2 matrices. Taking the square root of minus this equation, S_P becomes S_{NG}. So once again we can make our world-sheet action into a Klein-Gordon action for the fields X^μ, minimally coupled to a world-sheet metric.

It might occur to you to add an Einstein-Hilbert term $\sqrt{-\gamma}R$ with R the Ricci scalar built from γ_{ab}, but this is locally trivial, it is a topological invariant of the world-sheet.

However, there is a new feature. The equation of motion (4.7) doesn't fully determine γ_{ab}, because if we multiply a solution $\gamma_{ab}(\tau,\sigma)$ by any function $e^{2\omega(\tau,\sigma)}$, it is still a solution. In fact, the action also has this symmetry. So every physical motion of the string has many different realizations in terms of the fields $X^\mu(\tau,\sigma)$, $\gamma_{ab}(\tau,\sigma)$ — first, by choosing different coordinate systems, and second by rescaling the metric locally. This rescaling symmetry is generally called Weyl invariance in the string literature; in the GR literature it often referred to as conformal invariance, but we will reserve this for something slightly different though related. Weyl invariance is not a symmetry of GR in $D = 4$ (or any dimension greater than 2), though there is a more complicated theory with this symmetry, one which is unphysical but perhaps of some theoretical interest.

We've already dealt with the idea that the coordinate invariance is the price we pay for a covariant description, and the Weyl invariance is the price we pay for a nice action. It has another benefit as well. The anomaly in Lorentz invariance that we found in $D \neq 26$ ultimately comes from an anomaly in coordinate invariance (because the Lorentz frame entered through our choice of coordinates). We will have to worry about other possible anomalies later. Anomalies in coordinate invariance are messy and inconvenient, and what we can do is to keep coordinate invariance clean and shuffle any anomaly into the Weyl invariance.

In summary, the symmetries of the Polyakov action are Poincaré (Lorentz plus translations), and world-sheet coordinate and Weyl. The latter two are local symmetries, depending on τ,σ, and the are (like all other gauge symmetries) redundancies of description: nothing physical depends on them.

Euclidean world-sheets and conformal gauge

In quantum theory we are interested in transition amplitudes like

$$\langle 1|e^{-iHt}|2\rangle\,,\ t > 0\,. \tag{4.9}$$

Consider the analytic continuation $t \to -it_E$, $t_E > 0$. This becomes

$$\langle 1|e^{-Ht_E}|2\rangle \,. \tag{4.10}$$

Now, the eigenvalues of H are bounded below for a stable system, so this is well-defined: the transition amplitudes have a good continuation to Euclidean time, and we can evaluate them there and continue back. The path integral for the Euclidean amplitudes is usually better defined. For example, it is what lattice gauge theorists calculate numerically. We will do the same, starting with the Euclidean path integral and giving a rule later for the continuation.

What this means is that we have to replaces the metric γ_{ab}, which had signature $(-1, 1)$ like h_{ab}, with a metric g_{ab} of Euclidean signature. So the action is

$$S_P \to \frac{1}{4\pi\alpha'} \int d^2\sigma \, \sqrt{\det g}\, g^{ab} \partial_a X^\mu \partial_b X_\mu \,. \tag{4.11}$$

The overall sign flip of S is due to inconsistent conventions in the literature, but with this sign the path integral is weighted by e^{-S}. We will label the world-sheet coordinates σ^1, σ^2. We will keep the spacetime metric Lorentian, but we will have to make a rule for integrating over X^0.

The first step is to fix the gauge. The metric has three independent components (2×2 symmetric) and we have three choices to make (two coordinates and the Weyl scale). So it seems that we can set the metric to a convenient value,

$$g_{ab} = \delta_{ab} \,. \tag{4.12}$$

In fact we can always do this at least locally on the world-sheet; I won't go through this but it's done around 3.3.5 of the book.[6]

In this unit gauge the action is

$$S_P = \frac{1}{4\pi\alpha'} \int d^2\sigma \, (\partial_1 X^\mu \partial_1 X_\mu + \partial_2 X^\mu \partial_2 X_\mu) \,. \tag{4.13}$$

When you look at X^μ you should think ϕ. That is, it is a set of scalar field in the two dimensions of the world sheet, with the index μ labeling the different fields. Moreover, this is just a free massless field theory again.

Thus, it's straightforward to quantize canonically as before, but I will wait until introducing a new notation below. In QFT, most of the work is in introducing interactions, but we won't be doing that (we would have them if spacetime were curved, but we are going to

[6] Up to here we've mostly been on chap. 1, now we are moving on to chap. 2 but I am pulling a few ideas from further on.

use the flat spacetime as our example). What we are going to do that is nontrivial and new is to focus on a certain symmetry, conformal invariance, and its generators in the quantum theory, and also on the properties of composite operators. These are both things that have importance beyond string theory, so you will learn also some interesting parts of QFT.

Complex coordinates and conformal transformations

There is something that is very useful to do here, which is to collect our two real coordinates into one complex coordinate,

$$w = \sigma^1 + i\sigma^2 \,, \quad \bar{w} = \sigma^1 - i\sigma^2 \,. \tag{4.14}$$

For the closed string, which σ^1 is periodic with period 2π, so it is also useful to define $z = e^{-iw}$. We also define

$$\partial_w = \tfrac{1}{2}(\partial_1 - i\partial_2) \,, \quad \partial_{\bar{w}} = \tfrac{1}{2}(\partial_1 + i\partial_2) \,. \tag{4.15}$$

These satisfy $\partial_w w = 1$, $\partial_w \bar{w} = 0$, and so on.

In terms of these,

$$S_P = \frac{1}{2\pi\alpha'} \int d^2w \, \partial_w X^\mu \partial_{\bar{w}} X_\mu \,, \tag{4.16}$$

where I define $d^2w = 2d\sigma^1 d\sigma^2$.

In these variables, canonical quantization gives

$$X^\mu = x^\mu - i\frac{\alpha'}{2}p^\mu \ln|z^2| + i\sqrt{\frac{\alpha'}{2}} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{1}{m}\left(\frac{\alpha_m^\mu}{z^m} + \frac{\tilde{\alpha}_m^\mu}{\bar{z}^m}\right) \,, \tag{4.17}$$

where we have abbreviated $z = e^{-iw} = e^{-i\sigma_1+\sigma_2} \rightarrow e^{-i(\tau-\sigma)}$. The commutators are the canonical commutators

$$[\alpha_m^\mu, \alpha_n^\nu] = [\tilde{\alpha}_m^\mu, \tilde{\alpha}_n^\nu] = m\eta^{\mu\nu}\delta_{m+n,0} \,. \tag{4.18}$$

This is like before, except that we have D sets of oscillators. Our Hilbert space is too big (like having four polarizations in a covariant description of the photon), and we will have to reduce it in a covariant way later. By the way, you can replace σ^2 by $i\sigma^0$ to get back to the more familiar Lorentzian time.

Now, this action has an important symmetry. We are in the gauge

$$ds^2 = dw d\bar{w} \,. \tag{4.19}$$

Suppose that we adopt new coordinates such that $w = f(w')$. That is, it is a *holomorphic* function, not a general one. Then

$$ds^2 = \frac{\partial w}{\partial w'} \frac{\partial \bar{w}}{\partial \bar{w}'} dw' d\bar{w}' \,, \tag{4.20}$$

with no $dw'dw'$ or $d\bar{w}'d\bar{w}'$. But now by a Weyl transformation we can bring this to $dw'd\bar{w}'$, the same as we started with. So there is a combination of a holomorphic coordinate change and a Weyl transformation that leaves us in this gauge, once again we have failed to fully fix it. Before we introduced an additional noncovariant condition, but now we will deal with it differently. This symmetry will also be there for strings in curved spacetime, it plays a central role.

The upshot is that the action (4.16), even after the metric field has been fixed, has a symmetry under *conformal transformations*. These are holomorphic changes of coordinate. In terms of the real coordinates σ^1, σ^2, these locally preserve angles and ratios of lengths.

The conformal invariance of the action (4.16) is 'obvious,' the Jacobian from d^2w canceling those from ∂_w and $\partial_{\bar{w}}$, but let's go through the pedantic exercise of verifying it. Let

$$X^\mu(w, \bar{w}) \to X^\mu(w'(w), \bar{w}'(\bar{w})) \,. \tag{4.21}$$

Then

$$\begin{aligned}
S_P &= \frac{1}{2\pi\alpha'} \int d^2w \, \partial_w X^\mu(w', \bar{w}') \partial_{\bar{w}} X_\mu(w', \bar{w}') \,, \\
&= \frac{1}{2\pi\alpha'} \int d^2w \, \partial_w w' \partial_{\bar{w}} \bar{w}' \partial_{w'} X^\mu(w', \bar{w}') \partial_{\bar{w}'} X_\mu(w', \bar{w}') \,, \\
&= \frac{1}{2\pi\alpha'} \int d^2w' \, \partial_{w'} X^\mu(w', \bar{w}') \partial_{\bar{w}'} X_\mu(w', \bar{w}') \,, \\
&= \frac{1}{2\pi\alpha'} \int d^2w \, \partial_w X^\mu(w, \bar{w}) \partial_{\bar{w}} X_\mu(w, \bar{w}) \,.
\end{aligned} \tag{4.22}$$

In the last line we have just renamed the variable of integration.

5 The Virasoro Algebra

The Virasoro generators

Outline: basis of infinitesimal symmetries, invariance of action, conserved currents and charges, mode expansion, algebra of charges.

A general principle in mechanics is that associated with every symmetry is a conserved quantity, Noether's theorem. This is worked out in chap. 22 of Srednicki, and in a slightly different way in chap. 2.3 of the text. Srednicki eq. 22.27, Noether's theorem in general form, is the same as eq. 2.3.4, even though they look different. I will use the result in the book, not going through the derivation, but then verifying that it gives a conserved current.

First, we want to consider the infinitesimal transformations. We will work in the z coordinate which makes the closed string periodicity manifest. A general holomorphic transformation is

$$z' = z + \sum_{n=-\infty}^{\infty} \epsilon_n z^{n+1}\,, \tag{5.1}$$

with an infinite number of parameters ϵ_n. The transformation of $\bar{z}$ is the conjugate of this. The transformation (4.21) is

$$X^\mu(z,\bar{z}) \to X^\mu(z,\bar{z}) + \sum_{n=-\infty}^{\infty} \epsilon_n z^{n+1}\partial_z X^\mu(z,\bar{z}) + \sum_{n=-\infty}^{\infty} \epsilon_n \bar{z}^{n+1}\partial_{\bar{z}} X^\mu(z,\bar{z})\,. \tag{5.2}$$

The variation of the Lagrangian density

$$L_P = \frac{1}{2\pi\alpha'}\partial_z X^\mu \partial_{\bar{z}} X_\mu\,, \tag{5.3}$$

for given n, is

$$\begin{aligned}
&\frac{1}{2\pi\alpha'}\left(\partial_z(\epsilon_n z^{n+1}\partial_z X^\mu)\partial_{\bar{z}}X_\mu + \partial_z X^\mu \partial_{\bar{z}}(\epsilon_n z^{n+1}\partial_z X^\mu)\right) \\
&= \frac{1}{2\pi\alpha'}\left(\partial_z(\epsilon_n z^{n+1}\partial_z X^\mu)\partial_{\bar{z}}X_\mu + \partial_z X^\mu \epsilon_n z^{n+1}\partial_z\partial_{\bar{z}} X^\mu\right) \\
&= \frac{1}{2\pi\alpha'}\partial_z\left(\partial_z X^\mu \epsilon_n z^{n+1}\partial_{\bar{z}} X^\mu\right)\,,
\end{aligned} \tag{5.4}$$

showing again that this is a symmetry (we also have the complex conjugate term). The prescription in the book to get the conserved current is to replace the constant ϵ_n with $\epsilon_n\rho(\sigma)$. The variation of the action only comes from terms with derivatives of ρ, which show up when we try to move $\bar{z}$ to the right in going from line one to line two:

$$\frac{1}{2\pi\alpha'}\partial_z X^\mu (\partial_{\bar{z}}\rho)\epsilon_n z^{n+1}\partial_z X^\mu\,. \tag{5.5}$$

The theorem is that the coefficient of $\partial_{\bar{z}}\rho$ is $\epsilon_n j_n^{\bar{z}}/2\pi i$, so

$$j_{nz} = \frac{i}{\alpha'} z^{n+1}\partial_z X^\mu \partial_z X_\mu\,. \tag{5.6}$$

There is no term of the form $\epsilon_n\partial_z\rho$ so $j_{n\bar{z}} = 0$. For a conserved current,

$$\partial_z j_{\bar{z}} + \partial_{\bar{z}} j_z = 0\,, \tag{5.7}$$

and here this reduces to $\partial_{\bar{z}} j_{nz} = 0$. It is easy to verify this because $\partial_{\bar{z}}$ kills every factor in the current j_{nz}, using the equation of motion

$$\partial_{\bar{z}}\partial_z X^\mu = 0\,. \tag{5.8}$$

Similarly the coefficient of $\bar{\epsilon}_n \partial_z \rho$ gives a current

$$j_n^z = \frac{i}{\alpha'} \bar{z}^{n+1} \partial_{\bar{z}} X^\mu \partial_{\bar{z}} X_\mu \tag{5.9}$$

with vanishing $j_n^{\bar{z}}$ component.

Let me also give another way to think about this. The world-sheet energy-momentum tensor is given by varying the metric,

$$\begin{aligned} T_{ab} &= -\frac{\sqrt{4\pi}}{\sqrt{g}} \frac{\delta}{\delta g^{ab}} S_P \\ &= -\frac{1}{\alpha'}(\partial_a X^\mu \partial_b X_\mu - \tfrac{1}{2} g_{ab} \partial^c X^\mu \partial_c X^\mu)\,, \end{aligned} \tag{5.10}$$

giving

$$T_{w\bar{w}} = 0\,, \quad T_{ww} = -\frac{1}{\alpha'} \partial_w X^\mu \partial_w X_\mu\,, \quad T_{\bar{w}\bar{w}} = -\frac{1}{\alpha'} \partial_{\bar{w}} X^\mu \partial_{\bar{w}} X_\mu\,. \tag{5.11}$$

The fact that $T_{w\bar{w}} = 0$ is no accident, it holds in any conformally invariant theory. Note that we can also write this as the vanishing of the trace of T: $g^{ab} T_{ab} = 4T_{w\bar{w}} = 0$. But then conservation of T implies that

$$0 = g^{ab}\partial_a T_{bw} = 2\partial_w T_{\bar{w}w} + 2\partial_{\bar{w}} T_{ww} = 2\partial_{\bar{w}} T_{ww}\,. \tag{5.12}$$

So T_{ww} is holomorphic, and $T_{\bar{w}\bar{w}}$ is antiholomorphic. The current j_{nw} is then basically $(\delta_n w) T_{ww}$. That is, it translates w by something proportional to e^{-inw}.

For the conserved charge, we take

$$\int_0^{2\pi} d\sigma^1\, j^0 \tag{5.13}$$

(the integral of the time component around the string). Up to some normalizing factor that I don't care about (I could have rescaled the original ϵ_n), this becomes

$$L_n = \oint_C \frac{dz}{2\pi} j_{nz} = \oint_C \frac{dz}{2\pi i z} z^{n+2} T_{zz}\,. \tag{5.14}$$

(This final expression for L_n in terms of T is standard, some of my intermediate expressions for j_{nz} may be off by an i in the normalization.) Recalling that $z = e^{-i\sigma_1 + \sigma_2}$, a constant time contour is a circle of constant radius in z, conventionally taken counterclockwise. The mode expansion (4.17) becomes

$$\partial_z X^\mu = -i\sqrt{\frac{\alpha'}{2}} \sum_{m=-\infty}^{\infty} \frac{\alpha_m^\mu}{z^{m+1}}\,, \tag{5.15}$$

where I have defined $\alpha_0^\mu = p^\mu \sqrt{\alpha'/2}$ to make things simple. So if you insert the current (5.6) and the mode expansion (5.15) for each $\partial_z X$, you get

$$L_n = \frac{1}{2} \sum_{m=-\infty}^{\infty} \alpha_{n-m}^\mu \alpha_{\mu m}\,. \tag{5.16}$$

These are the Virasoro generators, and they play a big role in the covariant quantization of the string, or in any 2-dimensional QFT with conformal invariance. Notice that for $n \neq 0$ the two mode operators commute so we don't care about their order. For $n = 0$ we will define L_0 by normal ordering, putting the lowering operator on the right,

$$L_0 = \frac{1}{2}\alpha_0^\mu \alpha_{\mu 0} + \sum_{m=1}^{\infty} \alpha_{-m}^\mu \alpha_{\mu m} = \frac{\alpha'}{4} p^\mu p_\mu + \sum_{m=1}^{\infty} \alpha_{-m}^\mu \alpha_{\mu m}\,. \tag{5.17}$$

This is just a definition, adding a constant would still leave it conserved, but we will have to be careful when we use it.

Also, the right-movers give

$$\begin{aligned}
\tilde{L}_n &= \frac{1}{2} \sum_{m=-\infty}^{\infty} \tilde{\alpha}_{n-m}^\mu \tilde{\alpha}_{\mu m}\,, \\
\tilde{L}_0 &= \frac{1}{2}\tilde{\alpha}_0^\mu \tilde{\alpha}_{\mu 0} + \sum_{m=1}^{\infty} \tilde{\alpha}_{-m}^\mu \tilde{\alpha}_{\mu m} = \frac{\alpha'}{4} p^\mu p_\mu + \sum_{m=1}^{\infty} \tilde{\alpha}_{-m}^\mu \tilde{\alpha}_{\mu m}\,.
\end{aligned} \tag{5.18}$$

The Virasoro algebra

For symmetry generators, the commutator algebra closes. Thus we are interested in

$$L_m L_n - L_n L_m = \frac{1}{4} \sum_{k,l=-\infty}^{\infty} (\alpha_{m-k}^\mu \alpha_{\mu k} \alpha_{n-l}^\nu \alpha_{\nu l} - \alpha_{n-l}^\nu \alpha_{\nu l} \alpha_{m-k}^\mu \alpha_{\mu k})\,. \tag{5.19}$$

(A constant has a vanishing commutator so I don't have to worry about the ordering in L_0.) Now, the first term cancels against the second if I can move the two α's on the the left to the right, so I just have to pick up the commutator terms, e.g. $\alpha_{\mu k} \alpha_{n-l}^\nu = \alpha_{n-l}^\nu \alpha_{\mu k} + k\delta_\mu^\nu \delta_{k+n-l,0}$. The second term, inserted, kills the l sum and leaves

$$\frac{1}{4} \sum_{k=-\infty}^{\infty} k\alpha_{m-k}^\mu \alpha_{\mu\, n+k} \tag{5.20}$$

Now, we can rename $k \to k - n$ so this becomes

$$\frac{1}{4} \sum_{k=-\infty}^{\infty} (k - n) \alpha^{\mu}_{m+n-k} \alpha_{\mu\, k} \tag{5.21}$$

We can also rename it to $-k + m$, so the sum becomes

$$\frac{1}{4} \sum_{k=-\infty}^{\infty} (m - k) \alpha^{\mu}_{k} \alpha_{\mu\, n+m-k} \tag{5.22}$$

If you average these two expressions you get

$$\frac{1}{8} \sum_{k=-\infty}^{\infty} (m - n) \alpha^{\mu}_{k} \alpha_{\mu\, n+m-k} = \frac{1}{4}(m - n) L_{m+n} \,. \tag{5.23}$$

The other three commutator terms make an equal contribution (this is obvious in retrospect, since the current is symmetric in the two $\partial_z X$'s), so we have

$$L_m L_n - L_m L_n = (m - n) L_{m+n} + \text{constant} \,. \tag{5.24}$$

The constant comes from two places: the ordering in L_0, and also from *double commutators*. I.e there are two kinds of term: those with two oscillators, which we've found, and those with zero oscillators, which we now find. The simplest way to evaluate this unambiguously is to act on the oscillator vacuum:

$$(L_m L_n - L_n L_m)|0, 0\rangle = \frac{1}{4} \sum_{k,l=-\infty}^{\infty} (\alpha^{\mu}_{m-k} \alpha_{\mu k} \alpha^{\nu}_{n-l} \alpha_{\nu l} - \alpha^{\nu}_{n-l} \alpha_{\nu l} \alpha^{\mu}_{m-k} \alpha_{\mu k})|0, 0\rangle \,. \tag{5.25}$$

It's still a bit tedious, but here are some shortcuts: first, to get a constant out all the oscillator modes have to sum to zero in pairs, which implies $m + n = 0$. So let $n = -m$ and $m \geq 0$. Second, any term where a lowering operator or α_0 hits $|0, 0\rangle$ gives zero. So only the first term on the right contributes, and only for $l < 0$ and $n - l < 0$. There are then two equal terms, one of which has $k = -l$ and the other $m - k = -l$. These are equal in the end, so we get

$$\frac{1}{2} \delta^{\nu}_{\mu} \delta^{\mu}_{\nu} \sum_{k=1}^{m-1} (m - k) k |0, 0\rangle = \frac{D}{12}(m^3 - m)|0, 0\rangle \,. \tag{5.26}$$

The commutators require $\mu = \nu$, but then sum over μ to get D. Finally, this is the *Virasoro algebra*,

$$L_m L_n - L_n L_m = (m - n) L_{m+n} + \frac{D}{12}(m^3 - m) \delta_{m+n,0} \,. \tag{5.27}$$

The right-movers satisfy the same algebra. For the open string there is only one copy.

In the text I derive this in a more abstract way, and show that in any conformally invariant theory the algebra of the generators has this form, except that D is replaced by a general constant c known as the *central charge.*

You've probably never seen an infinite dimensional algebra before, but this one is pretty tame. Notice that for $m = 0$ we get $L_0 L_n = L_n(L_0 - n)$, so acting with L_n shifts the value of L_0 by $-n$. So we can do the same thing we do for $SU(2)$: divide the generators into one that we diagonalize (L_0), and then those that lower it ($n > 0$), and those that raise it ($n < 0$).

In the open string, L_0 is just the world-sheet Hamiltonian. In the closed string the Hamiltonian is $L_0 + \tilde{L}_0$, while the $L_0 - \tilde{L}_0$ generates translations of σ.

The conformal group is the set of isomorphisms of flat spacetime that leave the flat metric $d\sigma^a d\sigma^a$ (or $\eta_{ab} d\sigma^a d\sigma^b$ in the Lorentzian case) invariant. We see that in $d = 2$ dimensions there are an infinite number of such transformations. In $d \geq 3$ the number is finite: d translations and $d(d-1)/2$ Lorentz transformations (rotations and boosts), plus the overal rescaling of the metric and d 'special conformal transformations.' Together, these form the group $SO(d, 2)$ in the Lorentzian case and $SO(d+1, 1)$ in the Euclidean case (recall that the Lorentz group is $SO(d-1, 1)$.

6 Old covariant quantization

(Sec. 4.1 of the text). Now we have the tools we need to develop string theory. First up is the spectrum. So far we have D sets of oscillators, and this is too many. The light cone quantization gave two fewer sets, because we haven't yet fixed all the gauge symmetry. Also, we have a problem with the inner product. First off, we have

$$\langle 0, k | 0, k' \rangle = (2\pi)^D \delta^D(k - k') \tag{6.1}$$

by momentum conservation. For compactness I'll usually omit this factor and concentrate on the contribution of the nonzero modes. We have (for fixed μ)

$$\| \alpha^\mu_{-m} |0\rangle \|^2 = \langle 0 | \alpha^\mu_m \alpha^\mu_{-m} | 0 \rangle = \langle 0 | (\alpha^\mu_{-m} \alpha^\mu_m + [\alpha^\mu_m, \alpha^\mu_{-m}]) | 0 \rangle = m \eta^{\mu\mu} \, . \tag{6.2}$$

(Recall that $\alpha^{\mu\dagger}_m = \alpha^\mu_{-m}$, because these are the Fourier modes of a real variable X^μ). This is negative for $\mu = 0$, inconsistent with the probability interpretation of $\| \, \|$. This is the price of a covariant description, and the reason why we get local symmetries in spacetime.

Our Hilbert space is too big. The Virasoro generators L_m are the operators that generate the left-over gauge symmetry. This is supposed to act trivially on physical quantities, so for physical states $|\psi\rangle, |\psi'\rangle$ we need at least that

$$\langle \psi | L_m | \psi' \rangle = 0 \tag{6.3}$$

for all m (in a sec we'll add a constant for L_0). It would be natural to insist that L_m annihilate physical states for all m, but this is too strong. The weakest condition that we can impose that would give the above is that the lowering generators annihilate physical states,

$$L_m|\psi\rangle = 0\,, \quad m > 0\,. \tag{6.4}$$

For $m < 0$ the matrix element (6.3) then vanishes by action to the left, as $L_m^\dagger = L_{-m}$. Also for L_0 we need

$$L_0|\psi\rangle = -A|\psi\rangle \tag{6.5}$$

where we include a possible ordering constant A [I'm sorry, in class 2/1 I flipped the sign of A!]. A state satisfying the conditions (6.4, 6.5) is called *physical.*

This splitting of the algebra might seem a bit ad hoc, but it is the same thing that one does in the covariant quantization of the photon (though this is rarely discussed in the texts these days). A more general approach, BRST, is equivalent to this. In order to streamline the course I will omit BRST: it is an efficient way to package the unphysical (gauge) information, but I want to cover the physics first.

Now, any state of the form $L_{-n}|\chi\rangle$ for $n > 0$ is orthogonal to all physical states,

$$\langle\psi|L_{-n}|\chi\rangle = \langle L_n\psi|\chi\rangle = 0\,. \tag{6.6}$$

Such a state is called *spurious,* and a spurious physical state is called *null.* A null state is orthogonal to all physics states including itself, so it is essentially zero. We therefore need an equivalence relation, for two physical states $|\psi\rangle$, $|\psi'\rangle$:

$$|\psi\rangle \cong |\psi'\rangle \text{ if } |\psi\rangle - |\psi'\rangle \text{ is null.} \tag{6.7}$$

Again, similar things happen with the photon, which we'll see as a special case. We will call the equivalence classes the *observable* spectrum, since we've already used the word physical.

First at the lowest level (in terms of oscillation number), $|0,k\rangle$, L_m annihilates this for positive m because there is no lower state. Also there are no spurious states here, because there is nothing lower to raise. So we have only the condition

$$0 = (L_0 + A)|0,k\rangle = (\alpha' k^2/4 + A)\,. \tag{6.8}$$

If we're doing the closed string we have also the tilded condition, which is the same here; if we're doing the open string it's just $\alpha' k^2 - A$. So $M^2 = -k^2 = 4A/\alpha'$ or A/α'.

Things get more interesting at the next level; I'll do the open string so as to only have one set of oscillators to write. Then we have

$$e_\mu \alpha_{-1}^\mu |0,k\rangle, \tag{6.9}$$

taking a general linear combination e_μ. The L_0 condition is $M^2 = -k^2 = (1+A)/\alpha'$. In L_1 the only terms that we need are

$$\alpha_0^\mu \alpha_{\mu 1} = \sqrt{2\alpha'} p \cdot \alpha_1 \, . \tag{6.10}$$

The physical state condition is

$$0 = \sqrt{2\alpha'} p \cdot \alpha_1 e \cdot \alpha_{-1} |0,k\rangle = \sqrt{2\alpha'} p \cdot e |0,k\rangle \tag{6.11}$$

or $k_\mu e^\mu = 0$. (Here I'm using p^μ as the operator and k^μ as the eigenvalue, as in the text). The spurious states are of the form

$$L_{-1}|0,k\rangle = \sqrt{2\alpha'} k \cdot \alpha_{-1} |0,k\rangle \, . \tag{6.12}$$

So the spurious states have $e \propto k$, and this state is null precisely if $k^2 = 0$.

Now we have three cases depending on the value of A. If $A > -1$ then $M^2 > 0$ and we can go to the rest frame. Then the physical state condition is $e^0 = 0$, getting rid of our timelike mode with the negative norm. There is no null state, so the observable spectrum is just the $D-1$ spacelike oscillations, all of which have positive norm.

If $A < -1$ then this state has negative mass-squared, spacelike. We can go to the frame where the momentum points in the 1-direction. The state $\alpha_{-1}^0 |0,k\rangle$ is then physical, but it has a negative norm so this is clearly bad.

The case $A = -1$ is the most interesting. Now $M^2 = 0$, and also the spurious mode is physical. So we have

$$k^2 = 0 \, , \quad k \cdot e = 0 \, , \quad e^\mu \cong e^\mu + \gamma k^\mu \, . \tag{6.13}$$

Going to a frame in which $k_\mu = (1,1,0,0,0,\ldots), k^\mu = (-1,1,0,0,0,\ldots)$, this says that $e^1 = -e^0$ and $(e^0, -e^0, e^2, \ldots) \cong (e^0 - \gamma, -e^0 + \gamma, e^2, \ldots)$, and so we can choose γ to get to $(0,0,e^2,e^3,\ldots)$. So the observable Hilbert space consists of $D-2$ states with positive norm.

The case $A = -1$ gives the same number of states as the light-cone analysis, so this seems likely to be the right one: different gauges should agree. For $A > -1$ there is no obvious problem, but the disagreement with the light-cone counting suggests that something is wrong, and indeed the interactions do not work. (We will see later why $A = -1$ is required by the world-sheet symmetries: only in this case is the 'vertex operator' that governs string interactions conformally invariant: this is the value.)

Notice that this is the same as electrodynamics in covariant (Lorenz) gauge $\partial_\mu A^\mu = 0$. The field equation becomes $\Box A^\mu = 0$, so these are the L_1 and L_0 conditions. Moreover, solutions of the form $A_\mu = \partial_\mu \lambda$ for $\Box \lambda = 0$ satisfy both the equation of motion and the gauge condition, but they are pure gauge, so that A_μ and $A_\mu + \partial_\mu \lambda$ are equivalent. So string theory is giving some infinite component extension of ordinary gauge invariance; this is made more concrete in string field theory. It was once believed that gauge invariance is beautiful and fundamental, as we go to higher energies we would find a bigger and bigger gauge symmetry, spontaneously broken, so having this much bigger gauge invariance seems like a wonderful thing. But over the years we have learned that gauge symmetry is often the opposite of fundamental, it emerges from the dynamics, so who knows?

By the way, Weinberg's argument that the interactions must be those of a gauge field are simply based on the requirement that the equivalence relation extend to the interactions: equivalent states must have equal scattering amplitudes. If one wants a theory which is covariant but only has positive norm states then we need the equivalence relation. But this essentially means that the vector field is coupled to a conserved current, since a polarization proportional to k_μ gives zero, and if you work out the full action you get gauge theory or general relativity.

For the closed string you just get two copies of the above, one with everything tilded, so for $A = -1$ there are $(D-2)^2$ states as we found before. The general state is

$$e_{\mu\nu}\alpha^\mu_{-1}\tilde{\alpha}^\nu_{-1}|0, k\rangle \,. \tag{6.14}$$

We have the physical state conditions

$$k^\mu e_{\mu\nu} = k^\nu e_{\mu\nu} = 0 \,, \tag{6.15}$$

and the equivalence

$$e_{\mu\nu} \cong e_{\mu\nu} + k_\mu \tilde{\gamma}_\nu + \gamma_\mu k_\mu \,. \tag{6.16}$$

For the symmetric part this describes the covariant quantization of the graviton, and for the antisymmetric tensor part it describes the antisymmetric tensor field, which also must have a gauge invariance.

At the next level, where we excite one α_{-2} or two α_{-1} oscillators there is an interesting story. Taking the value $A = -1$ that we have already found, if $D > 26$ there is a negative norm physical state, bad. If $D < 26$ the norm is positive but there are more states than in the open string spectrum. Precisely if $D = 26$ the number of states matches the light-cone quantization, so only for this number are all the symmetries preserved. It gets a bit technical, but indeed only for $D = 26$ can we define sensible scattering amplitudes, where the negative norm states decouple.

The root of the problem is that for $D \neq 26$ there is an anomaly in the Weyl invariance. We are not going to go through it, because it would require us to develop a lot of machinery that we wouldn't use again. First we would need to keep carefully the determinant from gauge fixing, and write it in terms of a path integral over Faddeev-Popov ghosts. Then one finds that the ghosts have central charge -26, so the total vanishes for $D = 26$. One can then deduce from this that the measure is Weyl invariant only for $D = 26$. Also, the BRST charge only has the correct property (squaring to zero) in $D = 26$.

7 Local operators: OPE and conformal properties

Normal ordering

We have talked about $X^\mu(z, \bar{z})$, but we are going to need also things like $X^\mu(z, \bar{z})X^\nu(z, \bar{z})$, at the same point on the world-sheet. Now, the naive product does not make sense, it diverges because of the effect of the quantum fluctuations of all the Fourier modes of X. But there is a well defined finite object, which is roughly the product of X^μ and X^ν, which we will learn how to define. Similarly for products of three or more fields, and even exponentials $e^{ik\cdot X}$, and then things like $\partial_z X^\mu \partial_{\bar{z}} X^\nu e^{ik\cdot X}$. Products of local fields, suitably defined, are termed composite operators.

So we are going to define normal ordering, meaning that you write out all the mode expansions and move all the lowering operators to the right and the raising operators to the left. We are going to adopt the convention, which turns out to be very convenient, that p^μ is grouped with the lowering and x^μ with the raising. So the mode expansion (4.17) becomes

$$
\begin{aligned}
X^\mu &= X^{\mu-} + X^{\mu+} \\
X^{\mu-} &= -i\frac{\alpha'}{2}p^\mu \ln|z^2| + i\sqrt{\frac{\alpha'}{2}}\sum_{m=1}^{\infty}\frac{1}{m}\left(\frac{\alpha_m^\mu}{z^m} + \frac{\tilde{\alpha}_m^\mu}{\bar{z}^m}\right) \\
X^{\mu+} &= x^\mu + i\sqrt{\frac{\alpha'}{2}}\sum_{m=-\infty}^{-1}\frac{1}{m}\left(\frac{\alpha_m^\mu}{z^m} + \frac{\tilde{\alpha}_m^\mu}{\bar{z}^m}\right) .
\end{aligned}
\tag{7.1}
$$

Here $-$ and $+$ are lowering and raising, not to be confused with the light-cone $-$ and $+$ that we used before. So if we have two fields, the simple product is

$$
X^\mu(z, \bar{z})X^\nu(z', \bar{z}') = \{X^{\mu-}(z, \bar{z}) + X^{\mu+}(z, \bar{z})\}\{X^{\nu-}(z', \bar{z}') + X^{\nu+}(z', \bar{z}')\} , \tag{7.2}
$$

and the normal ordered product is

$$
\begin{aligned}
: X^\mu(z, \bar{z})X^\nu(z', \bar{z}') : = {} & X^{\mu-}(z, \bar{z})X^{\nu-}(z', \bar{z}') + X^{\mu+}(z, \bar{z})X^{\nu-}(z', \bar{z}') \\
& + X^{\nu+}(z', \bar{z}')X^{\mu-}(z, \bar{z}) + X^{\mu+}(z, \bar{z})X^{\nu+}(z', \bar{z}') .
\end{aligned}
\tag{7.3}
$$

When both are lowering or both are raising they commute, and the order doesn't matter, but notice the reversed order in the third term on the right.

Comparing the two expressions, we can write

$$X^\mu(z,\bar z)X^\nu(z',\bar z') = \,:X^\mu(z,\bar z)X^\nu(z',\bar z'): + \left[X^{\mu-}(z,\bar z), X^{\nu+}(z',\bar z')\right]. \tag{7.4}$$

Now, from the mode expansions we get

$$\begin{aligned}\left[X^{\mu-}(z,\bar z), X^{\nu+}(z',\bar z')\right] &= -\frac{\alpha'}{2}\eta^{\mu\nu}\left\{\ln|z|^2 + \sum_{m=1}^{\infty} m\left(\frac{z'^m}{z^m} + \frac{\bar z'^m}{\bar z^m}\right)\right\} \\ &= -\frac{\alpha'}{2}\eta^{\mu\nu}\ln(|z-z'|^2)\,. \end{aligned}\tag{7.5}$$

Actually the sum only converges when $|z| > |z'|$, but that is fine: Euclidean amplitudes only make sense when time-ordered (and this is what the path integral gives automatically), and $|z| = e^{\sigma^2}$. So we write

$$\mathrm{T}X^\mu(z,\bar z)X^\nu(z',\bar z') = -\frac{\alpha'}{2}\eta^{\mu\nu}\ln|z-z'|^2 + :X^\mu(z,\bar z)X^\nu(z',\bar z'): \tag{7.6}$$

and this is always true: on the left side the instruction is to put the field with the larger time argument to the left. Both sides are actually symmetric if we swap z,μ with z',ν. In this chapter I will be pedantic and always write the time-ordering symbol, but if I ever forget it should be understood.

We can now see the problem if we try to bring two operators together: the log blows up. The problem is on the left, where the first operator acts by creation and the second annihilates the same mode; the sum over modes diverges as we bring the operators together. The normal-ordered operator is protected from this. By the way, in the book I introduce normal ordering in a different way, via the path integral rather than operators, but for X^μ the result is equivalent to raising/lowering ordering so I will not make the distinction.[7] By the way, we luck out because we are in free field theory. In an interacting theory normal ordering is not enough to define composite operators, we need to go through the whole renormalization program to define them. Such operators have many uses in physics, but the subject is usually passed over in QFT texts; Peskin and Schroeder is an exception.

Let us go a bit further: set $z' = 0$ and Taylor expand in z:

$$\mathrm{T}X^\mu(z,\bar z)X^\nu(0,0) = -\frac{\alpha'}{2}\ln|z|^2 + \sum_{m,n=0}^{\infty}\frac{z^m\bar z^n}{m!\,n!}:(\partial_z^m\partial_{\bar z}^n X^\mu(0,0))X^\nu(0,0):. \tag{7.7}$$

The Taylor expansion is good because the normal ordering makes everything smooth. We can simplify this expression because the equation of motion gives zero unless at least one of

[7] For other fields they are not the same; the path integral version is superior for building operators with simple conformal properties, but we will not need this. I used different notations for the different kinds of ordering, but since they are all the same for XX we won't need this.

m and n vanishes, but this is a special case and the double sum is more typical. Now, notice that every operator on the right is local at 0. So this is like a Taylor expansion but unlike a Taylor expansion there is a nonanalytic term: this is called an *operator product expansion* and is always possible in QFT. More generally, as we will see, you get non-integer powers of z and $\bar{z}$, and if there is no conformal symmetry you get much more general $z, \bar{z}$ dependence.

In string theory, the operators

$$: e^{ik_\mu X^\mu(z,\bar{z})} := e^{ik\cdot X^+(z,\bar{z})} e^{ik\cdot X^-(z,\bar{z})} \tag{7.8}$$

will play a valuable role. In your QFT course you have never exponentiated a local field operator because it is a rather singular thing in $D = 4$, but it $D = 2$ it has useful propoerties.

Now, let's consider the product of two such operators. We will need the Campbell-Baker-Hausdorf lemma: if A and B are two matrices such that $[A, B]$ commutes with both A and B, then

$$e^A e^B = e^B e^A e^{[A,B]} \,. \tag{7.9}$$

Here's a quick derivation. First, I claim that

$$e^{sA} B e^{-sA} = B + s[A, B] \,. \tag{7.10}$$

Proof: it's obvious at $s = 0$. Now take d/ds of both sides, and on the left get $e^{sA}(AB - BA)e^{-sA} = [A, B]$ since e^{sA} commutes with $[A, B]$; now integrate from $s = 0$. We will use this below for $s = 1$. Now,

$$e^A B^n e^{-A} = (e^A B e^{-A})^n \tag{7.11}$$

and so by power series $e^A f(B) e^{-A} = f(e^A B e^{-A})$ for any f, and in particular

$$e^A e^B e^{-A} = e^{e^A B e^{-A}} = e^{B+[A,B]} = e^B e^{[A,B]} \tag{7.12}$$

where the last step follows because B and $[A, B]$ commute. Multiply by e^A on the right and you're done.

Now let's apply this (assume for convenience that $|z_1| > |z_2|$),

$$\begin{aligned}
\text{T}\!: e^{ik_1\cdot X(z_1,\bar{z}_1)} :: e^{ik_2\cdot X(z_2,\bar{z}_2)} : &= e^{ik_1\cdot X^+(z_1,\bar{z}_1)} e^{ik_1\cdot X^-(z_1,\bar{z}_1)} e^{ik_2\cdot X^+(z_2,\bar{z}_2)} e^{ik_2\cdot X^-(z_2,\bar{z}_2)} \\
&= e^{ik_1\cdot X^+(z_1,\bar{z}_1)} e^{ik_2\cdot X^+(z_2,\bar{z}_2)} e^{ik_1\cdot X^-(z_1,\bar{z}_1)} e^{ik_2\cdot X^-(z_2,\bar{z}_2)} e^{-[k_1\cdot X^-(z_1,\bar{z}_1),k_2\cdot X^+(z_2,\bar{z}_2)]} \\
&= : e^{ik_1\cdot X(z_1,\bar{z}_1)+ik_2\cdot X(z_2,\bar{z}_2)} : e^{\frac{1}{2}\alpha' k_1\cdot k_2 \ln|z_1-z_2|^2} \\
&= |z_1 - z_2|^{\alpha' k_1\cdot k_2} : e^{ik_1\cdot X(z_1,\bar{z}_1)+ik_2\cdot X(z_2,\bar{z}_2)} :
\end{aligned} \tag{7.13}$$

Now we can derive an OPE. The normal-ordered expression on the right is smooth as $z_1 \to z_2$, so we Taylor expand,

$$\text{T}\!: e^{ik_1\cdot X(z_1,\bar{z}_1)} :: e^{ik_2\cdot X(z_2,\bar{z}_2)} := |z_1 - z_2|^{\alpha' k_1\cdot k_2} : e^{i(k_1+k_2)\cdot X(z_2,\bar{z}_2)} : + O(|z_1 - z_2|^{\alpha' k_1\cdot k_2+1}) \,. \tag{7.14}$$

Notice that the power of the singularity depends on k_1 and k_2. You can carry the expansion further as in Eq. (7.7).

Dimensions and tensors

For a conformal transformation $z'(z)$, we have

$$X^\mu(z,\bar z) \to X^\mu(z',\bar z')\,. \tag{7.15}$$

Then

$$\partial_z X^\mu(z,\bar z) \to \partial_z X^\mu(z',\bar z') = \frac{\partial z'}{\partial z}\partial_{z'} X^\mu(z',\bar z')\,. \tag{7.16}$$

(Think of the field as the derivative of X^μ with respect to its first argument, and on the left the field is evaluated at z and on the right at z').

A local operator $\mathcal{O}$ is called a *tensor of weight* $(h,\tilde h)$ if

$$\mathcal{O}(z,\bar z) \to \left(\frac{\partial z'}{\partial z}\right)^h \left(\frac{\partial \bar z'}{\partial \bar z}\right)^{\tilde h} \mathcal{O}(z',\bar z')\,. \tag{7.17}$$

So $\partial_z X^\mu$ is a tensor of weight $(1,0)$ and $\partial_{\bar z} X^\mu$ is a tensor of weight $(0,1)$. Not every field is a tensor, many have transformations involving higher derivatives. I should call this a conformal tensor, the conformal transformation arose as a combination of a Weyl transformation and a coordinate transformation, so it's not just the coordinate transformation property.

More generally, though, we can take the special case $z' = \lambda z$ for constant λ. Then any local operator $\mathcal{A}$ is said to have weight $(h,\tilde h)$ if

$$\mathcal{A}(z,\bar z) \to \lambda^h \bar\lambda^{\tilde h} \mathcal{A}(z',\bar z')\,. \tag{7.18}$$

For λ real, this is a scale transformation, and we call $h+\tilde h$ the dimension of $\mathcal{A}$. For $\lambda = e^{i\theta}$ it is a rotation, and we call $h-\tilde h$ the spin of $\mathcal{A}$. So h counts the number of z derivatives and $\tilde h$ counts the number of $\bar z$ derivatives.

Now let $\mathcal{A}_i$ be a complete set of local operators, with weights $(h_i,\tilde h_i)$ (I will give a concrete construction soon). The OPE states that

$$\mathrm{T}\,\mathcal{A}_i(z,\bar z)\mathcal{A}_j(0,0) = \sum_k c^k_{ij}(z,\bar z)\mathcal{A}_k(0,0)\,. \tag{7.19}$$

(Here and henceforth, time ordering is so ubiquitous that I will assume it, and indicate explicitly if an equation is not time ordered. Occasionally I will put the T in for emphasis.) For now you can think of this as an asymptotic expansion, but it is actually better than this, as I will explain later. Now make the transformation $z' = \lambda z$ on the operators on both sides,

$$\lambda^{h_i+h_j}\bar\lambda^{\tilde h_i+\tilde h_j}\mathrm{T}\,\mathcal{A}_i(\lambda z,\bar\lambda\bar z)\mathcal{A}_j(0,0) = \sum_k c^k_{ij}(z,\bar z)\lambda^{h_k}\bar\lambda^{\tilde h_k}\mathcal{A}_k(0,0)\,. \tag{7.20}$$

Equality still holds because this is a symmetry. But now we can use the OPE to rewrite the left side as

$$\lambda^{h_i+h_j}\bar{\lambda}^{\tilde{h}_i+\tilde{h}_j}\sum_k c_{ij}^k(\lambda z,\bar{\lambda}\bar{z})\mathcal{A}_k(0,0)\,. \tag{7.21}$$

Equating the k term on both sides gives

$$c_{ij}^k(\lambda z,\bar{\lambda}\bar{z})=\lambda^{h_k-h_i-h_j}\bar{\lambda}^{\tilde{h}_k-\tilde{h}_i-\tilde{h}_j}c_{ij}^k(z,\bar{z})\,. \tag{7.22}$$

Now set $z=\bar{z}=1$ and rename $\lambda\to z$ to get

$$c_{ij}^k(z,\bar{z})=z^{h_k-h_i-h_j}\bar{z}^{\tilde{h}_k-\tilde{h}_i-\tilde{h}_j}c_{ij}^k(1,\bar{1})\,. \tag{7.23}$$

So the z dependence is fully determined by the dimensions,

$$\mathrm{T}\,\mathcal{A}_i(z,\bar{z})\mathcal{A}_j(0,0)=\sum_k z^{h_k-h_i-h_j}\bar{z}^{\tilde{h}_k-\tilde{h}_i-\tilde{h}_j}c_{ij}^k(1,1)\mathcal{A}_k(0,0)\,. \tag{7.24}$$

and we see that operators of smallest dimension $(h_k+\tilde{h}_k)$ are most important in the sum. (Notation: we can omit the (1,1).)

By the way, we don't worry about δ-function terms in OPE's, which vanish at finite separation. These are at best convention-dependent - we have to define what we mean by the two operators at the point - and more often meaningless since the product just diverges. None of our applications of the OPE will depend on such terms.

Now let us apply this to the product (7.14). Writing $h(k),\tilde{h}(k)$ for the weights of $:e^{ik\cdot x}:$, we get

$$h(k_1+k_2)-h(k_1)-h(k_2)=\tilde{h}(k_1+k_2)-\tilde{h}(k_1)-\tilde{h}(k_2)=\frac{\alpha'}{2}k_1\cdot k_2\,, \tag{7.25}$$

which has the unique solution

$$h(k)=\tilde{h}(k)=\frac{\alpha'}{4}k\cdot k\,. \tag{7.26}$$

You might have thought the dimension would be zero because there are no derivatives, we have just taken a function of X^μ, but the normal ordering sneaks in some conformal dependence. This is an elegant result, which has implications beyond string theory (for example, it determines the temperature dependence of the tunneling rate between edges of fractional quantum Hall systems).

So a complete set $\mathcal{A}_i$ would be

$$:e^{ik\cdot X}\prod_{m=1}^{\infty}\prod_{\mu}(\partial_z^m X^\mu)^{N_{m\mu}}\prod_{m=1}^{\infty}\prod_{\mu}(\partial_{\bar{z}}^m X^\mu)^{\tilde{N}_{m\mu}}:\,. \tag{7.27}$$

These have weights

$$h = \frac{\alpha'}{4}k^2 + N\,, \quad \tilde{h} = \frac{\alpha'}{4}k^2 + \tilde{N}\,, \tag{7.28}$$

where

$$N = \sum_{m=1}^{\infty}\sum_{\mu} N_{m\mu}\,, \quad \tilde{N} = \sum_{m=1}^{\infty}\sum_{\mu} \tilde{N}_{m\mu}\,. \tag{7.29}$$

are the total number of $z, \bar{z}$ derivatives. The resemblance to equations (5.17, 5.18) is no accident, as we will see. It will be important to know which of these are tensors; we will develop this in the next chapters.

Further properties

Here I'd like to mention, mostly without derivation, some further results, which are useful for tying things together. More details are given in chapter 2 of the text.

For a tensor operator, the OPE with T is

$$\mathrm{T}\, T_{zz}(z)\mathcal{O}(0,0) = \frac{h}{z^2}\mathcal{O}(0,0) + \frac{1}{z}\partial_z\mathcal{O}(0,0) + \text{nonsingular}\,. \tag{7.30}$$

This is a Ward identity: it relates matrix elements of the conserved quantity to the transformation laws of the other fields. It is derived by an extension of Noether's theorem, see chapter 2 if you're interested. For non-tensor operators, there are additional terms with higher powers of $1/z$. For T with itself,

$$\mathrm{T}\, T_{zz}(z)T_{zz}(0,0) = \frac{c}{2z^4} + \frac{2}{z^2}T_{zz}(0,0) + \frac{1}{z}\partial_z T_{zz}(0,0) + \text{nonsingular} \tag{7.31}$$

so it is not a conformal tensor, and its transformation property is governed by the central charge. This OPE actually has exactly the same content as the Virasoro algebra; again, this is basically a Ward identity, and is derived by a contour integration argument in ch. 2.6.

Normal ordering satisfies Wick's theorem. Let $\mathcal{F}$ represent any product of operators, not necessarily at the same point. Then

$$\mathrm{T}\,\mathcal{F} = \,:\sum(\text{contractions of } \mathcal{F}):\,. \tag{7.32}$$

This means to sum over all ways of choosing $0, 1, 2, \ldots$ pairs $X^\mu(z_i, \bar{z}_i)X^\nu(z_j, \bar{z}_j)$ in $\mathcal{F}$ and replacing them with $-\frac{1}{2}\alpha' \ln|z_i - z_j|^2$. In our basic result (7.6) the terms have 1 and 0 subtractions. Also, if we have something like $:\mathcal{F}::\mathcal{G}:$ and we want to normal-order the product, then

$$\mathrm{T}\,:\mathcal{F}:\,:\mathcal{G}:\, = \,:\sum(\text{cross-contractions between } \mathcal{F} \text{ and } \mathcal{G}):\,. \tag{7.33}$$

In class I work out as an example the TX and Te^{ikX} OPEs. I suggest that you try the TT OPE.

8 Vertex operators

The state-operator isomorphism

Consider a 2-dimensional CFT (conformally invariant field theory) on the infinite cylinder, with some initial state specified by the boundary condition in the past. Under $w \to z$ this maps to the plane, where the initial condition now represents some weighting for the fields at the origin. Thus we have an isomorphism

$$\text{States of a CFT quantized on a circle} \quad \longleftrightarrow \quad \text{Local operators}\,. \tag{8.1}$$

We are more used to the first kind of space in quantum theory, but in a CFT they must be the same. For the free field theory we are looking at, it is easy to guess the isomorphism from symmetry. The ground state with momentum k^μ should map to the simplest operator with this momentum. Thus

$$|0,k\rangle \quad \longleftrightarrow \quad :e^{ik\cdot X}:\,. \tag{8.2}$$

In the text I derive this, by showing that the RHS is annihilated by the lowering operators, and also I try to be systematic about the normalization. Now, on the left we can excite with raising operators $\alpha^\mu_{-m}, \tilde\alpha^\mu_{-m}$ and on the right we can multiply by $\partial^m_z X^\mu, \partial^m_{\bar z} X^\mu$ (and normal order the product), so we guess

$$\alpha^\mu_{-m}, \tilde\alpha^\mu_{-m} \quad \longleftrightarrow \quad \partial^m_z X^\mu, \partial^m_{\bar z} X^\mu\,. \tag{8.3}$$

With this dictionary, the $L_0, \tilde L_0$ eigenvalues (5.17, 5.18) of the state are identical to the weights (7.28) of the operator, which makes sense. Again, derivations and normalizations are left to the text.

For a CFT in d spacetime dimensions, the same construction gives

$$\text{States of a CFT quantized on } S^{d-1} \quad \longleftrightarrow \quad \text{Local operators}\,, \tag{8.4}$$

which arises in AdS/CFT duality.

This isomorphism gives a simple derivation of the OPE for a general CFT (ch. 2.9). It is simply the insertion of a complete set of states! This also implies that it converges; the radius of convergence is the distance to the nearest other operator in the matrix element. So we can think of a CFT abstractly: the OPE coefficients c^k_{ij} and weights $h_i, \tilde h_i$ determine everything. They have a consistency condition, associativity (fig. 2.9 of text), and another one (modular invariance) that we will meet later. Sometimes it is possible to construct a CFT abstractly in this way. Idse Heemskerk, Joao Penedones, James Sully and I have recently been able to do this in $d = 2, 4$ with additional conditions motivated by AdS/CFT.

(1) How do we know that a set of operators is complete? Two answers: it closes under the OPE, and it is isomorphic to a complete basis for the Hilbert space (though modular invariance will give a further notion of completeness). (2) Do the weights h have to be real? In general, yes, since these are the eigenvalues of the world-sheet Hamiltonian L_0. In Euclidean CFT's that have a sensible Lorentzian continuation with unitary amplitudes there must be a good inner product under which L_0 is Hermitian, but in intrinsically Euclidean theories that only describe spatial behavior this property (reflection positivity) is not necessary and various funny things can happen. In this case there is sometimes another subtlety that we can't always diagonalize h. In 'logarithmic conformal field theories' we can only bring L_0 to Jordan normal form. But for string theory we are mostly interested in unitary CFT's. In chap. 2.9 I derive some results that hold only in unitary (reflection positive) CFT's.

Vertex operators

In the sum over string world-sheets, we can use the state/operator conformal transformation to replace the incoming and outgoing strings, which are semi-infinite cylinders, with disks with local operators. So we end up integrating over compact surfaces with local 'vertex operators' $\mathcal{V}(z,\bar{z})$. Summing over world-sheets will include an integration,

$$\int d^2z\, \mathcal{V}(z,\bar{z})\,. \tag{8.5}$$

Now, this has to be conformally invariant (surviving from the coordinate and Weyl invariance of the original path integral), so $\mathcal{V}(z,\bar{z})$ *must be a tensor of weight* $(1,1)$.

The tensor condition precisely corresponds to the statement that L_m annihilates the corresponding state, and the condition that $h=\tilde{h}=1$ is precisely the condition that we found in the old covariant quantization. For the states $|0,k\rangle \leftrightarrow\, : e^{ik\cdot X} :$, the weights are $h=\tilde{h}=\alpha' k^2/4$,, so we gets $k^2 = 4/\alpha'$. In the spacelike-positive metric this is $-M^2$, so we recover the same tachyonic value found in the light-cone quantization. For the states

$$\alpha^\mu_{-1}\tilde{\alpha}^\nu_{-1}|0,k\rangle \;\leftrightarrow\; :\partial_z X^\mu \partial_{\bar{z}} X^\nu e^{ik\cdot X}: \tag{8.6}$$

the weights are $h=\tilde{h}=1+\alpha' k^2/4$, so we again get the result that the first excited states are massless tensors.

For open strings the vertex operators live in the boundary [illustrate], and must have weight 1 in terms of $\partial y'/\partial y$ where y is a coordinate along the boundary. The isomorphism is

$$\text{States of a CFT quantized on a segment} \quad \longleftrightarrow \quad \text{Local operators on the boundary}\,. \tag{8.7}$$

The tachyon vertex op. is $: e^{ik\cdot X} :$ for $M^2 = -1/\alpha'$, and the first excited state is $: e^{ik\cdot X}\partial_y X^\mu :$, massless.

Later, we will back up to the original action, before gauge fixing, and discuss vertex operators in that context.

9 Tree amplitudes I: Preliminaries

The sphere

So now let's jump in and calculate a scattering amplitude. Based on the previous discussion, the tree level scattering of n closed string tachyons would be

$$S_{n\,\text{tachyon}}(k_1, \ldots, k_n) = \int \mathcal{D}X\, \mathcal{D}g\, e^{-S_P} \prod_{i=1}^{n} \int d^2\sigma_i \sqrt{g(\sigma_i)} e^{ik_i \cdot X(\sigma_i)}\,. \tag{9.1}$$

The path integral runs over all metrics with the topology of a sphere. For each vertex operator there is an integral over position, with the usual determinant factor, so as to make the expression coordinate invariant.

I haven't given any formal introduction to path integrals, but we'll figure out what to do as we go along. (In the Big Book I lay out the most general case and then specialize to the simple examples. Here, I focus on the simple examples).

First, we need to fix gauge. Earlier I asserted that we could locally always fix the metric to a specific form (and a justification is given around eq. 3.3.5 of the text). Here I will quote a theorem: any metric on the sphere can be brought by a combination of coordinate and Weyl transformations to the round constant-curvature metric

$$ds^2 = \frac{4dzd\bar{z}}{1 + z\bar{z}}\,, \tag{9.2}$$

(or to any other chosen form); the normalization here is chosen to give unit radius.

So, as before, we can forget about the metric, though this will not be true for one loop amplitudes. But there is still a subtlety: choosing the metric does not fix all the symmetry. On the sphere we do not have the full conformal symmetry that we have been discussing, but the *Möbius transformations*

$$z \to \frac{\alpha z + \beta}{\gamma z + \delta} \tag{9.3}$$

are one-to-one of the sphere into itself (including the point at infinity), provided that $\alpha\delta - \beta\gamma$ is nonvanishing. The parameters $\alpha, \beta, \gamma, \delta$ are arbitrary complex numbers. However, only three are independent, because if we multiply all by a common constant we get the same transformation; it is conventional to use this to set $\alpha\delta - \beta\gamma = 1$. So we have some more gauge symmetry to fix, and the simplest way to do it is to fix the positions of any three of the vertex operators.

Integral $\mathcal{D}X$

Operator evaluation

So now we have the path integral over the X's, and then we will take care of the integral over the unfixed vertex operator positions. I am going to do this in two ways. First, I claim that

$$\int \mathcal{D}X \, e^{-S_P} \prod_{i=1}^{n} \sqrt{g} e^{ik_i \cdot X(\sigma_i)} = \langle 0,0| \mathrm{T} \prod_{i=1}^{n} : e^{ik_i \cdot X(\sigma_i)} : |0,0\rangle \,. \tag{9.4}$$

I should mention, by the way, that the overall normalization will be discussed separately in the next chapter. The main point in writing this equation is the fact that a path integral with extra factors in the integrand is equal to the time-ordered matrix element of the corresponding operators. I have inserted the initial and final states $|0,0\rangle$ using the state operator mapping: this is the same as a trivial factor 1 in the path integral.

I haven't told you how I have defined the vertex operators on the left, you can just assume that I have done so so that the right-hand side comes out as it does. If you're willing to assume this, jump ahead to Eq. (9.10); otherwise, read on. [This is an attempt to improve on the discussion in class on 2/1.] First, Wick's theorem (7.32) can also be written as

$$: \mathcal{F} : = \sum (\text{subtractions of } \mathrm{T}\mathcal{F}) \,, \tag{9.5}$$

where a subtraction means to replace a given pair of X's with $\frac{1}{2}\alpha' \ln |z_i - z_j|^2$. Equivalently, for any propagator connecting two fields in $: \mathcal{F} :$, use the regular propagator plus $\frac{1}{2}\alpha' \ln |z_i - z_j|^2$: this has a smooth limit as $z_i \to z_j$, so it is well-defined, and it is completely equivalent to creation-annihilation ordering. (The smearing I introduced in class is not really necessary). So we have two equivalent ways of defining normal ordering: in terms of creation-annihilation ordering, and in terms of a modification of the propagator for self-contractions. To extend it to curved world-sheet in a covariant way, creation-annihilation ordering isn't really useful; instead, we use the second kind of description, adding $\frac{1}{2}\alpha' \ln d^2(\sigma, \sigma')$ to the propagator for self-contractions. Here $d^2(\sigma, \sigma')$ is the geodesic distance between the two points. This definition by construction preserves coordinate invariance, which is always nice to do. Now, writing the metric as $ds^2 = e^{2\omega(z,\bar{z})} dz d\bar{z}$, we have

$$\ln d^2(\sigma, \sigma') = 2\omega(z, \bar{z}) + \ln |z - z'|^2 \,, \tag{9.6}$$

to leading order in $z - z'$. So the covariantly regulated operators differ from the earlier normal ordering just by the ω factor. For exponentials, you know from problem set 3 that the contractions add up in a nice way, so that

$$(e^{ik\cdot X})_{\mathrm{r}} = e^{-\alpha' k^2 \omega(z,\bar{z})/2} : e^{ik\cdot X} : . \tag{9.7}$$

That is, the finite difference between the subtractions between two factors of $ik\cdot X$ is $-\alpha' k^2\omega$ times a symmetry factor of $\frac{1}{2}$, and these exponentiate. So this exactly cancels the $\sqrt{g}$ when we are at the correct mass $4/\alpha'$, and the result is Weyl invariant, independent of ω. This is no surprise, but we should emphasize: the external strings have to be on shell, as in the S matrix. Off-shell quantities aren't really physical, until we specify exactly how they are constructed and that can often be complicated.

A questoin about Pauli-Villars... This is not so different: one subtracts a propagator with large mass. Now, this falls exponentially, so only affects contractions within a vertex operator. Moreover, the massive propagator at short distance works out to essentially

$$-\frac{\alpha'}{2}\ln \Lambda^2 d^2(\sigma,\sigma') \tag{9.8}$$

so it's almost the same as the above. Only the large mass factor Λ is different, so that (using the simple combinatorics of exponentials)

$$(e^{ik\cdot X})_{\mathrm{PV}} = \Lambda^{+\alpha' k^2/2}(e^{ik\cdot X})_{\mathrm{r}} \tag{9.9}$$

So after using PV regularization we must renormalize by the multiplicative factor $\Lambda^{-\alpha' k^2/2}$, and then we get the same result as the covariant normal ordering above.

Now, our earlier result (7.13) generalizes immediately to n exponentials,

$$\mathrm{T}\prod_i :e^{ik_i\cdot X(z_i,\bar z_i)}: \;=\; :e^{i\sum_i k_i\cdot X(z_i,\bar z_i)}:\prod_{i<j}|z_i - z_j|^{\alpha' k_i\cdot k_j} \tag{9.10}$$

Inserting this into the matrix element (9.4) gives

$$\begin{aligned}
&\prod_{i<j}|z_i - z_j|^{\alpha' k_i\cdot k_j}\langle 0,0|\!:e^{i\sum_i k_i\cdot X(z_i,\bar z_i)}:\!|0,0\rangle \\
&= \prod_{i<j}|z_i - z_j|^{\alpha' k_i\cdot k_j}\langle 0,0|e^{i\sum_i k_i\cdot x}|0,0\rangle \\
&= (2\pi)^D\delta^D(\textstyle\sum_i k_i)\displaystyle\prod_{i<j}|z_i - z_j|^{\alpha' k_i\cdot k_j}\,.
\end{aligned} \tag{9.11}$$

In the second line, the lowering operators and p^μ act trivially to the right, the raising operators act trivially to the left, and only the zero mode x^μ remains. In the last line, we have used translation invariance to deduce what the answer must be, up to a convenient normalization.

We can also state the result for arbitrary excited state vertex operators, due to Wick's theorem it's just like in a Feynman graph: sum all contractions between pairs of X's in *distinct* vertex operators, with the propagator $-\frac{1}{2}\alpha'\ln|z-z'|^2$. (We can define the coordinate invariant regulator in such a way that the ω's from the regulator just cancel the explicit ones from the verter operator as above.) For the exponential vertex operators, the sum exponentiates to give the result (9.11).

Path integral evaluation

For the low order examples that we will study, the operator methods such as we have just used will be enough, and we could charge ahead to use the result (9.11) to get the string S-matrix, but it's a nice bit of practice to repeat the calculation using path integrals. For tachyon vertex operators this is just a few steps, because the path integral is nearly gaussian:

$$\int \mathcal{D}X \; \exp\left\{ -\frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{g} g^{ab} \partial_a X_\mu \partial_b X^\mu + i \sum_i k_i \cdot X(\sigma_i) \right\} \tag{9.12}$$

As usual we first solve the classical equation of motion,

$$\begin{aligned} 0 &= \partial_a (g^{ab} \sqrt{g} \partial_b X^\mu_{\rm cl}) + 2\pi\alpha' i \sum_i k_i^\mu \delta^2(\sigma - \sigma') \\ &= 2\partial_z \partial_{\bar z} X^\mu_{\rm cl} + i \sum_i k_i^\mu \delta^2(z - z', \bar z - \bar z')\,. \end{aligned} \tag{9.13}$$

Now we use

$$\partial_z \partial_{\bar z} \ln |z|^2 = 2\pi \delta^2(z, \bar z)\,. \tag{9.14}$$

You've essentially shown this in homework 2, from the momentum space Green's function. If we do the $\bar z$ derivative the left-hand side becomes $\partial_z(1/\bar z)$, which looks like it might be zero but is (a standard result in complex analysis) a delta function. Maybe the simplest way to see this is to regulate,

$$\partial_z \partial_{\bar z} \ln(|z|^2 + \epsilon) = \frac{\epsilon}{(|z|^2 + \epsilon)^2} \tag{9.15}$$

and the right-hand side goes to zero everywhere but the origin as $\epsilon \to 0$, but with area 2π; note the normalization below Eq. (4.16). So

$$X^\mu_{\rm cl}(z, \bar z) = -i \frac{\alpha'}{2} \sum_i k_i^\mu \ln |z - z_i|^2\,. \tag{9.16}$$

So, shifting the variable of integration $X^\mu = X^\mu_{\rm cl} + Y^\mu$, the path integral becomes

$$\int \mathcal{D}Y \exp\left\{ -\frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{g} g^{ab} (\partial_a Y_\mu \partial_b Y^\mu + 2\partial_a Y_\mu \partial_b X^\mu_{\rm cl} + \partial_a X_{\mu\rm cl} \partial_b X^\mu_{\rm cl}) + i \sum_i k_i \cdot (Y(\sigma_i) + X_{\rm cl}(\sigma_i)) \right\}. \tag{9.17}$$

Now, we have one extra step as compared to the usual QFT: we separate out the constant mode of Y:

$$Y^\mu(\sigma) = y^\mu + Y'^\mu(\sigma)\,, \quad \mathcal{D}Y = d^D y\, \mathcal{D}Y'\,. \tag{9.18}$$

We have to do this because y^μ has no quadratic term, it is not gaussian. It appears only multiplying $i\sum_i k_i^\mu$ in the action, so the y integral gives $(2\pi)^D\delta^D(\sum_i k_i)$. Now we can integrate by parts and use the equation of motion for $X_{\rm cl}$ to show that all terms in the exponent that are linear in Y' cancel, and the quadratic term in $X_{\rm cl}$ cancels half of the source term, so we are left with

$$(2\pi)^D\delta^D(\textstyle\sum_i k_i)e^{i\sum_i k_i\cdot X_{\rm cl}(\sigma_i)/2}\displaystyle\int \mathcal{D}Y' \exp\left\{-\frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{g}g^{ab}\partial_a Y'_\mu\partial_b Y'^\mu\right\}. \qquad (9.19)$$

As usual, the path integral has factorized and the last term is a constant, independent of the vertex operators.

There was one subtlety: as $z\to\infty$, $X^\mu_{\rm cl}$ is proportional to $\ln|z|^2\sum_i k_i^\mu$. If we didn't have the delta function from the y^μ integration, there would have been an extra term from the integration by parts. We made our lives simpler by doing the y integral to generate the delta function before integrating by parts. The fool who wrote the text didn't do this, so there are some annoying factors starting in 6.2.9, which drop out in 6.2.17.

Again, we are for now leaving the overall normalization aside, so we just get

$$(2\pi)^D\delta^D(\textstyle\sum_i k_i)\displaystyle\prod_{i,j}|z_i-z_j|^{\alpha' k_i\cdot k_j/2}. \qquad (9.20)$$

As in the operator evaluation, the $i=j$ terms are canceled by the normal ordering, and so we get

$$(2\pi)^D\delta^D(\textstyle\sum_i k_i)\displaystyle\prod_{i<j}|z_i-z_j|^{\alpha' k_i\cdot k_j}. \qquad (9.21)$$

just as before, (9.11). By adding a general source term for X^μ in the exponential, we can use this as a generating functional as in field theory, again getting Wick's theorem that for more general vertex operators we are to sum over all contractions.

Gauge fixing

According to the discussion of the Möbius group, we are to integrate the amplitude (9.11, 9.21) over

$$\int d^2z_4\ldots d^2z_n\,. \qquad (9.22)$$

Now let us look at the simplest nontrivial case, three tachyon vertex operators, where there is no integral left to do, and the answer is just given by Eq. (9.11)

$$|z_{12}|^{\alpha' k_1\cdot k_2}|z_{13}|^{\alpha' k_1\cdot k_3}|z_{23}|^{\alpha' k_2\cdot k_3}. \qquad (9.23)$$

For brevity I have left off the delta function, and let $z_{ij} = z_i - z_j$.

We have left off one factor. Essentially, we have been saying that

$$\frac{\int \mathcal{D}g \prod_{i=1}^{3} d^2\sigma_i}{\int \mathcal{D}\zeta\, \mathcal{D}\omega} = 1 \tag{9.24}$$

In the numerator are the integrals we gauge fixed, and in the denominator the gauge overcounting. However, we know from non-Abelian gauge theories that there is generally a Jacobian J to take into account, the Faddeev-Popov (FP) determinant. The text spends a lot of effort doing this systematically, but we are going to take the easy way out and figure out what this determinant must be.

There is a problem with the expression (9.23): it depends on where we have fixed the vertex operators, but this was supposed to be just a gauge choice. However, this simplifies a bit, since

$$2k_1 \cdot k_2 = (k_1 + k_2)^2 - k_1^2 - k_2^2 = k_3^2 - k_1^2 - k_2^2 = -\frac{4}{\alpha'}\,, \tag{9.25}$$

using momentum conservation plus the fact that we are on the mass-shell. So the three-point amplitude becomes

$$|z_{12} z_{13} z_{23}|^{-2}\,. \tag{9.26}$$

This does not depend on the vertex operator momenta, and so the dependence on the arbitrary choice of positions is canceled if we insert a factor of $J = |z_{12} z_{13} z_{23}|^2$, and this completely determines J.

We are left with the three-tachyon amplitude just equal to a constant, which was inevitable because there is nothing it can depend on, all the invariants are fixed as in Eq. (9.25). As to the value of the constant, we have to defer this just a little longer.

We can now write the n-point amplitude

$$\int d^2 z_4 \ldots d^2 z_n\, |z_{12} z_{13} z_{23}|^2 \prod_{i<j} |z_{ij}|^{\alpha' k_i \cdot k_j}\,. \tag{9.27}$$

It is not manifest that this is independent of the fixed positions, or that it is symmetric in the integrated and unintegrated operators, but in fact it is as a consequence of the FP procedure, as one can show by a Möbius transformation. It is conventional to set $z_1 = 0$, $z_2 = 1$ and to take the limit $z_3 \to \infty$; in particular, all the monomials involving z_3 cancel upon using momentum conservation and the mass shell condition, and we are left with

$$\int d^2 z_4 \ldots d^2 z_n \prod_{\substack{i<j \\ i,j \neq 3}} |z_{ij}|^{\alpha' k_i \cdot k_j}\,. \tag{9.28}$$

10 Tree amplitudes, II

The Virasoro-Shapiro amplitude

For $n = 4$, the amplitude becomes

$$\int d^2 z_4 \, |z_4|^{\alpha' k_1 \cdot k_4} |1 - z_4|^{\alpha' k_2 \cdot k_4} \, . \tag{10.1}$$

This integral can be done, with the result $(2\pi)^D \delta^D(\sum_i k_i)$ times

$$2\pi \frac{\Gamma(1 + \alpha' k_1 \cdot k_4/2)\Gamma(1 + \alpha' k_2 \cdot k_4/2)\Gamma(1 + \alpha' k_3 \cdot k_4/2)}{\Gamma(-\alpha' k_1 \cdot k_4/2)\Gamma(-\alpha' k_2 \cdot k_4/2)\Gamma(-\alpha' k_3 \cdot k_4/2)} \, . \tag{10.2}$$

Just to give a hint of where this comes from, we can represent

$$|z_{ij}|^{\alpha' k_i \cdot k_j} = \frac{1}{\Gamma(-\alpha' k_i \cdot k_j/2)} \int_0^\infty dt \, t^{-1-\alpha' k_i \cdot k_j/2} e^{-t|z_{ij}|^2} \, , \tag{10.3}$$

providing two of the six Γ functions, and few gaussian integrals and further uses of the Γ function integral yield the result.

For the three-point function, all the kinematic invariants were fixed by the mass-shell condition and momentum conservation. Here, we are describing a two particle to two particle scattering, so there are two invariants, which we can think of as the center of mass energy and the scattering angle.[8] We can take for example $k_1 \cdot k_4$ and $k_2 \cdot k_4$ to be the independent variables, you can check that all the other $k_i \cdot k_j$ are determined in terms of these. It is conventional to describe the scattering process in terms of the invariants named for my Ph. D. advisor, Stanley Mandelstam:

$$s = -(k_1 + k_2)^2 = -(k_3 + k_4)^2 \, , \quad t = -(k_1 + k_3)^2 = -(k_2 + k_4)^2 \, , \quad u = -(k_1 + k_4)^2 = -(k_2 + k_3)^2 \, . \tag{10.4}$$

These satisfy $s + t + u = \sum_i M_i^2 = -16/\alpha'$, so again two are independent. If we take $k_{1,2}$ to be incoming (time components positive) and $k_{3,4}$ to be outgoing,

$$s = E^2 \, , \quad t = (4M^2 - E^2)(1 - \cos\theta)/2 \, , \quad u = (4M^2 - E^2)(1 - \cos\theta)/2 \, , \tag{10.5}$$

where E is the center of mass energy, $M^2 = -4/\alpha'$ is the common mass-squared of the external particles, and θ is the scattering angle between particles 1 and 3 [draw]. In terms of these, the integral becomes

$$(10.2) = 2\pi \frac{\Gamma(-1 - \alpha' s/4)\Gamma(-1 - \alpha' t/4)\Gamma(-1 - \alpha' u/4)}{\Gamma(2 + \alpha' s/4)\Gamma(2 + \alpha' t/4)\Gamma(2 + \alpha' u/4)} \, . \tag{10.6}$$

[8] The fact that these particles are tachyons can be finessed, just let the spatial momenta be large enough that $\vec{k}^2 + M^2$ is positive for each particle, and the energies are then real.

The first thing to notice is that the amplitude has poles whenever one of the Γ functions in the numerator has a nonpositive integer as its argument, meaning that s, t, or u takes one of the values $-4/\alpha', 0, 4/\alpha', 8/\alpha', \ldots$. These are precisely the masses-squared of the closed string states, and is just what general principles require, these are resonances whenever any pair of external particles is degenerate with a single particle. In QFT, we would have separate graphs giving the s, t, and u channel poles [draw], but here they all come from a single expression.

The graphs suggest that the pole in the four-particle amplitude is related to the square of the three-particle amplitude, and indeed this follows from general principles of quantum theory. We are calculating the S-matrix

$$S_{ij} \equiv \langle i \text{ outgoing} | j \text{ incoming} \rangle \,. \tag{10.7}$$

We can separate this into an amplitude for no scattering to occur, the identity matrix, and an amplitude for scattering to occur, which we are calculating, $S = I + iT$. Now, conservation of probability means that S is unitary, i.e.

$$I = S^\dagger S = (I - iT^\dagger)(I + iT) \implies T^\dagger T = 2\,\mathrm{Im}\, T \,. \tag{10.8}$$

As developed in the text, Eq. 6.4.13 and chap. 9.1, this requires the poles to be there, and determines their normalization in terms of the square of the three point function. This allows us to normalize the scattering amplitudes.

The result is that unitarity is satisfied if we associate a factor $g_{\rm c}$ to each tachyon vertex operator, and a factor $8\pi i/\alpha' g_{\rm c}^2$ to the overall path integral. In other words,

$$\begin{aligned} S_{3\,\text{tachyon}}(k_1, k_2, k_3) &= (2\pi)^D \delta^D(\textstyle\sum_i k_i) \frac{8\pi i g_{\rm c}}{\alpha'} \,, \\ S_{4\,\text{tachyon}}(k_1, k_2, k_3, k_4) &= (2\pi)^D \delta^D(\textstyle\sum_i k_i) \frac{8\pi i g_{\rm c}^2}{\alpha'} \times (10.2 \text{ or } 10.6) \,. \end{aligned} \tag{10.9}$$

Various normalization constants that we have not worried about can all be absorbed into $g_{\rm c}$. We see that the three-tachyon amplitude is proportional to this, it is some sort of coupling constant that will appear in all three-closed string couplings, with the subscript because we are looking at closed strings here. It is an important result that this coupling constant is determined by the value of the dilaton field, $g_{\rm c} \propto e^{\Phi}$, as we will discuss later. By the way, the factor of i required by unitarity may be thought of as coming from the rotation to Euclidean space, you may recall a similar factor when you do QFT integrals.

The poles, by the way, arise from the vertex operator integral when the integrated vertex operator approaches one of the other three vertex operators, and so can be analyzed using the OPE. Also, the string amplitude (10.1), like the Γ function integral (10.3), has a limited range of convergence but can be analytically continued past the pole; this again is an artifact of the rotation to Euclidean world-sheets, but poses no difficulties.

One final point of interest in this amplitude is the high energy behavior. Using Stirling's approximation

$$\Gamma(x) \sim \sqrt{2\pi} e^{-x} x^{x-1/2}, \tag{10.10}$$

one can show that if one takes the center of mass energy to infinity at fixed angle, meaning that s and t go to infinity with a fixed ratio, the 4-tachyon amplitude goes to zero as $e^{-O(s)}$. For pointlike particles it generally behaves as a power, but for an extended object as we increase the energy it becomes less likely that it scatters without breaking. So the elastic amplitude falls exponentially, but the total cross section actually grows, as I will note shortly.

An interesting limit is s large with t fixed. This is known as the Regge regime. Using Stirling's approximation, you can show that the amplitude goes as

$$s^{2+\alpha' t/2}, \tag{10.11}$$

times a function of t. This is known as Regge behavior. Notice, by the way, that t is negative for s-channel scattering, Eq. (10.5). For exchange of a single particle of spin J, one would get s^J. To see this, note that in the s-channel a particle of spin J will give a spin-J spherical harmonic, which is a polynomial in $\cos\theta$ (and therefore t) with maximum power J. So if the exchange is in the t channel we get s^J. In string theory we have particles of all integer spin, so this infinite tower is summing up to something that looks like a single particle of variable spin.

The unitarity relation also determines the total (rather than elastic) cross section. The $(T^\dagger T)_{ii}$ term with initial and final states the same ($t \to 0$) is the squared amplitude summed over intermediate states, which is essentially the total cross section. This gets related to the imaginary part of T_{ii}; there is an extra $1/s$ from the kinematics, so the total cross section grows as s here. This growth can't go on forever, eventually higher order terms become important.

Regge behavior, as well as the 'duality' property that a single meromorphic function contains all of the s, t, and u poles, played a role in the initial discovery of string theory as a possible theory of the strong interaction.

Open string amplitudes

Open string amplitudes work largely the same way. Tree level amplitudes reduce to vertex operators on the edge of a disk. A disk is half a sphere, so it is convenient to take the upper half of the z-plane, with the vertex operators integrated along the real axis (I called the coordinate along the real axis y in the book, so I will stick with that). This is left invariant by the Möbius transformations with $\alpha, \beta, \gamma, \delta$ real, and again one scales out so we can fix three vertex operators. The expectation value of exponential vertex operators is almost the same,

$$\langle 0,0|\mathrm{T}\prod_{i=1}^{n} :e^{ik_i\cdot X(y_i)}: |0,0\rangle_{\rm open} = (2\pi)^D\delta^D(\textstyle\sum_i k_i)\displaystyle\prod_{i<j}|y_i - y_j|^{2\alpha' k_i\cdot k_j}\,. \tag{10.12}$$

Note the factor of two in the exponent compared to the closed string. The effect of the boundary on the XX propagator is to add an image charge term, and when the points are on the boundary this just doubles the result (implicitly I always use boundary normal ordering, with the image term, for open string vertex operators).

The full amplitude is this, times a gauge-fixing Jacobian J', integrated over $n-3$ coordinates and mulitiplied by the appropriate normalization factors. From the three-tachyon amplitude we can deduce that $J = |y_{12}y_{13}y_{23}|$. The normalization factors work out to $g_{\rm o}$ for each tachyon vertex operator, and an overall factor of $i/\alpha' g_{\rm o}^2$ (see the Big Book for details). There is one additional feature. We are considering the oriented string, the coordinate group does not include orientation-changing transformations (which would take z to a function of $\bar z$). We therefore cannot change the cyclic ordering, so there are two terms, e.g. $y_1 = 0, y_2 = 1, y_3 = \infty$ and $y_1 = 0, y_3 = 1, y_2 = \infty$. In the present case they are just equal, so we get

$$S_{3\,\rm tachyon(open)}(k_1,k_2,k_3) = (2\pi)^D\delta^D(\textstyle\sum_i k_i) 2ig_{\rm o}/\alpha'\,. \tag{10.13}$$

We will see that the open string coupling $g_{\rm o} \propto e^{\Phi/2}$, so it is essentially the square root of the closed string coupling.

Moving to the 4-point amplitude, there are six cyclic orderings. Let us take $y_1 < y_4 < y_2 < y_3$, and get the others by permutation. Choosing $y_1 = 0, y_2 = 1, y_3 = \infty$ as before, this term becomes

$$(2\pi)^D\delta^D(\textstyle\sum_i k_i)\displaystyle\frac{ig_{\rm o}^2}{\alpha'}\int_0^1 dy_4\, y_4^{2\alpha' k_1\cdot k_4}(1-y_4)^{2\alpha' k_2\cdot k_4}\,. \tag{10.14}$$

The integral is Euler's beta function

$$\frac{\Gamma(1+2\alpha' k_1\cdot k_4)\Gamma(1+2\alpha' k_2\cdot k_4)}{\Gamma(-2\alpha' k_3\cdot k_4)} = \frac{\Gamma(-1-\alpha' u)\Gamma(-1-\alpha' t)}{\Gamma(2+\alpha' s)}\,. \tag{10.15}$$

The other orderings give the permutations of s, t, u, three distinct terms twice each. Each term has poles in two channels, at the positions of the open string masses-squared. The total Veneziano amplitude is $(2\pi)^D\delta^D(\sum_i k_i)$ times

$$\frac{2ig_{\rm o}^2}{\alpha'}\left(\frac{\Gamma(-1-\alpha' u)\Gamma(-1-\alpha' t)}{\Gamma(2+\alpha' s)} + \frac{\Gamma(-1-\alpha' t)\Gamma(-1-\alpha' s)}{\Gamma(2+\alpha' u)} + \frac{\Gamma(-1-\alpha' s)\Gamma(-1-\alpha' u)}{\Gamma(2+\alpha' t)}\right)\,. \tag{10.16}$$

Gauge amplitudes and Chan-Paton factors.

At the next level of the open string, the vertex operator for a massless vector of polarization e_μ is

$$e_\mu \colon \dot{X}^\mu e^{ik\cdot X} \colon . \tag{10.17}$$

The dot indicates a derivative with respect to the y coordinate along the boundary. The excitation of the first harmonic of the open string translates into a first derivative of X^μ, as discussed before. We are discussing Neumann boundary conditions, where the normal derivative vanishes, so it must be a tangent derivative as indicated. When we get to D-branes, the tangent derivative vanishes for the Dirichlet coordinates, and the normal derivative appears in the massless vertex operator.

The polarization must satisfy $k \cdot e = 0$, as we have discussed, and if $e_\mu \propto k_\mu$ the amplitude should vanish, in order to satisfy the equivalence relation required by the covariant quantization. We can see right away why this is,

$$k_\mu \colon (\partial_y X^\mu) e^{ik\cdot X} \colon = -i\partial_y \colon e^{ik\cdot X} \colon \tag{10.18}$$

is a total derivative, so its amplitudes vanish when integrated around the boundary of the disk. The null states obtained from L_{-1} all work this way, for the higher L_{-n} it's a little more complicated but it still works: by a contour argument (chap. 9.1) one can move these from the null vertex operator over to the other physical state vertex operators, where they become lowering operators and give zero.

The simplest amplitude would be one massless vector and two open string tachyons,

$$\int \mathcal{D}X \, e^{-S_P} \colon e \cdot \dot{X} e^{ik_1\cdot X(y_1)} \colon \colon e^{ik_2\cdot X(y_2)} \colon \colon e^{ik_3 \cdot X(y_3)} \colon . \tag{10.19}$$

The exponentials give a factor of $\prod_{i<j} |y_i - y_j|^{2\alpha' k_i \cdot k_j}$ as before. The additional X can contract with either the second or third exponential, giving a factor of

$$-2i\alpha' e_\mu \partial_{y_1} (k_2^\mu \ln|y_{12}| + k_3^\mu \ln|y_{13}|) = -2i\alpha' \left(\frac{e\cdot k_2}{y_{12}} + \frac{e \cdot k_3}{y_{13}} \right) = -2i\alpha' \frac{e\cdot k_2 y_{13} + e\cdot k_3 y_{12}}{y_{12} y_{13}} . \tag{10.20}$$

Now, the physical state condition gives us $0 = e\cdot k_1 = -e\cdot k_2 - e \cdot k_3$, or $e \cdot k_2 = -e\cdot k_3 = e\cdot (k_2 - k_3)/2$. So we can write everything as

$$-i\alpha' e \cdot (k_2 - k_3) \frac{y_{23}}{y_{12} y_{13}} . \tag{10.21}$$

In addition we have the following additional factors as usual:

$$(2\pi)^D \delta^D(\textstyle\sum_i k_i) \, ,$$

$$
\begin{aligned}
& g_o^2 \frac{-ig_o}{\sqrt{2\alpha'}} \frac{i}{\alpha' g_o^2}\,, \\
& |y_{12} y_{13} y_{23}|\,, \\
\prod_{i<j} |y_i - y_j|^{2\alpha' k_i \cdot k_j} = {} & |y_{23}|^{-2}\,.
\end{aligned} \tag{10.22}
$$

In the second line, I have introduced the new information that the massless vertex operator normalization is $-ig_o/\sqrt{2\alpha'}$. The factor of i is required by unitarity (in particular $\mathrm{Im}(T) = 0$ for the three-point function), the sign is a convention; also, $e \cdot e = 1$. This can be obtained from the state-operator mapping (it conforms to the rule that the open string CFT is simple when $\alpha' = \frac{1}{2}$) or by comparing with the pole in the four-point function (which you are working out in the homework). In the fourth line I have used the mass-shell conditions for the vector and two tachyons to evaluate the momentum dot products, $k_1 \cdot k_2 = k_1 \cdot k_3 = 0, k_1 \cdot k_2 = -2/\alpha'$.

As it should, the choice of the arbitrary coordinates y_i drops out — almost. First if $y_1 < y_2 < y_3$, or any cyclic permutation of this, the various y-dependent terms multiply out to -1 and we get

$$
-i(2\pi)^D \delta^D(\textstyle\sum_i k_i) \dfrac{g_o}{\sqrt{2\alpha'}}\, e \cdot (k_2 - k_3)\,. \tag{10.23}
$$

We should check gauge invariance: replacing the polarization of the photon vertex operator with its momentum k_1 gives zero as it should. Indeed, you should recognize this as the form of the vector-scalar-scalar vertex in scalar electrodynamics.

However, as with the 3-tachyon amplitude we must add separately the other cyclic ordering, for which the -1 becomes $+1$ and everything cancels. There is a simple interpretation to this: the photon couples to the endpoints of the open string [draw], and one endpoint has charge $+1$ and the other -1 so the total charge is zero.

Now is a good time to return to something I mentioned rather quickly at the end of chap. 3, the Chan-Paton degrees of freedom. One way to motivate these is historically: in string theory's first incarnation, open strings were supposed to be mesons, which we now interpret as a quark and antiquark, which are roughly speaking at the ends of a tube of color flux (a tube because it can't spread due to confinement). Now, the q and $\bar{q}$ can be u, d, or s (for the light quarks then known), so we need an index to distinguish the different kind of mesons, one at each end. So open string states would be denoted

$$
|\text{oscillator state}, k; I, J\rangle\,, \tag{10.24}
$$

where I, J take values u, d, s or $1, 2, 3$ (and we'll generalize to $1, \ldots, n$). We are no longer interpreting the string this way, but it's interesting to see if the idea is still viable. If we we trying to deal with real mesons we'd have some kind of mass term at the endpoints, since these quarks have different masses, but conformal invariance doesn't allow this. In fact the only reasonable Hamiltonian for these degrees of freedom is zero, which certainly preserves conformal invariance. In other words, whatever state the endpoint is in it just stays in that state. This sounds rather trivial, but in fact these states make all the difference.

One might also have hit on this idea just by asking, what is the most general Poincaré invariant and conformally invariant thing we can do? The modern interpretation, as I described before, is this: the open string has to end on something, a spacetime-filling D-brane. But there can be more than one D-brane on top of each other, so the Chan-Paton degree of freedom labels which D-brane it ends on.

Now, in order for an interaction to happen, the right endpoint of each string must be in the same state as the left endpoint of the next one [draw]. If each string has a Chan-Paton wavefunction $\lambda^{(i)}_{IJ}$ then the amplitude has an additional factor

$$\lambda^{(1)}_{IJ}\lambda^{(2)}_{JK}\lambda^{(3)}_{KI} = \mathrm{Tr}(\lambda^{(1)}\lambda^{(2)}\lambda^{(3)}) \tag{10.25}$$

if the cyclic order is 123, and so on. It is useful to introduce a complete basis λ^a of $n \times n$ matrices, and unitarity requires them to be Hermitean, so we would then write this as

$$\mathrm{Tr}(\lambda^{a_1}\lambda^{a_2}\lambda^{a_3})\,. \tag{10.26}$$

But now we can go back and the two cyclic orderings no longer cancel. Instead, we get

$$-i(2\pi)^D\delta^D(\textstyle\sum_i k_i)\dfrac{g_o}{\sqrt{2\alpha'}}\, e\cdot(k_2-k_3)\mathrm{Tr}(\lambda^{a_1}[\lambda^{a_2},\lambda^{a_3}])\,. \tag{10.27}$$

This is exactly the interaction of a gauge boson with two scalars in the adjoint representation. The gauge group is $U(n)$, generated by the Hermitean $n \times n$ matrices. Actually we should separate into traceless matrices and the identity, which close separately, $U(n) = SU(n) \times U(1)$. The $U(1)$ interaction is zero here because the commutator vanishes. From the world-sheet point of view $U(n)$ is a global symmetry, but this shows the general phenomenon that global symmetries of the world-sheet become gauge symmetries in spacetime. It is a general principle, still without exceptions, that string theory does not give rise to exact continuous global symmetries.

Incidentally, if we go back to the Veneziano amplitude, we had six terms before, which collected into three forms, and now each is multiplied by the trace in a different order.

We can now go on to the three-gauge boson amplitude

$$-\frac{g_o}{\sqrt{8\alpha'^5}}|y_{12}y_{13}y_{23}|\int \mathcal{D}X\, e^{-S_P} :e_1\cdot\dot{X}e^{ik_1\cdot X(y_1)}::e_2\cdot\dot{X}e^{ik_2\cdot X(y_2)}::e_3\cdot\dot{X}e^{ik_3\cdot X(y_3)}:\,. \tag{10.28}$$

In addition, we must include the Chan-Paton factors and sum over cyclic orderings. Each $\dot{X}$ can be contracted with an exponential or with one of the other $\dot{X}$, so there are a number of terms to collect. There are two types of terms: those where all three $\dot{X}$ contract with exponentials, which are of order k^3, and those where two contract with each other and the third with an exponential, of order k^1. In the end, the amplitude is $i\frac{g_o}{\sqrt{2\alpha'}}(2\pi)^D\delta^D(\sum_i k_i)$ times

$$e_1\cdot(k_2-k_3)e_2\cdot e_3+e_2\cdot(k_3-k_1)e_3\cdot e_1+e_3\cdot(k_1-k_2)e_1\cdot e_2+\frac{\alpha'}{2}e_1\cdot(k_2-k_3)e_2\cdot(k_3-k_1)e_3\cdot(k_1-k_2)\,. \tag{10.29}$$

Again the two cyclic orderings enter with opposite signs to give $\mathrm{Tr}(\lambda^{a_1}[\lambda^{a_2},\lambda^{a_3}])$.

One can check that if one replaces any of the polarizations e_i with the corresponding momentum k_i this vanishes, and in fact this full determines the form, up to the separate normalizations of the k^1 and k^3 terms. Thus, the three terms of order k^1 are precisely the three gauge boson amplitude from non-Abelian gauge theory. Since these states are massless we should be able to describe their long-distance physics by an effective actions, and indeed the amplitude we have found is what would be obtained from a spacetime action

$$\int d^Dx\left(-\frac{\alpha'}{2g_o^2}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu})-\frac{4i\alpha'^2}{3g_o^2}\mathrm{Tr}(F_\mu{}^\nu F_\nu{}^\omega F_\omega{}^\mu)\right)\,. \tag{10.30}$$

The first term is the usual action for non-Abelian gauge theory, which would be renormalizable in four dimensions (of course here $D=26$). The second term is suppressed at length scales long compared to $\sqrt{\alpha'}$. This is as we expect for effective actions, all terms allowed by symmetry, with the underlying length scale appearing as required by dimensional analysis. Normally this is an infinite sum, but for the 3-point amplitude only two distinct terms contribute. Effective actions always contain nonrenormalizable terms, but the resulting divergences are cut off by the physics at the higher scale, as we will see in the next chapter.

Gravitational amplitudes

After all of the above, the gravitational amplitudes are almost no work. The massless closed string vertex operator is

$$\frac{2g_c}{\alpha'}e_{\mu\nu}\!:\partial_z X^\mu\partial_{\bar z}X^\nu:\,. \tag{10.31}$$

The physical state conditions are $k^\mu e_{\mu\nu}=k^\nu e_{\mu\nu}=0$, and the null state equivalence relation is

$$e_{\mu\nu}\sim e_{\mu\nu}+a_\mu k_\nu+k_\mu b_\nu \tag{10.32}$$

for arbitrary a, b. By separating $e_{\mu\nu}$ into traceless symmetric, antisymmetric, and identity pieces we get the graviton, two-form, and dilaton amplitudes, but we can do them all at once.

Contractions between z-derivatives and $\bar{z}$-derivatives vanish because

$$\partial_z \partial_{\bar{z}'} \ln |z - z'|^2 = \partial_z \frac{1}{\bar{z}' - \bar{z}} = -2\pi\delta^2(z - z') \,. \tag{10.33}$$

Such a delta function never contributes because the terms from the exponentials give zero as $z \to z'$; recall that we are analytically continuing from the convergent regime. Because of this vanishing, the calculation factorizes into two copies of the open string amplitude, one from z and one from $\bar{z}$. The result for a massless tensor and two tachyons is then

$$-i\pi g_{\mathrm{c}} e_{1\mu\nu}(k_2 - k_3)^\mu (k_2 - k_3)^\nu \,. \tag{10.34}$$

Similarly, for the three-tensor amplitude,

$$i\pi g_{\mathrm{c}} e_{1\mu\nu} e_{2\sigma\rho} e_{3\lambda\omega} T^{\mu\sigma\lambda} T^{\nu\rho\omega} \,, \tag{10.35}$$

where

$$T^{\mu\sigma\lambda} = (k_2 - k_3)^\mu \eta^{\sigma\lambda} + (k_3 - k_1)^\sigma \eta^{\lambda\mu} + (k_1 - k_2)^\lambda \eta^{\mu\sigma} + \frac{\alpha'}{8}(k_2 - k_3)^\mu (k_3 - k_1)^\sigma (k_1 - k_2)^\lambda \,, \tag{10.36}$$

is essentially the three-photon coupling, with a slightly different normalization for the k^3 term. This amplitude contains k^2, k^4, and k^6 terms.

Let us focus first on the k^2 terms, which are the most important at long distance. These can be obtained from the effective action

$$\frac{1}{8\pi^2 g_{\mathrm{c}}^2} \int d^{26}x \sqrt{-G} e^{-2\Phi} \left(R - \frac{1}{12} H_{\mu\nu\lambda} H^{\mu\nu\lambda} + 4\partial_\mu \Phi \partial^\mu \Phi \right) \,, \tag{10.37}$$

expanding around

$$\Phi(x) = 0 \,, \quad G_{\mu\nu}(x) = \eta_{\mu\nu} \,. \tag{10.38}$$

There is much to say here. First, the first term is the Einstein-Hilbert action for general relativity, if we set $\Phi = 0$. In $D = 4$ the coefficient is $1/16\pi G_{\mathrm{Newton}}$, but the 16π is very particular to $D = 4$, so it has become conventional to use the normalization $1/2\kappa^2$ in any number of dimensions. So

$$\kappa = 2\pi g_{\mathrm{c}} \,. \tag{10.39}$$

Now, κ has units of length$^{(D-2)/2}$, so $\kappa/\alpha'^{(D-2)/4}$ is some sort of dimensionless coupling.

The second thing to notice is that the action is multiplied by an overall $e^{-2\Phi}$. There is a reason for this, which we will get to when we study strings in curved spacetime (right after one-loop amplitudes), but for now let's just note an important consequence that the dimensionless coupling that we have just noted is not fixed, by a field redefinition $\Phi(x) \to \Phi(x) + \Phi_0$ we can change it. Equivalently we can expand around a different solution

$$\Phi(x) = \Phi_0\,, \quad G_{\mu\nu}(x) = \eta_{\mu\nu}\,. \tag{10.40}$$

So there is no dimensionless coupling in the theory, but different vacua can have different effective couplings. By the way, the S-matrix only determines the action up to field redefinitions, and so we can write this in different ways (in particular removing the Φ-dependence from the Einstein-Hilbert action), but we will save this discussion for later.

The third thing is the kinetic term for the antsymmetric tensor field, built from

$$H_{\mu\nu\lambda} = \partial_\mu B_{\nu\lambda} + \partial_\nu B_{\lambda\mu} + \partial_\lambda B_{\mu\nu}\,, \tag{10.41}$$

where $B_{\mu\nu}$ and $H_{\mu\nu\lambda}$ are each fully antisymmetric (i.e. they are forms, and $H = dB$). This has a gauge invariance

$$B_{\mu\nu}(x) \to B_{\mu\nu}(x) + \partial_\mu \zeta_\nu(x) - \partial_\nu \zeta_\mu(x)\,, \tag{10.42}$$

which we can see in the antisymmetric part ($a = -b$) of the equivalence relation (10.32).

It's interesting, and not obvious from the Lagrangians, that the three-graviton vertex from general relativity is the square of the Yang-Mills vertex. This factorization property, which we've seen in the spectrum and three-point amplitudes, extends to the higher amplitudes as well. It's a bit nontrivial: the integral over the complex plane doesn't immediately look like the square of the integral along the real axis, but after rotating some contours it works. This is a very clever observation by Kawai, Lewellen, and Tye, and it turns out that it projects down to the field theory and has been valuable in the recent exploration of gravity and supergravity loop amplitudes.

Finally, the k^4 and k^6 terms mean that there are Riemann-squared and Riemann-cubed terms in the action, as well as higher powers of H. Incidentally, in supersymmetric theories some of these terms are forbidden, in the most symmetric case the first correction is a four-Riemann term.

11 Loop amplitudes

Our main goal in studying the loop divergences is seeing what happens to the UV infinities, but along the way we will encounter some important CFT concepts, such as modular invariance, as well as an interesting connection between open and closed strings. We will only need to look at the simplest loop amplitudes: one loop of closed strings (the torus) and one loop of open strings (the annulus) [draw].

Riemann surfaces

All possible topologies for two-dimensional oriented surfaces are easily described: the consist of spheres with h handles attached and b holes (boundaries) drilled in them. Thus the disk is a sphere with one hole, an annulus is a sphere with two holes, and so on. For theories with closed strings only we'd have handles but no holes. The Euler number χ is

$$\chi = 2 - 2h - b\,. \tag{11.1}$$

This is -2 for the sphere, -1 for the disk, 0 for the torus and annulus, and negative for higher order surfaces. On a surface without boundary, this can be written as a curvature invariant,

$$\chi = \frac{1}{4\pi}\int d^2\sigma\sqrt{g}R\,. \tag{11.2}$$

With boundaries, the definition includes a boundary integral as well (ex. 1.3 in the Big Book), whose form is determined by the requirement that χ is a topological invariant. The Euler number enters in a number of ways: it governs the string coupling dependence (which we'll get to in the next chapter) and it enters through the Reimann-Roch (RR) theorem.

The RR theorem has to do with the zero modes of various differential operators, and for us the relevance is this. After we fix the coordinate and Weyl invariances by choosing the metric, we may have a few additional gauge transformations to fix as we've seen for the sphere and disk. In addition, it may not be possible to fully fix the metric, rather one can fix it to some form depending on some parameters, and these must still be integrated over. The RR theorem states that *the number of unfixed gauge transformations minus the number of (real) parameters remaining in the metric is equal to* 3χ. For the sphere this is 6, as we found (3 complex = 6 real). For the disk it was 3. For the torus and annulus it is zero, but not so trivially, in that the separate numbers are nonzero.

The torus

Since $\chi = 0$ the average curvature is zero, and it is a theorem that by a coordinate and Weyl transformation we can bring it everywhere to unit form, or in complex coordinates to the form $dwd\bar{w}$. However, to do this we have to allow the coordinate region to vary: the periodicities are

$$w \equiv w + 2\pi \equiv w + 2\pi\tau \tag{11.3}$$

for some complex parameter τ [draw: fig. 5.1]. We can assume that $\operatorname{Im}\tau > 0$, since τ and $-\tau$ generate the same lattice (also, $\operatorname{Im}\tau = 0$ would not generate a torus). In this description it looks like τ is not in the metric but in the periodicity, but if we define $w = \sigma^1 + \tau\sigma^2$ then σ^1 and σ^2 both have fixed periodicity 2π and the parameter τ gets put into the metric,

$$dwd\bar{w} = (d\sigma^1)^2 + 2\tau_1 d\sigma^1 d\sigma^2 + (\tau_1^2 + \tau_2^2)(d\sigma^2)^2\,, \quad \tau = \tau_1 + i\tau_2\,. \tag{11.4}$$

The way I've stated RR corresponds to the latter. This is two real parameters in the metric, and it's obvious what are the two unfixed coordinate transformations, just translations of the origin of coordinates. The main point is that after gauge fixing we will still have an integral over τ.

The range of the τ integral is important. First, τ and $\tau + 1$ generate the same group of identifications. Second, τ and $-1/\tau$ generate the same group of identifications: if I define $w' = -w/\tau$ then the identifications above become

$$w' \equiv w' - 2\pi/\tau \equiv w' - 2\pi\,. \tag{11.5}$$

These two transformations generate the modular group:

$$\tau \to \frac{a\tau + b}{c\tau + d}\,, \quad a, b, c, d \in \mathbf{Z}\,, \quad ad - bc = 0\,. \tag{11.6}$$

This looks like the Möbius transformations but with integer parameters.

In order to avoid overcounting, restricting $|\tau| \leq 1$ fixes the first transformation and $|\tau| > 1$ fixes the second. This leaves an integral over the fundamental region, with edges identified [draw: fig. 5.2].

But there's one more wrinkle: we haven't fixed all the gauge symmetry. If we shift $w \to w + u$ for any complex constant u, the metric and periodicity remain the same. So we will have to fix one vertex operator position (or, as you'll see, deal with this another way). This is an illustration of the general theorem: 2 real parameters in the metric minus two extra gauge symmetries $= 3 \times 0$.

By the way, there is a nice analogy with the point particle, summing over all paths with the topology of a circle. We had the field $\eta(\tau) \sim \sqrt{-\gamma_{\tau\tau}}$, and the invariant length of the path is$\int d\tau\, \eta(\tau) = l$, which of course we can't change by a coordinate choice. So if we go to the gauge $\eta = 1$ we have to let the range of the τ coordinate vary, $\tau \equiv \tau + l$. Or we can go to the gauge $\eta = l$, and the coordinate range is fixed as $\tau \equiv \tau + l$. Either way, after modding out the coordinate transforms there is one parameter left to integrate.

$\int \mathcal{D}X$ on the torus

According to our general formalism, the one-loop amplitude will be

$$\int d^2\tau\, J(\tau) \int \mathcal{D}X\, e^{-S_P(\tau)} \int d^{2(n-1)}\sigma \prod_{i=1}^{n} \sqrt{g(\sigma_i)} \mathcal{V}_i(\sigma_i)\,. \tag{11.7}$$

I've fixed one vertex operator position to get rid of last coordinate symmetry. Now, here's a nice surprise: we can learn everything that we need to from this amplitude, including some unexpected things, purely from the vacuum case $n = 0$. Of course the higher amplitudes are interesting and we'll say a few words about them later.

Again we can evaluate the X path integral either by operator or path integral methods, both are developed in chap. 7.[9] We will stick to operator methods here. In the $dwd\bar{w}$ coordinates, we can think of the torus as formed by taking a closed string, evolving it for a Euclidean time $2\pi\tau_2$, twisting by an angle $2\pi\tau_1$, and then tracing, setting it equal to its initial state and summing over states:

$$\int \mathcal{D}X \, e^{-S_P(\tau)} = \mathbf{Tr}\left(e^{2\pi i\tau P - 2\pi\tau H}\right) \,, \tag{11.8}$$

where H is the world-sheet Hamiltonian,P is the world-sheet translation operator, and the trace runs over all states of the world-sheet CFT; I made it boldface to distinguish it from the Chan-Paton trace. Now, H and P must be constructed from integrals of the world-sheet energy-momentum tensor, so must work out to be Virasoro generators, and in fact

$$H = L_0 + \tilde{L}_0 - \frac{D}{12} \,, \quad P = L_0 - \tilde{L}_0 \,. \tag{11.9}$$

This is derived in 2.6.10 of the text, and is mostly a matter of tracing through the definitions and conventions. The one tricky thing is the $D/12$ term, which comes about because T_{zz} is not a tensor, so picks up a piece in going from the z frame to the w frame. But rather than repeat the demonstration in the book, I will just assert it, and we will see in a little while that this constant must be there for consistency. So, defining $q = e^{2\pi i\tau}$ we have

$$\int \mathcal{D}X \, e^{-S_P(\tau)} = (q\bar{q})^{-D/24}\mathbf{Tr}\left(q^{L_0}\bar{q}^{\bar{L}_0}\right) \,. \tag{11.10}$$

The states in the CFT Hilbert space are labeled by their momenta k and their oscillator levels $N_{m\mu}, \tilde{N}_{m\mu}$, so we can write the trace

$$\mathbf{Tr} = V \int \frac{d^D k}{(2\pi)^D} \prod_{\mu=0}^{D-1} \prod_{m=1}^{\infty} \left(\sum_{N_{m\mu}=0}^{\infty} \sum_{\tilde{N}_{m\mu}=0}^{\infty} \right) \,. \tag{11.11}$$

The factor of the spacetime volume V comes from putting the system in a big box and approximating the momentum sum by an integral. Also, since $L_0, \tilde{L}_0$ are sums their exponentials are products, and we get

$$\int \mathcal{D}X \, e^{-S_P(\tau)} = (q\bar{q})^{-D/24} V \int \frac{d^D k}{(2\pi)^D} (q\bar{q})^{\alpha' k^2/4} \prod_{\mu=0}^{D-1} \prod_{m=1}^{\infty} \left(\sum_{N_{m\mu}=0}^{\infty} q^{mN_{m\mu}} \sum_{\tilde{N}_{m\mu}=0}^{\infty} \bar{q}^{m\tilde{N}_{m\mu}} \right)$$

[9]Remarkable: it's week 6 and we're in chapter 7. This is not so much due to a breakneck pace (I hope) as to the judicious omission of unnecessary information.

$$= iV(4\pi^2\alpha'\tau_2)^{-D/2}(q\bar{q})^{-D/24}\prod_{m=1}^{\infty}(1-q^m)^{-D}(1-\bar{q}^m)^{-D}\,. \tag{11.12}$$

Note that the sums converge because $|q| < 1$. There's one subtlety: the k^0 integral does not converge because it's the wrong-sign gaussian, so we have to make the rule that it is defined by contour rotation, which by the way gives the i required by unitarity.[10]

Modular invariance and ghosts

The result can be compactly expressed in terms of the Dedekind η function

$$\eta(\tau) = e^{i\pi\tau/12}\prod_{m=1}^{\infty}(1-q^m)\,, \tag{11.13}$$

i.e.

$$\int \mathcal{D}X\, e^{-S_P(\tau)} = iV(4\pi^2\alpha'\tau_2|\eta(\tau)|^4)^{-D/2}\,. \tag{11.14}$$

Now, there is an important check we need to make. We have seen that $\tau+1$ and $-1/\tau$ describe the same metric as τ, and so the path integral should be the same for all. This is easy for $\tau' = \tau+1$ because q is invariant, so $\eta(\tau') = e^{i\pi/12}\eta(\tau)$ and the phase cancels out in the absolute value. For $\tau' \to -1/\tau$, we need the identity

$$\eta(\tau') = (-i\tau)^{1/2}\eta(\tau)\,. \tag{11.15}$$

Also

$$\mathrm{Im}(-1/\tau) = \mathrm{Im}(-\tau^*/|\tau|^2) = \tau_2/|\tau|^2\,, \tag{11.16}$$

so it follows that $\tau_2|\eta(\tau)|^4$ is invariant. By the way, this would not work if we did not have the constant $-D/12$ in Eq. (11.9): we could have deduced it by being careful form the start, but these things are overdetermined and modular invariance is a powerful constraint.

Now, I am going to make a guess as to what the determinant is, namely

$$J(\tau) = \frac{4\pi^2}{2\cdot 8\pi^2\tau_2}|\eta(\tau)|^4 \tag{11.17}$$

The factor of $|\eta(\tau)|^4$ just cancels the effect of the two extra sets of oscillators, and can be obtained from a sum over Faddeev-Popov ghost states. The factor of $8\pi^2\tau_2$ is the area

[10] By the way, in the previous class the question arose whether there is a Lorentzian form for the amplitudes that converges without analytic continuations. I think that the light-cone framework pioneered in S. Mandelstam, Nucl. Phys. B64: 205-235, 1973 would probably accomplish this. I do not know of any manifestly covariant Lorentzian description.

of the torus, $\int dw d\bar{w}$ and corrects for the extra coordinate freedom,[11] which we have left unfixed because we have no vertex operators. The factor of 2 is from one more coordinate overcounting $\sigma \to -\sigma$. We'll be able to confirm these normalizations by comparison with quantum field theory. Unfortunately I can't give a simple argument for the $4\pi^2$ in the numerator, it comes from the determinant for the metric and the width of the strings (the 2π choice for the periodicity was arbitrary, after all, it can be changed by rescaling the coordinate). But we will give a physics derivation of this soon.

In all we have the the torus amplitude without vertex operators,

$$iV \int \frac{d^2\tau}{4\tau_2} (4\pi^2\alpha'\tau_2)^{-13} |\eta(\tau)|^{-48} \,. \tag{11.18}$$

I've set $D = 26$ for simplicity, but we'll see below that this amplitude is consistent only in this dimension. The τ integral runs over the fundamental region.

Comparison with field theory

The best way to get a feel this is to look at the corresponding amplitude in quantum field theory. First, if we sum over all paths connecting two points x and y, you would not be surprised if we got a propagator

$$\int \frac{d^D k}{(2\pi)^D} e^{ik\cdot(x-y)} \int_0^\infty dl\, e^{-l(k^2+m^2)} = \int \frac{d^D k}{(2\pi)^D} \frac{e^{ik\cdot(x-y)}}{k^2+m^2} \,. \tag{11.19}$$

Now suppose that we have a closed loop, analogous to the torus. This should contribute

$$V \int \frac{d^D k}{(2\pi)^D} \int_0^\infty \frac{dl}{2l} e^{-l(k^2+m^2)} = iV \int_0^\infty \frac{dl}{2l} (4\pi l)^{-D/2} e^{-lm^2} \,, \tag{11.20}$$

where the factor $2l$ accounts for coordinate overcounting as on the torus. To make this look more familiar, let us do the l integral before the k integral. This diverges logarithmically, as $l \to 0$ so let's put a small lower cutoff ϵ on l, to get

$$\begin{aligned} -\frac{V}{2} \int \frac{d^D k}{(2\pi)^D} \ln[\epsilon(k^2+m^2)] &= -\frac{1}{2} \mathrm{Tr} \ln[\epsilon(-\partial^2+m^2)] \\ &= -\frac{1}{2} \ln\det[\epsilon(-\partial^2+m^2)] \,. \end{aligned} \tag{11.21}$$

Now, this is a single closed loop. If we sum over any number of such loops the sum exponentiates to give

$$\det{}^{-1/2}[\epsilon(-\partial^2+m^2)] \,. \tag{11.22}$$

[11] Recall our convention $dw d\bar{w} = 2 d\mathrm{Re}w\, d\mathrm{Im}w$, and similarly for $d\tau d\bar{\tau}$, though the 2's cancel between these.

You might recognize this as the path integral over a scalar field,

$$\int \mathcal{D}\phi\, e^{-\frac{i}{2}\int d^Dx\,(\partial_\mu\phi\partial^\mu\phi+m^2\phi^2)}\,, \tag{11.23}$$

where ϵ comes in as a regulator from normalizing $\mathcal{D}\phi$. So we have connected the path sum with the usual field integral.

This overall normalizing factor has the interpretation of the vacuum to vacuum amplitute

$$\langle 0|e^{-i\int d^Dx\,\mathcal{H}}|0\rangle = e^{-i\rho_0 V} \tag{11.24}$$

where ρ_0 is the vacuum energy. The dependence on ϵ reflects a dependence on unknown high energy physics.

The l-integral above diverges at small l, and this is a UV divergence. What we will see is that the string amplitude is very similar in form to the particle amplitude, but with a UV cutoff built in.

I claim

Let us expand at large τ_2, meaning small q. We have

$$\eta(\tau) = q^{1/24}(1-q+\ldots)\,, \tag{11.25}$$

and so

$$|\eta(\tau)|^{-48} = (q\bar{q})^{-1} + 24q^{-1} + 24\bar{q}^{-1} + 24^2 + \ldots\,. \tag{11.26}$$

Now, if we change variables to

$$\pi\alpha'\tau_2 = l\,, \tag{11.27}$$

then

$$(q\bar{q})^{-1} = e^{4l/\alpha'} = e^{-M_0^2 l} \tag{11.28}$$

where M_0^2 is the tachyon mass-squared. So the leading behavior of the string amplitude is exactly the same as the field theory amplitude with the mass-squared. The q^{-1} and $\bar{q}^{-1}$ give zero under $\int d\tau_1$, so the next term is the 24^2 which is just the contribution of the massless states.

By the way, in D dimensions the expansion

$$|\eta(\tau)|^{-2(D-2)} = (q\bar{q})^{-(D-2)/24}[(1+2(D-2)(q+\bar{q})+(D-2)^2q\bar{q}+\ldots]\,, \tag{11.29}$$

and the states at the second level only come out massless if $D = 26$. So we only get consistency between the known masslessness of these states (from the tensor property of the vertex operator) and the requirements of modular invariance if $D = 26$. This is very close to how Lovelace first discovered the need for a critical dimension (to be precise, he was looking at the analogous result for the annulus).

So, the string calculation looks like the sum of the field theory calculations from the individual particles, *except* for the range of integration: the would-be UV divergent region $\tau_2 \to 0$ is cut off. Further, the cutoff does not spoil anything else, like unitarity or the decoupling of the null states. If we had cut off by hand the l integral in field theory these would fail, but here the cutoff is innocuous because the cut edges of the fundamental region are identified. The only true asymptotic region is $\tau_2 \to \infty$ ($\sim l \to \infty$), which we recognize as an IR region.

Other limits

Introducing vertex operators is not too hard, either in the path integral or operator form. Their correlators involve *ϑ-functions,* which respect the periodicities of the torus. But we just want to understand the fate of the UV divergences, which we can see pictorially.

Let's consider the torus with four vertex operators, which is the string version of the divergent gravitational loop graph with which we began. The amplitude is an integral over the vertex operator positions and over the modulus τ, and the question is where divergences arise. We've seen that the only asymptotic region for τ has an IR interpretation. The other potentially dangerous region is when two or more vertex operators come together. This appears to correspond to the problematic region of the Feynman graph, where all the interactions come together in spacetime, but it's actually very different.

Let's start with the case of two vertex operators $\mathcal{V}_{1,2}$ approaching on another. Locally this is the same as on the sphere, and we know what happens there: we do get a divergence, which produce poles in $(k_1 + k_2)^2$. But these poles are an IR effect, a resonance between $k_1 + k_2$ and the mass of the intermediate string. Thus this is an IR divergence. We can make this more intuitive by a conformal transformation, which keeps the vertex operators at constant separation but produces a long tube between them and the rest of the amplitude [fig. 9.7].

This works for any number of vertex operators approaching, but the case that all four approach is interesting. Now we get poles in $(k_1 + k_2 + k_3 + k_4)^2$, but this is zero by momentum conservation, so we are forced to sit on top of the pole. The interpretation is that the one-loop vacuum energy discussed above produces a term $-\rho_0\sqrt{-G}$ in the spacetime Lagrangian, which in perturbation theory $G_{\mu\nu} = \eta_{\mu\nu} = h_{\mu\nu}$ has an amplitude to produce a zero momentum graviton which connects on to the rest of the amplitude. The problem is that we are expanding around the wrong vacuum, flat spacetime is no longer a solution due to ρ_0, and the cure, as explained by Fischler and Susskind, is to perturb the background to a solution. This illustrates the general principle that IR divergences come about because we have asked the wrong question, and are removed by figuring out what is the right question.

This is all there is to it: all possible divergent regions of integration space are IR in nature: one can break the world-sheet up into vertices and propagators, in such a way that l never goes to zero for any propagator, and all potentially divergent regions are $l \to \infty$.

By the way, for supersymmetric theories one expects that flat spacetime is an exact solution, so the IR divergences should cancel without shifting the background. There are formal arguments that this is true, but it is a bit of an embarrassment that the technology has not been developed adequately to make this explicit. Roughly speaking, the breaking up of the world-sheet into pieces takes us off shell, and off-shell supersymmetry is much more complicated than on-shell. The main research effort in this area uses the 'pure spinor' formalism, but I wonder whether some of the modern on-shell methods used in gauge theory might be more effective.

The annulus

We can describe the annulus as the region

$$0 < \mathrm{Re}\, w < \pi\,, \quad w \sim w + 2\pi i t\,, \tag{11.30}$$

where the real parameter t cannot be removed by a coordinate transformation. The RR theorem tells us to expect one unfixed coordinate transformation, which is the translation in the imaginary direction. There is no modular group, the range of t is 0 to ∞, very similar to l for the circle, and the rest of the calculation is very similar to the torus, with just one set of operators, yielding the amplitude

$$\int_0^\infty \frac{dt}{2t}(8\pi^2\alpha' t)^{-13}\eta(it)^{-24}\,. \tag{11.31}$$

Again, $t \to \infty$ is dominated by the lightest string states and is an IR region. Now, however, nothing excludes the region $t \to 0$. In fact there is a divergence here, but it is again an IR divergence. Pictorially, this region looks like a *closed* string appearing from the vacuum, propagating a distance, and disappearing. Indeed by a modular transformation one can rewrite the amplitude (11.31), derived from loops of open strings, as a sum over closed string poles, with a divergence from the massless pole.

As before, the cure is to expand around a corrected solution to the field equations, and in supersymmetry theories (the Type I theory in this case) the divergence just cancels because there is no net amplitude for a massless closed string to appear from the vacuum. Historically (and logically) the annulus is important because it shows that (i) even if we start from just open strings, we must get closed strings, and (ii) only for $D = 26$ does one get poles after the modular transformation, other values of D give branch cuts with no physical interpretation (the calculation is similar to the torus discussion, but this is where Lovelace actually did it). By the way, these earlier results (i,ii) were originally obtained by looking at the annulus with two vertex operators on each boundary, rather than the vacuum amplitude.

12 String in curved spacetime

(Chap. 3.7 of the text.)

The nonlinear sigma model

Now we come to the next big subject. Let's start with the most naive thing, putting a general curved metric into the Polyakov action,

$$S_P \to \frac{1}{4\pi\alpha'} \int d\tau d\sigma \sqrt{\det g}\, g^{ab} G_{\mu\nu}(X) \partial_a X^\mu \partial_b X^\nu \,. \tag{12.1}$$

The first thing to notice is that things are more complicated: the action is no longer quadratic in X so we don't just have free field theory in 1+1 dimensions, we have interactions. This is known as a *nonlinear sigma model,* for historic reasons that are no longer particularly apt. Now, in QFT you do this by perturbation theory in the coupling, but what is the coupling here? Well, if we expand around a point x_0 in spacetime

$$G_{\mu\nu}(X) = G_{\mu\nu}(X_0) + (X-x_0)^\lambda \partial_\lambda G_{\mu\nu,\lambda}(X_0) + \frac{1}{2}(X-x_0)^\lambda (X-x_0)^\sigma \partial_\lambda \partial_\sigma G_{\mu\nu}(X_0) + \ldots \,, \tag{12.2}$$

and define $X - x_0 = Y$, we see that the Y^3 interaction is proportional to the first derivative of the metric, the Y^4 is proportional to the second derivative, and so on. So the interactions are weak when the metric is slowly varying, and strong when it changes quickly. The loop expansion parameter is essentially to α'/L^2 where L is the typical scale of variation. Equivalently it is α' times the curvature of the spacetime. So there are two expansions, in world-sheet loops (controlled by the curvature) and in string loops (controlled by the coupling).

The action (12.1) is classically Weyl invariant, but again there could be an anomaly. In fact, if we look at the case that the metric is nearly flat, $G_{\mu\nu}(X) = \eta_{\mu\nu} + \chi_{\mu\nu}(X)$, then the perturbation of S_P looks exactly like a graviton vertex operator, which should not be a surprise: roughly speaking the curved spacetime action exponentiates the vertex operators. We know that the vertex operators get quantum corrections to their Weyl transformations, so that in position space we get

$$\Box \chi_{\mu\nu} = \partial^\mu \chi_{\mu\nu} = \eta^{\mu\nu} \chi_{\mu\nu} = 0 \,. \tag{12.3}$$

This is just a linearized and gauge-fixed version of Einstein's equation. In fact, we should get a coordinate invariant equation, because a change $X'(X)$ is just a change of variables in the path integral. This is correct, the condition for this theory to be Weyl invariant is

$$0 = \mathbf{R}_{\mu\nu} + \alpha' O(\mathbf{Riemann}^2)\,. \tag{12.4}$$

(I do part of this calculation in the text, but I don't develop the full world-sheet perturbation theory because it isn't needed later.) I am using boldface for spacetime curvatures, to distinguish them from world-sheet curvatures. Thus we get the vacuum Einstein equation, to first approximation, from the condition for the 1+1 dimensional nonlinear sigma model to be Weyl invariant. (This is all we could get, because at one loop the RHS has to involve two derivatives, and it must be a tensor.) This curious fact is given a physical interpretation by the embedding in string theory, but with higher-derivative corrections. Note the implication, that it only makes sense to discuss strings in spacetimes that satisfy the equations of motion, the theory only wants to deal with physically meaningful questions.

By the way, a rigid scaling of lengths is a special case of a Weyl transformation. This is governed by the β-function of the field theory, so we have

$$\mu \frac{\partial}{\partial \mu} G_{\mu\nu} = -\alpha' \mathbf{R}_{\mu\nu} \tag{12.5}$$

to one loop on the world-sheet. This may seem a bit odd: instead of a coupling constant one has a whole function $G_{\mu\nu}$ (unless there is a lot of symmetry), but it all makes sense. This is known as *Ricci flow* and also plays an interesting role in mathematics. The minus sign, by the way, means that positive curvature is asymptotically free. The n-sphere, for example, becomes free at short distance and strongly coupled at long distance, actually developing a mass gap as in QCD. But this has limited relevance to string theory, since we want the cases where the β-function is zero.

Other massless backgrounds

It is natural to generalize the world-sheet action to

$$S_P \to \frac{1}{4\pi\alpha'} \int d\tau d\sigma \,\sqrt{\det g}\, \left[\left(g^{ab} G_{\mu\nu}(X) + i\epsilon^{ab} B_{\mu\nu}(X)\right) \partial_a X^\mu \partial_b X^\nu + \alpha' R \Phi(X)\right]\,. \tag{12.6}$$

My convention is that ϵ^{ab} is a tensor, not a density, so that $\epsilon^{12} = -\epsilon^{21} = 1/\sqrt{\det g}$. This is the most general action with two derivatives (though in the last term these act on the metric, not on X). It depends on the three fields G, B, Φ which appear to correspond to the massless closed string states. Indeed, the condition for Weyl invariance, partly worked out in the text, is exactly the equations of motion from the *spacetime* action

$$\mathbf{S} = \frac{1}{2\kappa^2} \int d^{26}x \sqrt{-G} e^{-2\Phi} \left(\mathbf{R} - \frac{1}{12} H_{\mu\nu\lambda} H^{\mu\nu\lambda} + 4\partial_\mu \Phi \partial^\mu \Phi\right)\,, \tag{12.7}$$

where again bold distinguishes spacetime from worldsheet.

The fields B and Φ couple to the worldsheet in interesting ways, which explain part of the form of the spacetime action. Notice that the $\sqrt{\det g}$'s cancel out so B just couples like a form, $\frac{i}{2\pi\alpha'}\int d^2\sigma B_{\mu\nu}(\partial_1 x^\mu \partial_2 dx^\nu)$. This is invariant under the gauge transformation $\delta B_{\mu\nu} = \partial_\mu \zeta_\nu - \partial_\nu \zeta_\mu$, as is the spacetime action. Again, it is the analog for a string of the way that a point particle couples to a vector potential.

The field Φ couples to the world-sheet curvature, and for $\Phi(X)$ equal to a constant Φ_0 this term just becomes $\chi\Phi_0$ where Φ_0 is the Euler number (this is for a world-sheet without boundary, with a boundary we'd also have surface terms). That is, the path integral weight is $e^{-\chi\Phi_0}$. This why e^{Φ_0} governs the closed string coupling: adding a handle to any surface corresponds to emitting and absorbing an extra closed string, and decreases χ by 2. Similarly, adding a strip corresponds to emitting and absorbing an extra open string, and decreases χ by 1, so the open string coupling goes as $e^{\Phi_0/2}$. Also, the spacetime action at tree level can be thought of as arising from the sphere, and indeed is proportional to $e^{-2\Phi_0}$. You might recall that in field theory the coefficient of the action is the inverse of the loop expansion parameter, because it governs the saddle point expansion of the functional integral.

By the way, we are always free to make field redefinitions, these don't change the physics. The gravitational term in the action (12.7) is unconventional, because it depends on the field Φ, but we can fix this by defining

$$G_{\mu\nu} = e^{\Phi/6}\tilde{G}_{\mu\nu} \tag{12.8}$$

so that the action becomes

$$\mathbf{S} = \frac{1}{2\kappa^2}\int d^{26}x\sqrt{-\tilde{G}}\left(\tilde{\mathbf{R}} - \frac{1}{12}e^{-\Phi/3}H_{\mu\nu\lambda}\tilde{H}^{\mu\nu\lambda} - \frac{1}{6}\partial_\mu\Phi\tilde{\partial}^\mu\Phi\right) . \tag{12.9}$$

(By the way, in the text I give a slightly more general form for this redefinition, which would be useful for focussing on a vacuum with some given nonzero value of Φ_0, but this is not usually needed.) I've put tildes in wherever the metric appears, even implicitly through raising indices. The metric $\tilde{G}_{\mu\nu}$ is known as the Einstein metric and $G_{\mu\nu}$ as the string metric. The Einstein metric makes the gravitational dynamics standard and simple (the dilaton and metric don't mix),while the string metric makes the role of the dilaton clearest, and is often simpler for describing certain solutions (black branes). Which metric do we talk about in GR? Well, if there were really a massless dilaton we would have to worry about this, but since quantum corrections will produce a potential for Φ and fix its value we don't have to worry about this.

Compactification

If spacetime has more than four dimensions, it must be that around here the metric is approximately the product,

$$G_{MN} = \begin{bmatrix} \eta_{\mu\nu} & 0 \\ 0 & G_{mn}(x^p) \end{bmatrix} . \tag{12.10}$$

Here I am using M, N for all the dimensions, μ, ν for the four large ones, and m, n, p for the compact ones. Of course, the large dimensions are not exactly flat, but that is on much longer scales. The extra dimensions, 22 or however many, must be highly compact. How small? If they were larger than 10^{-15} cm we would be seeing lots of massive particles corresponding to Kaluza-Klein states, whose wavefunctions that depend on the extra dimensions, so at least this small (with branes involved they can get a bit larger). In the simplest case they would be close to the Planck scale, perhaps just below, because this is where the dimensionless gauge and gravitational couplings meet (fig. 18.1). Again, with branes there are more possibilities (fig. 18.2). In studies of compactification, it is usually assumed that the extra dimensions are large enough that we can study them using just the long distance, e.g. action (12.9), although string corrections are often needed to provide specific effects.

The simplest solution for the compact space is that Φ be constant, $H_{\mu\nu\lambda}$ vanish, and $G_{\mu\nu}$ have vanishing $\mathbf{R}_{\mu\nu}$. Ricci-flat spaces, if they have an additional property required by supersymmetry (Kähler), are known as Calabi-Yau spaces. There is one simple example: periodic flat space. All others are complicated and nonsymmetric. Fortunately the flat periodic case is very interesting (not for realistic models, but to see some interesting phenomena), and we turn to it next.

13 Toroidal compactification

Kaluza-Klein theory

We can learn a lot about string physics without dealing with curved backgrounds at all, just by making some of the directions periodic. I will focus entirely on the case of a single periodic dimension, leaving the generalization to the text. We should first think about the effect this has in ordinary general relativity. We'll use M, N for the $D = d + 1$ coordinates of spacetime, and $\mu, \nu = 0, \ldots, d-1$ for the d noncompact directions. For x^d we have the periodicity

$$x^d \equiv x^d + 2\pi R . \tag{13.1}$$

The metric $G^{(D)}_{MN}$ separates into $G_{\mu\nu}$, $G_{\mu d} = G_{d\mu}$, and G_{dd}, which are respectively a metric, a gauge field, and a scalar from the point of view of the noncompact dimensions (e.g. what a low energy observer would see, who could not resolve the extra dimensions). This separation is best made manifest by parameterizing the metric as

$$ds^2 = G^{(d)}_{\mu\nu} dx^\mu dx^\nu + e^{2\sigma}(dx^d + A_\mu dx^\mu)^2 \,. \tag{13.2}$$

This is the most general metric invariant under translations of x^d, with $G^{(d)}_{\mu\nu}, A_\mu$, and σ being general functions of the noncompact coordinates x^μ. The most general coordinate transformation that leaves this form invariant is reparameterizations $x'^\mu(x^\nu)$ and also $x'^d = x^d + \lambda(x^\nu)$. The first is d-dimensional coordinate invariance, and the second is an effective gauge transformation

$$A'_\mu = A_\mu - \partial_\mu \lambda \,. \tag{13.3}$$

If we insert the Ansatz (13.2) into the action we get

$$\frac{1}{2\kappa^2}\int d^D x\, \sqrt{-G^{(D)}} e^{-2\Phi} \mathbf{R}^{(D)} = \frac{\pi R}{\kappa^2}\int d^d x\, \sqrt{-G^{(d)}} e^{-2\Phi+\sigma} \left(\mathbf{R}^{(d)} - 4\partial_\mu \sigma \partial^\mu \Phi - \frac{1}{4} e^{2\sigma} F_{\mu\nu}F^{\mu\nu} \right) . \tag{13.4}$$

Not surprisingly, we get a Maxwell action, since it must be gauge invariant. (This is just one term, for the full action see eq. 8.1.9 of the text.)

The physical circumference of the periodic dimension is $2\pi R e^\sigma$. The parameter R is arbitrary and can be changed by a redefinition of σ; to study a specific vacuum it is convenient to set $\sigma = 0$ by a shift, and we will do this in the rest of the section. Any constant value of σ solves the field equation, so we have a family of solutions (actually a two-modulus family, since we have both σ and Φ). This troubled Kaluza, Klein, and Einstein. Such continuous families commonly arise with the simply and more symmetric solutions to string theory, but quantum corrections generically eliminate the degeneracy.

In string theory we get another gauge field as well, from $B_{MN} \to B_{\mu\nu}, B_{\mu d}$.

The Ansatz (13.2) is the most general that is independent of x^d, but excited states can depend on x^d. They must be periodic of course, so their momenta are quantized, $p_d = n/R$. If we have a massless state, $p^M p_M = 0$, then $-p^\mu p_\mu = p^d p^d = n^2/R^2$, so the effective d-dimensional mass-squared is n/R.

Strings on a torus

The metric is still flat, so the action is quadratic and the X^μ are free fields. The periodicity of spacetime only enters into the periodicity of the mode expansion. For fields there was one effect, the quantization of momentum, but here there are two: the center-of-mass momentum must be quantized, but also a closed string can wrap around the periodic dimension, we need only come back to an equivalent point as we go around the string [draw],

$$X^d(\tau, \sigma + 2\pi) = X^d(\tau, \sigma) + 2\pi R w \,, \quad w \in \mathbf{Z} \,. \tag{13.5}$$

The mode expansion for the noncompact theory was

$$X^\mu = x^\mu - i\alpha' p^\mu \sigma^2 + \text{oscillator terms} = x^\mu - i\frac{\alpha'}{2} p^\mu \ln|z^2| + \text{oscillator terms}\,. \quad (13.6)$$

The X^d oscillations are unaffected by the periodicity, but for the zero modes we now have

$$X^d = x^d - i\frac{\alpha' n}{R}\sigma^2 + Rw\sigma^1 + \text{oscillator terms}\,. \quad (13.7)$$

This can be written in a nice way if we define

$$p_L^d = \frac{n}{R} + \frac{wR}{\alpha'}\,, \quad p_R^d = \frac{n}{R} - \frac{wR}{\alpha'}\,. \quad (13.8)$$

so that

$$X^d = x^d - i\frac{\alpha'}{2} p_L^d \ln z - i\frac{\alpha'}{2} p_R^d \ln \bar{z} + \text{oscillator terms}\,. \quad (13.9)$$

The Virasoro generators are then

$$L_0 = \frac{\alpha' p_L^2}{4} + \sum_{m,M} m N_{mM}\,, \quad \tilde{L}_0 = \frac{\alpha' p_R^2}{4} + \sum_{m,M} m \tilde{N}_{mM}\,. \quad (13.10)$$

The mass-shell condition $L_0 + \tilde{L}_0 - 2 = 0$ becomes

$$M_{(d)}^2 = -p^\mu p_\mu = \frac{n^2}{R^2} + \frac{w^2 R^2}{\alpha'^2} + \frac{2}{\alpha}\left(-2 + \sum_{m,M} m(N_{mM} + \tilde{N}_{mM})\right)\,. \quad (13.11)$$

The first term is from the compact momenta, just as in field theory; the second is from the tension of a stretched string (circumference $\times$ tension $= 2\pi R/2\pi\alpha'$); the third is from $2A$ and the oscillators, just as in noncompact space (note that all D directions appear); these three terms add in quadratures. Also, the constraint $L_0 = \tilde{L}_0$ becomes

$$nw + \sum_{m,M} m(N_{mM} - \tilde{N}_{mM}) = 0\,. \quad (13.12)$$

It is a nice exercise to verify that the one-loop amplitude is still modular invariant, see p. 237 of the text. Under $\tau \to -1/\tau$, the sums over w and n in the trace over states just interchange (after one step, a Poisson resummation). This is a useful consistency check, especially when we get to more complicated situations. In fact, formally if we have closure and associativity of the OPE, then factorization (a generalization of the OPE) and one-loop modular invariance are the only extra properties needed to guarantee a consistent CFT.

T-duality

The mass formula (13.11) has a notable symmetry. When R is large, winding states are heavy, but states of fixed momentum quantum n become light. When R is small, winding states are light but momentum states are heavy. In fact, the formula is invariant under

$$R \leftrightarrow \frac{\alpha'}{R}\,, \quad n \leftrightarrow w\,. \tag{13.13}$$

Moreover, this applies not just to the mass formula but to the interactions as well. We see that it has the effect of leaving p_L^d unchanged and flipping the sign of p_R^d. So if we write $X^d(z,\bar{z}) = X_L^d(z) + X_R^d(\bar{z})$, then

$$X'^d(z,\bar{z}) = X_L^d(z) - X_R^d(\bar{z}) \tag{13.14}$$

has all the same OPE's as X^d (the minus signs come in pairs), but it describes the transformed theory.

This equivalence is known as T-duality, and it is important for many reasons. First, it is a simple example of *emergent spacetime*. As $R \to 0$, we would expect to see only 25 spacetime dimensions at low energy, and this what does happen in field theory. But somehow here the degrees of freedom reorganize themselves so that a new dimension appears. Second, this is an example of a minimum length appearing in string theory - the effective minimum value of R is $\sqrt{\alpha'}$. Third, it is an example of *stringy geometry,* that strings see spacetime differently than point particles. A richer example of this is known as *mirror symmetry,* which can be understood as a fiberwise T-duality.

As an aside, the action (13.4) has two scalar fields Φ, σ. It is natural to wonder whether there is some similarity between them. We know that σ arose from the radius of a compact dimension — is it possible that Φ does as well? And, (going back to the language where $R = \sqrt{\alpha'}$ is fixed and the different solutions are parameterized by σ), we have a symmetry $\sigma \to -\sigma$ — is it possible that such a symmetry exists for Φ as well? Both of these require us to understand the physics at large Φ, strong coupling, so the go beyond the perturbative understanding in this quarter. Also, they can't really be posed for the bosonic string, with its tachyonic instability. But when we get to the superstring theories, I will tell you the answer, and it is surprisingly subtle: the answer is sometimes yes and sometimes no, for each question.

Duality of form theories

The transformation between the fields X^d and X'^d can be thought about in another way. Consider the path integral for a single massless scalar in two flat dimensions,

$$\int DX\, e^{-\frac{1}{2}\int d^2\sigma\, \partial_a X \partial_a X}\,. \tag{13.15}$$

Let's change variables to $V_a = \partial_a X$. But because V_a is a gradient its curl automatically vanishes. Locally the reverse is true, if the curl vanishes it's a gradient (globally things are more subtle, but here I'm just looking at the local degrees of freedom). So we need a functional delta function to set $\epsilon_{ab}\partial_a V_b = 0$. We'll write this in integral form, $\int d\lambda e^{i\lambda x} \propto \delta(x)$,

$$\int DXD\lambda \, e^{-\int d^2\sigma \, \frac{1}{2}V_a V_a + i\lambda\epsilon_{ab}\partial_a V_b} \, . \tag{13.16}$$

Now do the gaussian integral over V and you're left with

$$\int D\lambda \, e^{-\int d^2\sigma \, \frac{1}{2}\partial_a\lambda\partial_a\lambda} \, . \tag{13.17}$$

This looks like the original path integral, but with λ in place of X. What is the relation between these? The equation of motion for V_a from the action (13.16) is

$$i\lambda\epsilon_{ab}\partial_b\lambda = V_a = \partial_a X \tag{13.18}$$

which is exactly (13.14), identifying $\lambda = X'$.

We've discussed q-form gauge potentials, in d dimensions. The field strength $F_{(q+1)} = dC_{(q)}$ satisfies a Bianchi identity $dF_{(q+1)} = 0$ (vanishing curl). If there are no sources, its field equation can be written $d * F_{(q+1)} = 0$, where $*F_{(q+1)}$ is the $(d-q-1)$-form obtained by contracting $F_{(q+1)}$ with the ϵ tensor. We can interchange the Bianchi identity and field equation: by exactly the series of steps above we can convert a path integral over a q-form gauge field into one over a $q' = (d-q-2)$-form. The example above is $d = 2$, $q = q' = 0$. Another interesting case is $d = 4$, $q = q' = 1$, which converts the usual description of the free Maxwell theory into one with a magnetic instead of electric vector potential. The case $d = 4$, $q = 2$, $q' = 0$ shows that a 2-form has the same degrees of freedom as a scalar in $d = 4$. Another notable case is $d = 3$, $q = 1$, $q' = 0$, so a gauge field is dual to a scalar in $d = 3$.

All this is for the free Abelian theory. We now have strong evidence that in some cases this extends to interacting and non-Abelian theories, but we don't know how to actually transform one path integral into the other as above, in most cases.

Winding charge and enhanced gauge symmetry

Notice that T-duality interchanges the KK gauge field from that with $B_{\mu d}$. We can see this from the vertex operators

$$\partial_z X^\mu \partial_{\bar{z}} X^d + \partial_z X^d \partial_{\bar{z}} X^\mu = -\partial_z X'^\mu \partial_{\bar{z}} X'^d + \partial_z X'^d \partial_{\bar{z}} X'^\mu \, . \tag{13.19}$$

By the way, we noted in KK compactification that momentum states are charged under the gauge field from $G_{\mu d}$, but nothing was charged under that from $B_{\mu d}$. String theory fixed this: the momentum is the KK charge n and the winding number w is the B charge. The $\int B$ coupling looks like an $\int A$ coupling after we integrate over the periodic direction.

We get even more gauge symmetry at the self-dual radius. If we take $n = w = 1$ and excite one $\tilde{\alpha}^{\mu}_{-1}$ oscillator, we satisfy the constraint (13.12) and get the mass formula

$$\frac{(\alpha' - R^2)^2}{R^2 \alpha'^2} \,. \tag{13.20}$$

This goes to zero right at the self-dual radius $R = \sqrt{\alpha'}$. There are four such massless vector states from all cases $|n| = |w| = 1$. Also, they are charged under the $U(1)$'s from $G_{\mu d}$ and $B_{\mu d}$. We must therefore be getting a non-Abelian symmetry, and we are, $SU(2) \times SU(2)$. To see this we can tabulate the charges of the four states under A_{KK} and A_B, and then look at $\frac{1}{2}(A_{KK} \pm A_B)$. In fact, in this case T duality is just an element of the group, though more generally you can have T duality without extra gauge symmetry.

This phenomenon of enhancement of Abelian symmetries to non-Abelian due to charged wrapped strings or branes happens in many situations.

Orbifolds

I just want to briefly describe a more general construction that is extremely useful in generating new solutions to string theory. String theory in the periodic space is obtained from the noncompact one be (a) requiring that states be invariant under translation by $2\pi R$ and (b) including strings which are closed only up to translation by any multiple $2\pi w R$. This can be extended to *any* subgroup of the spacetime symmetry, and gives a consistent string theory still based on free fields on the world-sheet. This is known as *orbifolding.* (It also can be applied to curved solutions with symmetries, for which the world-sheet fields are interacting.)

For example, consider the group generated by two operations, translation by $2\pi R$ and reflection $X^d \to -X^d$. Now the plane $X^d = 0$, whose points are fixed by the reflection, becomes a spacetime boundary, and so also does $X^d = \pi R$ under the combination of reflection followed by $2\pi R$ translation: the X^d direction is a segment rather than a circle, $0 \le x^d \le \pi$. States must be invariant under this reflection, but in addition there are twisted states that are stuck near the fixed plane [draw]. The fixed planes are like mirrors (which can even reflect gravity!) but more generally (when the reflection acts on more than one dimension) they are a sort of conical singularity.

This is a very rich construction, for example many Calabi-Yau manifolds have special cases where they become orbifolds of flat spacetime. Also, one can construct a simple example of topology change where we have a big compactified dimension which is periodic (a circle), make it smaller, and then expand it out again but the dimension has become a segment: topology change! The details take a little while, the situation is depicted in fig. 8.2.

14 D-branes

Now let us consider how T-duality acts on open strings. These have no winding quantum number, so as $R \to 0$ there are no states becoming light, and no emergent dimension. But then if we have a theory with both open and closed strings, as $R \to 0$ the closed strings think that they live in 25+1 dimensions, but the open strings think they live in 24+1 dimensions — how can this be? It must be that the open string is confined to a 24-dimensional hyperplane. More precisely, its endpoints are stuck, since the oscillators still exist in all directions. To show this, take the boundary again to lie along the real z axis, so

$$0 = \partial_n X^d = (\partial_z - \partial_{\bar{z}})X^d = (\partial_z + \partial_{\bar{z}})X'^d = \partial_t X'^d \,, \tag{14.1}$$

where t means tangent to the boundary and n means normal. So X'^d is constant along the boundary, the endpoint can't move in this direction. The Neumann boundary condition becomes Dirichlet.

We could compactify additional dimensions and take their radii to zero, each time getting an additional Dirichlet dimension and reducing the dimension of the hyperplane by one. We will come back to this idea later.

Wilson line breaking and D-brane coordinates

What is this hyperplane? In fact, it must be a dynamical object, it can't be rigid because spacetime itself is dynamical. In fact, I claim that the open string states $\alpha^d_{-1}|0,k\rangle$ correspond to ripples of the hyperplane, just as the closed string graviton states correspond to ripples of spacetime.

To make this more evident, let us first return to the purely Neumann case and introduce a useful idea, Wilson line breaking. Let us first return to field theory, compactified on a circle, and suppose that we have a gauge field A_M already in the higher dimensional theory (so this is not the KK gauge field). A constant value for A_d, independent off all coordinates, gives zero field strength so it solve the field equation. In fact it seems trivial, pure gauge, but it is not. To see this, note that the covariant derivative in the d-direction is $D_d = \partial_d - iA_d$. For a charged state whose wavefunction is proportional to $e^{inx^d/R}$ this becomes $i(\frac{n}{R} + A_d)$. The wave equation for the charged field ϕ is then

$$0 = D_M D^M \phi - M^2\phi = \partial_\mu \partial^\mu \phi - (n/R - A_d)^2 \phi - M^2 \phi \,, \tag{14.2}$$

so the effective d-dimensional mass is

$$M^2_{(d)} = M^2 + (n/R - A_d)^2 \,. \tag{14.3}$$

For $A_d = 0$ this is just the familiar mass due to compact momentum, but now we see that it is shifted. Defining $A_d = -\theta/2\pi R$, the quantum is shifted from n to $n + \theta/2\pi$, so there is definitely a physical effect. The point is that to gauge this away we would need $\phi \to e^{i\theta x^d/2\pi R}$, and this is not periodic. A gauge invariant measure of this background is the Wilson line

$$e^{i\int_0^{2\pi R} A_M dx^M} = e^{-i\theta} \,. \tag{14.4}$$

Now return to the open string, first with purely Neumann boundary conditions, and suppose the Chan-Paton index runs from 1 to n. The gauge field is an $n \times n$ matrix. Consider the solution where A_d is an x^M-independent diagonal matrix

$$A_d = -\frac{1}{2\pi R}\mathrm{diag}(\theta_1, \theta_2, \ldots, \theta_n) \,. \tag{14.5}$$

Now consider an open string whose Chan-Paton state is $|ij\rangle$ moving in this background. This couples to $A_{d,jj}$ with charge $+1$ and $A_{d,ii}$ with charge -1, so the shift is to $n + (\theta_j - \theta_i)/2\pi$. In particular, if $\theta_j - \theta_i$ is nonzero for all i, j (and none are multiples of 2π, then $n + (\theta_j - \theta_i)/2\pi$ is always nonzero except for the diagonal states $i = j$. So this adds mass to all the off-diagonal string states, and in particular to the previously massless gauge fields. So only the n diagonal fields are massless, and the gauge symmetry is broken from $U(n)$ to $U(1)^n$. If r of the θ_i are equal, then we get an $r \times r$ block of massless states, and a $U(r)$ gauge symmetry unbroken. In general then the unbroken subgroup is $U(r_1) \times U(r_2) \times \ldots$ where the total rank is n. (When I teach QFT, I like to assign a problem where $U(n)$ is broken by the expectation value of an adjoint (matrix) scalar, which is essentially what is happening here.)

So now let us take R to be very small and go to the T-dual picture where the radius is $R' = \alpha'/R$. So from this point of view, why are the off-diagonal states massive? The momentum energy gets exhanged with stretching energy, so the idea is that there must be n hyperplanes, with $x'^d = \theta R'$. (To see the normalization, note that the spectrum is periodic when any θ_i changes by 2π, shifting the n's by 1. So we have $x'^d = 2\pi\alpha' A^d$.) [draw] A constant $A_{d,ii}$ is a rigid shift of the i'th plane, while general configurations of the gauge field map to general shapes for the plane.

We call these D-branes, short for Dirichlet membrane. A D-brane has a $U(1)$ gauge field living on it, from A_μ. If r D-branes are coincident, the gauge group is $U(r)$. There are models that get the $SU(3) \times SU(2) \times U(1)$ of the Standard Model in this way.

It might seem that we have a large number of different string theories: one with n D25-branes, one with n D24-branes, and so on. However, it is plausible that the number is far fewer: if we can have one Dp-brane, we should be able to have any number ('cluster decomposition': physics is local, if I can have one here I can have one somewhere else too). Also, suppose we have a theory with only D1's. Then if we have a D1 wrapped on x^1, and another orthogonal to it, under a T-duality along x^1 the wrapped D1 loses a dimension and becomes a D0, and the unwrapped one gains a dimension and becomes a D2. So following this logic, there should be two theories at most, one with any number of even dimensional D-branes, and one with any odd number. This is actually what we will come to for the superstring, except that the fully Neumann branes have an extra consistency condition.

The D-brane action

We can calculate the interactions of the strings stuck to the branes with each other and with the closed string states by the usual technology, just changing the boundary condition for the open string propagator. This does not change things much, let me just outline a few steps. The vertex operators for the massless modes on the world-sheet are

$$:\partial_t X^\mu e^{ik_\mu X^\mu}:\,, \quad :\partial_n X^d e^{ik_\mu X^\mu}:\,. \tag{14.6}$$

The world-sheet XX propagator is

$$-\frac{\alpha'}{2}\ln|z_1 - z_2|^2 \mp \frac{\alpha'}{2}\ln|z_1 - \bar{z}_2|^2 \tag{14.7}$$

where the second term is an image charge, on the other side of the real axis. The upper sign is for the Neumann components, so it just gives a factor of 2 when both points are on the boundary as we used before. The lower sign is for Dirichlet, so it gives zero when one point is on the boundary but gives a nonzero contribution for the normal derivatives. Notice also that k_μ is nonzero only in the Neumann directions, the Dirichlet conditions set it to zero.

So now you can work out the amplitudes for these modes to scatter along the open string. There are also processes like two open strings colliding and turning into a closed string (disk with two boundary and one interior vertex operator). As for the closed string fields, the most important information is the long-wavelength action for the massless fields. Here again we can reason it out from general considerations.

We imagine a Dp-brane moving in a general closed string background G_{MN}, B_{MN}, Φ. To describe the brane we introduce coordinates on it, ξ^a for $a = 0, 1, \ldots, p$, and its embedding in spacetime is given by $X^M(\xi)$. So we are describing a general curved brane here. There is also a gauge field on the brane, $A_a(\xi)$, which is tangent to the brane. I claim that the action is

$$\mathbf{S}_p = -T_p \int d^{p+1}\xi\, e^{-\Phi}[-\det(G_{ab} + B_{ab} + 2\pi\alpha' F_{ab})]^{1/2} + \text{higher derivative terms}\,. \tag{14.8}$$

Here I have defined the ‘pullbacks’

$$G_{ab}(\xi) = G_{MN}\partial_a X^M \partial_b X^N\,, \quad B_{ab}(\xi) = B_{MN}\partial_a X^M \partial_b X^N\,. \tag{14.9}$$

Note that G_{ab} is the same as h_{ab} in the string action.

The dilaton dependence fits with the general scaling $e^{-\chi\Phi}$, where the lowest order open string amplitudes come from the disk, $\chi = 1$. We'll verify this normalization later. Next, if we set B and F to zero, this looks just like the Nambu-Goto action, extended to an object with more dimensions: it's just the invariant area swept out by the brane as it moves through spacetime. It's obviously invariant both under world-sheet and spacetime coordinate transformations. This action was first written down by Dirac (for flat spacetime) for $p = 2$ in 1962, actually eight years before Nambu and Goto wrote if for $p = 1$.

Next let's turn on F_{ab}. The way this appears inside the square root is a bit unexpected, though if we expand in powers of F we do get the usual F^2 term, as well as higher terms (agreeing with the cubic term that we found in the. To understand this form, again we can use T-duality. I want to start with a D$(p-1)$ brane, with the X'^2 direction compact and the X'^2 coordinate of the brane a function of x^1 [draw], flat spacetime. Pythagoras tells us that action is proportional to

$$\int dx^1 \sqrt{1 + (\partial_1 X'^2)^2}\,. \tag{14.10}$$

Now go to the T-dual description, where this becomes

$$\int dx^1 \sqrt{1 + (2\pi\alpha'\partial_1 A_2)^2}\,. \tag{14.11}$$

We complete this by gauge invariance to F_{12}, and then note that

$$\det \begin{bmatrix} 1 & 2\pi\alpha' F_{12} \\ 2\pi\alpha' F_{21} & 1 \end{bmatrix} = 1 + (2\pi\alpha' F_{12})^2\,. \tag{14.12}$$

We can extend to curved space in the obvious way.

This square root action for the gauge field is known as the Born-Infeld action. Born and Infeld were trying to solve the self-energy problem for the electron. For an electric field the action takes the form

$$\sqrt{1 - (2\pi\alpha' E)^2} \tag{14.13}$$

so there is an upper bound to the electric field. However, they did not calculate the energy correctly: if you work out the energy-momentum tensor from Noether's theorem there is a factor of $1 - (2\pi\alpha' E)^{-1/2}$ so there is still a divergence. Today of course we know that it is string theory that ultimately cures the self-energy problem, the finite size of the string, but the Dirac-Born-Infeld (DBI) action survives. There is an interesting interpretation to the maximum field: it is where the force on the string endpoint just balances the tension, any more and a string will stretch without bound; pair production and stretching will screen the field below the critical value.

Finally, B_{ab} (see 8.7.6 through 8.7.10). The string world-sheet action for B_{MN} and the gauge field A_M is

$$\frac{i}{2\pi\alpha'}\int_{\mathcal{M}} B_{MN}\partial_1 X^M \partial_2 X^N d^2\sigma + i\int_{\partial\mathcal{M}} dX^M A_M\,. \tag{14.14}$$

The B_{MN} gauge transformation now picks up a surface term, which is nicely offset by

$$\delta A_M = -\zeta_M/2\pi\alpha'\,. \tag{14.15}$$

Notice that there are two gauge invariances, the ordinary $\delta A_M = \partial_M\lambda$ and this one that has a vector gauge function. Only the combination $B_{ab} + 2\pi\alpha' F_{ab}$ is ζ_M-gauge invariant, so this is what must appear in the action. Notice that as a result of all this we have the combination $G_{ab} + B_{ab}$, which seems like a plausible result.

A lot of interesting string physics involves branes in background closed string fields, at long wavelength so we can use the effective action description for each. The DBI action then describes how the D-brane moves in the background, but also how it sources the various closed string fields.

The D-brane tension

Here's a nice physical application of the annulus amplitude with no vertex operators. Consider two separated parallel Dp-branes. The annulus amplitude, a loop of open strings with one end on each brane, gives the leading potential energy between them.

The fully Neumann vacuum annulus amplitude was

$$\int_0^\infty \frac{dt}{2t}\mathrm{Tr}e^{-2\pi(L_0-26/24)}J_{\mathrm{osc}}(t)\,, \tag{14.16}$$

where the Jacobian J just cancels two directions of oscillators. This is all just like the torus. Calculating the trace as before gives

$$iV_{26}\int_0^\infty \frac{dt}{2t}(8\pi^2\alpha' t)^{-13}\eta(it)^{-24}\,. \tag{14.17}$$

Now, for the p-branes there are two changes. The momentum integrals only run in the $p+1$ directions parallel to the branes, and there is an extra term in L_0 proportional to the separation y. Thus we get

$$iV_{p+1}\int_0^\infty \frac{dt}{2t}(8\pi^2\alpha' t)^{-(p+1)/2}\eta(it)^{-24}e^{-ty^2/2\pi\alpha'}\,. \tag{14.18}$$

The large-y behavior comes from t near zero, which is best analyzed by a modular transformation $t \to 1/t$. One finds (details left to the text) that the potential energy falls as y^{p-23}, which is the analog of an inverse square law (a sphere surrounding a p-dimensional object in 25 dimensions has $p-24$ dimensions, and the gradient of the potential falls as y^{p-24}.

In the text I work out the result from exchange of gravitons and dilations, and deduce what the tension must be. It has α''s for units, and lots of 2's and π's, but the most notable factor is a $1/g$, where $g = e^{\Phi}$ is the string coupling. The quickest way to see this is to note that the annulus is a pure number, independent of the string coupling. The graviton propagator goes as $e^{2\Phi}$, inverse to the action, so the coupling to each D-brane must contribute a $1/g$. A soliton or classical solution (like the Kaluza-Klein monopole that I mentioned) has an action of order $1/g^2$ (again, this is just the $e^{-2\Phi}$ in the action. The string tension has no g-dependence. So a D-brane is 'heavier' than a string and 'lighter' than a soliton. In string field theory there is an interesting description as a soliton of open string fields.

The KK monopole affects the metric at order 1. The D-brane affects it at order g, and the string at order g^2. So if we pile $1/g$ D-branes on top of each other they begin to look like some classical metric, known as a black brane. One of the keys to understanding black hole entropy, and to seeing AdS/CFT duality, is to take some large number N of D-branes, and adiabatically continuing between small gN and large gN. (To be precise, all of this is needs to be in the supersymmetric case to be meaningful). When gN is small the low energy excitations are the $U(N)$ gauge fields on the brane, when it is large it is the near-horizon geometry of the black brane, which is AdS space.

15 Superstrings

The theory we've been discussing so far does not have a stable vacuum, there is a tachyon. Also, it has only bosons, we need fermions too. We'll fix the second problem, and the first will take care of itself.

The generalization that we will make is not obvious, so to motivate it let us look at the form of the L_0 condition in covariant quantization,

$$-p_\mu p^\mu |\psi\rangle = \frac{1}{\alpha'}\left(-1 + \sum_{n=1}^{\infty}\sum_{i=2}^{D-1} nN_m^i\right)|\psi\rangle\,. \tag{15.1}$$

This has the form of a Klein-Gordon equation. Suppose we want to get a Dirac equation, with $p_\mu \Gamma^\mu$ on the left? Now, p_μ comes from the zero mode of $X^\mu(\sigma,\tau)$, so let's also try to get the Γ_μ the same way. Now, the Dirac matrices satisfy anticommutation relations,

$$\{\Gamma^\mu, \Gamma^\nu\} = 2\eta^{\mu\nu}\,, \tag{15.2}$$

so evidently we need a world-sheet fermion field. Notice however that its spacetime index is vector, bosonic. This seems to give a problem with spin-statistics, but we'll fix this with something known as a GSO projection. There is a different approach, introducing a field with a spacetime spinor index, which turns out to be equivalent, but its world-sheet structure is more intricate so we won't go there. The approach that we're taking is the *Ramond-Neveu-Schwarz* formalism, the other is the *Green-Schwarz* formalism.

World-sheet fermions

Let's look first at the world-sheet Dirac equation. In two dimensions the Dirac matrices are two-dimensional, e.g.

$$\gamma^\tau = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \gamma^\sigma = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{15.3}$$

where the lower case γ^a are on the world-sheet. We're back to a Lorentzian world-sheet here and we're only going to do a simplified version of light-cone quantization, because it's the quickest way to cover all the necessary physics in two lectures. The massless Dirac Lagrangian is

$$i\bar{\Psi}\gamma^a\partial_a\Psi = -i\Psi^\dagger \begin{bmatrix} \partial_\tau - \partial_\sigma & 0 \\ 0 & \partial_\tau + \partial_\sigma \end{bmatrix} \Psi\,, \tag{15.4}$$

giving the Dirac equation

$$\begin{bmatrix} \partial_\tau - \partial_\sigma & 0 \\ 0 & \partial_\tau + \partial_\sigma \end{bmatrix} \Psi = 0\,. \tag{15.5}$$

Notice that the two components of

$$\Psi = \begin{bmatrix} \psi \\ \tilde{\psi} \end{bmatrix} \tag{15.6}$$

don't mix, we can regard them as independent fields and in fact can have one with the other. The separate components are *Weyl fermions,* eigenstates of $\gamma^\tau\gamma^\sigma = \mathrm{diag}(1, -1)$. The Dirac equation just says that they are respectively left- and right-moving,

$$\psi(\tau + \sigma)\,, \quad \tilde{\psi}(\tau - \sigma)\,. \tag{15.7}$$

Notice also that there are no i's in the Dirac equation, in this basis, so it is consistent to impose reality,

$$\psi^\dagger = \psi\,, \quad \tilde{\psi}^\dagger = \tilde{\psi}\,. \tag{15.8}$$

This is a Majorana condition. So the smallest building block here is a Majorana-Weyl fermion. Note that the properties of fermions are very dimension-dependent. In four dimensions one can impose a Weyl condition or a Majorana condition but not both at once.

In more: in *two* dimensions we can have

1. A Dirac fermion, with two complex components.

2. A Weyl fermion, with one complex component.

3. A Majorana fermion, with two real components.

4. A Majorana-Weyl fermion, with one real component.

The Weyl and Majorana fermions have the same number of components, but they are different: the two components of the Majorana fermion move in opposite directions, while the Weyl fermion moves in a single direction.

In *four* dimensions we can have

1. A Dirac fermion, with four complex components.

2. A Weyl fermion, with two complex components.

3. A Majorana fermion, with four real components.

In this case there is an isomorphism between the Weyl and Majorana fermions.

General dimensions are discussed in the Appendix to volume 2. I note that fermion properties depend on the dimension mod 8, so that $D = 10$ is like $D = 2$ (except that there are 16 times as many components in each case).

The world-sheet action is then (in Lorentzian form, in coordinates $g_{ab} = \eta_{ab}$)

$$S = \frac{1}{4\pi\alpha'} \int d\tau\, d\sigma (\partial_\tau X^\mu \partial_\tau X_\mu - \partial_\sigma X^\mu \partial_\sigma X_\mu) + \frac{1}{4\pi i} \int d\tau\, d\sigma \left\{ \psi^\mu (\partial_\tau - \partial_\sigma) \psi_\mu + \tilde{\psi}^\mu (\partial_\tau + \partial_\sigma) \tilde{\psi}_\mu \right\} \tag{15.9}$$

As an aside, it follows from the equation of motion that

$$(\partial_\tau - \partial_\sigma)\Big(\psi_\mu (\partial_\tau + \partial_\sigma) X^\mu\Big) = 0\,, \quad (\partial_\tau + \partial_\sigma)\Big(\psi_\mu (\partial_\tau - \partial_\sigma) X^\mu\Big) = 0\,. \tag{15.10}$$

In the complex notation that we've been using so far,

$$\partial_{\bar{z}} T_{Fz} = 0\,, \quad \partial_z T_{F\bar{z}} = 0\,, \tag{15.11}$$

where $T_{Fz} = \psi_\mu \partial_z X^\mu$. Compare $\partial_{\bar{z}} T_{zz} = \partial_z T_{\bar{z}\bar{z}}$. Note that T_F is fermionic. A few comments

1. As with the energy-momentum tensor, we can define two infinite sets of conserved charges from the Fourier modes of T_F. Together with the Virasoro modes these form the superconformal algebra.

2. The $n = 0$ modes, together with L_0 and $\tilde{L}_0$, form a supersymmetry algebra, the first one found on this side of the Iron Curtain (by Ramond).

3. These world-sheet supersymmetry currents are spacetime scalars/worldsheet spinors.

4. These charges play the same role as the Viraosoro charges, they have to vanish on physical states. However, we will do a quick-and-dirty light-cone quantization.

5. We will later meet a different supersymmetry, a *spacetime* spinor, which is a global symmetry of the theory.

The open string

The quickest way to the spectrum is to generalize the light-cone quantization. In a covariant description we would have fields ψ^μ, $\tilde{\psi}^\mu$. This is what I do in chapter 10, but here I'll just keep ψ^i, $\tilde{\psi}^i$. Let me first consider the open string. We need boundary conditions are 0 and π which relate the right- and left-movers at each end of the string. The variation of the action

$$\frac{1}{4\pi i}\int d\tau\, d\sigma \left\{\psi^i(\partial_\tau - \partial_\sigma)\psi^i + \tilde{\psi}^i(\partial_\tau + \partial_\sigma)\tilde{\psi}^i\right\} \tag{15.12}$$

is equal to the equation of motion plus a surface term

$$\frac{1}{2\pi i}\int d\tau(\psi^i\delta\psi^i - \tilde{\psi}^i\delta\tilde{\psi}^i)|_{\sigma=0}^{\sigma=\pi}\,. \tag{15.13}$$

To satisfy this, we have at each end of the string *two* choices,

$$\psi^i(\tau,0) = \pm\tilde{\psi}^i(\tau,0)\,, \quad \psi^i(\tau,\pi) = \pm'\tilde{\psi}^i(\tau,\pi)\,. \tag{15.14}$$

At $\sigma = 0$ we can always take the + sign, by a field redefinition, but the sign at $\sigma = \pi$ them matters, and we will need both cases, the string Hilbert space will include both.

The + sign at both ends is known as the *Ramond* sector and gives the mode expansion

$$\psi^i(\tau,\sigma) = \sum_{n=-\infty}^{\infty} e^{-in(\tau-\sigma)}\psi^i_n\,, \quad \tilde{\psi}^i(\tau,\sigma) = \sum_{n=-\infty}^{\infty} e^{-in(\tau+\sigma)}\psi^i_n\,, \tag{15.15}$$

and satisfies both the equation of motion and the boundary condition. With the − sign at π, the *Neveu-Schwarz* sector, we get instead an expansion

$$\psi^i(\tau,\sigma) = \sum_{\substack{r=-\infty\\ r\in\mathbf{Z}+\frac{1}{2}}}^{\infty} e^{-ir(\tau-\sigma)}\psi^i_r\,, \quad \tilde{\psi}^i(\tau,\sigma) = \sum_{\substack{r=-\infty\\ r\in\mathbf{Z}+\frac{1}{2}}}^{\infty} e^{-ir(\tau+\sigma)}\psi^i_r\,. \tag{15.16}$$

The canonical commutators in the respective sectors are

$$\{\psi^i_m, \psi^j_n\} = \delta^{ij}\delta_{m+n,0}\,, \quad \{\psi^i_r, \psi^j_s\} = \delta^{ij}\delta_{r+s,0}\,. \tag{15.17}$$

The $n = 0$ mode of the Ramond sector gives us our Dirac equation, and in fact the states in the Ramond sector are our spacetime fermions, and those in the Neveu-Schwarz sector are bosons.

The Neveu-Schwarz sector

For $r > 0$, ψ^i_r is a lower operator and ψ^i_{-r} is a raising operator. Let us focus first on a fixed value of r and i. The anticommutation relations imply that $\psi^i_r\psi^i_r = 0$, so lowering twice annihilates any state (as does raising twice) and we can always start with a state $|\downarrow\rangle$ such that

$$\psi^i_r|\downarrow\rangle = 0\,. \tag{15.18}$$

We can also define

$$\psi^i_{-r}|\downarrow\rangle = |\uparrow\rangle\,, \tag{15.19}$$

and the algebra then implies

$$\psi^i_r|\uparrow\rangle = |\downarrow\rangle\,, \quad \psi^i_{-r}|\uparrow\rangle = 0\,. \tag{15.20}$$

Thus, each fermionic mode generates a two-state system. Notice that $\psi^i_{-r}\psi^i_r$ is a number operator,

$$\psi^i_{-r}\psi^i_r|\downarrow\rangle = 0\,, \quad \psi^i_{-r}\psi^i_r|\uparrow\rangle = |\uparrow\rangle\,. \tag{15.21}$$

Now, we see from the mode expansion that this oscillator has energy r. The natural ordering from the path integral is that the two fermionic states have average energy 0, or $-r/2$, $+r/2$ respectively. The zero point energy of the ground state has the opposite sign as for a bosonic oscillator.

The full spectrum is the tensor product of this over all r and i. I will now make the natural assertion that the mass formula (3.7) generalizes straightforwardly to

$$M^2 = \frac{1}{\alpha'}\sum_{m=1}^{\infty}\sum_{i=2}^{D-1} m\left(N^i_m + \frac{1}{2}\right) + \frac{1}{\alpha'}\sum_{r=1/2}^{\infty}\sum_{i=2}^{D-1} r\left(N^{\psi i}_r - \frac{1}{2}\right)\,. \tag{15.22}$$

The same heuristic that had $1 + 2 + \ldots = -\frac{1}{12}$ gives $\frac{1}{2} + \frac{3}{2} + \ldots = +\frac{1}{24}$, as you've shown in the homework, so the mass becomes

$$M^2 = \frac{1}{\alpha'}\left(-\frac{D-2}{8} + \sum_{m=1}^{\infty}\sum_{i=2}^{D-1} mN^i_m + \sum_{r=1/2}^{\infty}\sum_{i=2}^{D-1} rN^{\psi i}_r\right)\,. \tag{15.23}$$

The ground state $|0,k\rangle_{NS}$ is again a tachyon. The first excited states are obtain by acting with one of the $\psi^i_{-1/2}$,

$$\psi^i_{-1/2}|0,k\rangle_{NS}\,, \quad M^2 = \frac{1}{\alpha'}\left(-\frac{D-2}{8} + \frac{1}{2}\right) = \frac{10-D}{8\alpha'}\,. \tag{15.24}$$

The same argument as in the bosonic case shows that if the spectrum is Lorentz invariant than this must be massless, and so we get the critical dimension of the superstring, $D = 10$.

We still have the tachyon, but there is a new ingredient that we did not have in the bosonic string, a symmetry under reflection of $\psi^i \to -\psi^i$. The tachyon and photon have opposite charges under this symmetry, since they differ by one ψ excitation. In fact, modular invariance turns out to require that we keep only half the states, those with the same ψ-parity as the massless states, meaning an odd number of ψ excitations over the ground state. This is known as *the GSO projection*.

So the tachyon is gone, and the symmetry implies that this is consistent with the interactions.

So the massless spectrum of the NS sector is a gauge field, with gauge group $U(n)$ if we introduce Chan-Paton indices (later this will be different, in the unoriented string).

The Ramond sector

Now the ψ are integer-moded, so the zero point energy immediately cancels between the bosonic and fermionic modes, and the ground state is massless. However, it is degenerate. The modes ψ_0^i don't change the energy, so they act on a nontrivial space of ground states. Their algebra

$$\{\psi_0^i, \psi_0^j\} = \delta^{ij} \tag{15.25}$$

is that of the Dirac matrices, up to a factor of two. To figure out the spectrum, define

$$c_1 = (\psi_0^2 + i\psi_0^3)/\sqrt{2}\,, \quad c_2 = (\psi_0^4 + i\psi_0^5)/\sqrt{2}\,, \quad c_3 = (\psi_0^6 + i\psi_0^7)/\sqrt{2}\,, \quad c_4 = (\psi_0^8 + i\psi_0^7)/\sqrt{2}\,, \tag{15.26}$$

so that $c_1^\dagger = (\psi_0^2 - i\psi_0^3)/\sqrt{2}$, etc. Then

$$\{c_\alpha, c_\beta^\dagger\} = \delta_{\alpha\beta}\,, \quad \{c_\alpha, c_\beta\} = \{c_\alpha^\dagger, c_\beta^\dagger\} = 0\,. \tag{15.27}$$

Thus we have four copies of the fermionic oscillator algebra, generating 2^4 states. We can write out the effect of the operators $\psi_0^i\sqrt{2}$ as 16×16 matrices, which just satisfy the Γ-matrix (Clifford) algebra. As this construction shows, in $2k$ dimensions the Γ matrices are rank 2^k, e.g. 4 in $D = 4$. (In the present case the full dimension is 10 and so the Γ matrices are rank 32, but we gauge away half the components in going to the light-cone and things look 8-dimensional).

By the way, the spin matrix in the 2-3 plane works out to $-i\psi_0^2\psi_0^3 = c_1^\dagger c_1 - \frac{1}{2}$, whose eigenvalues are $\pm\frac{1}{2}$, so these states are indeed spinors (similarly in the 4-5, 6-7, 8,9 planes).

As in the two-dimensional case, this can be reduced to two Weyl representations: starting from the ground state which is annihilated by the lowering c_α, there are 8 states with an even number of excitations (0,2, or 4) and 8 states with an odd number. Again the GSO projection requires us to keep one or the other, though here we can choose which (more on this later). The 8 states with an even number of excitations are known as the 8_s representation of $SO(8)$, and the 8 states with an odd number are known as the 8_c. Under the rotation generators in the 2-3, 4-5, 6-7, and 8-9 planes, the spinors always have spins $\pm\frac{1}{2}$. For 8_s there are an even number of $+\frac{1}{2}$'s, and for 8_c there are an odd number.

Again, in the light-cone we classify states only under the $SO(8)$ that acts perpendicular to their momentum.

In all there are 8 massless fermions (Or $8n^2$ with Chan-Paton indices), the same as the number of massless bosons. In fact this theory is supersymmetric, there is a symmetry between the fermions and bosons. There are equal numbers of fermionic and bosonic states at every level. The massless sector is $D = 10$ supersymmetric Yang-Mills, whose spacetime Lagrangian is simply

$$\frac{1}{2g^2}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu} + i\bar{\chi}\Gamma^\mu D_\mu \chi)\,. \tag{15.28}$$

Here χ is a 16-component Majorana-Weyl spinor in $D = 10$. By the way $D = 10$ is the maximum in which supersymmetric Yang-Mills theory exists, which is probably more directly why 10 is so important.

If you dimensionally reduce to $D = 4$ (meaning that you take the fields to be independent of $x^{4,5,6,7,8,9}$ and integrate the action only over $x^{0,1,2,3}$), so the vectors in the reduced directions become scalars, this becomes the $D = 4$ gauge theory with $\mathcal{N} = 4$ supersymmetry which plays a central role in the basic AdS/CFT duality. This is the low energy field theory on n coincident D3-branes in the IIB superstring (to be defined).

Closed strings

As in the bosonic case, the closed string spectrum is the product of two copies of the massless open string spectrum. Here, we look at the massless case. We can independently choose periodic or antiperiodic conditions on ψ^i and $\tilde{\psi}^i$, so there are four sectors:

NS-NS

Taking both to be antiperiodic, we get vector×vector, which is exactly like the bosonic string, so the massless spectrum is again

$$G_{\mu\nu}\,,\ B_{\mu\nu}\,,\ \Phi\,. \tag{15.29}$$

R-R

Here the left- and right-moving states are each fermionic, so the product state is bosonic. Also, there is a simple result that the product of two fermions transforms like a sum of antisymmetric tensors, this is worked out in the appendix to volume 2. There are still four choices according to which GSO projection we make on each side. The resulting massless spectra are

$$\begin{aligned} 8_s \times 8_s &\to C\,,\ C_{\mu\nu}\,,\ C_{\mu\nu\lambda\rho}(SD)\,, \\ 8_c \times 8_c &\to C\,,\ C_{\mu\nu}\,,\ C_{\mu\nu\lambda\rho}(ASD)\,, \\ 8_s \times 8_c,\ 8_c \times 8_s &\to C_\mu\,,\ C_{\mu\nu\lambda}\,. \end{aligned} \tag{15.30}$$

The meaning of SD and ASD will be explained below.

For example, 8_s contains the states $|\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\rangle$ and $|-\frac{1}{2},-\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\rangle$, and if we take one of these on the right and the other on the left we get total spin zero in each plane, the scalar C.

NS-R, R-NS

Here, the massless fields have one spinor index and one vector, so they are gravitini, gauge fields for supersymmetry. (There is also a field with just a spinor index that can be separated out, the dilatino). So the theory must be supersymmetric. The spinor part of the vertex operator (which we'd need a lot of technology to construct) behaves like a world-sheet current associated with this symmetry.

16 The five superstring theories

After imposing modular invariance on the theory, there are only three distinct superstring theories. The first two have only closed strings (that is, there are no D9-branes in the vacuum, we can still have D-branes as excited states, and there will be open string attached to these). They are distinguished by which GSO projection we make in the R sector on each side. This would appear to give four choices. However, parity (e.g. reflection of x^9) interchanges these (note that it takes $c_4 \leftrightarrow c_4^\dagger$, so it interchanges the filled and empty states for this oscillator). So we don't count separately theories that are just interchanged by parity, the orientation of the coordinate axes is a convention.

The IIA theory has $8_s \times 8_c$ in the R-R sector and massless bosonic fields

$$\text{Type IIA}:\ G_{\mu\nu}\,,\ B_{\mu\nu}\,,\ \Phi\,,\ C_\mu\,,\ C_{\mu\nu\lambda}\,, \tag{16.31}$$

as well as two gravitini and two dilatini. Notice that this theory is invariant under the combined operation of spacetime parity and world-sheet parity, the first turning the R-R sector into $8_c \times 8_s$ and the second restoring it to $8_s \times 8_c$. The IIB theory has $8_s \times 8_s$ in the R-R sector and massless fields

$$\text{Type IIB}: \ G_{\mu\nu}\,, \ B_{\mu\nu}\,, \ \Phi\,, \ C\,, \ C_{\mu\nu}\,, \ C_{\mu\nu\lambda\rho}(SD)\,, \tag{16.32}$$

plus two gravitini and two dilatini. This theory is invariant under world-sheet parity.

An interesting fact is that T-duality interchanges these two theories. We if we T-dual in one direction, the relation between X and X' is to flip the sign of the right-moving part, i.e. a one-sided parity transformation. But this flips the R-moving GSO projection while leaving the L-moving projection the same, so it reverses IIA $\leftrightarrow$ IIB.

If we want to have open strings, i.e. D9-branes in the vacuum, there is the problem that the annulus gives us a $1/0$ divergence. The only way to cancel this is to introduce unoriented world-sheets, so that the Möbius strip cancels the divergence [draw], and this works only for $n = 2^5 = 32$. There is also a projection onto states that are even under world-sheet parity, which gives the gauge group $SO(32)$ not $U(32)$, and removes some of the forms, leaving the massless bosonic spectrum

$$\text{Type I}: \ G_{\mu\nu}\,, \ \Phi\,, \ C_{\mu\nu}\,, \ SO(32)A_\mu\,. \tag{16.33}$$

One gravitino and one dilatino survive the orientation projection.

Heterotic strings

We've seen that the closed string spectrum and amplitudes factorize into the product of left- and right-moving factors. This suggests a crazy way to make a new theory: use the left-moving part of the superstring and the right-moving part of the bosonic string! The idea is that there is no tachyon, because the $L_0 = \tilde{L}_0$ physical state condition would require the bosonic tachyon to match with a tachyonic state from the superstring, which isn't there.

There is an important point though: the zero mode spectrum

$$p_L = \frac{n}{R} + \frac{wR}{\alpha'}\,, \quad p_R = \frac{n}{R} - \frac{wR}{\alpha'} \tag{16.34}$$

doesn't factorize, the two sides are correlated through n and w. For very special periodic compactifications the zero modes factorize: this requires a multiple of 8 periodic dimensions.

The lattices (p_L^m, p_R^m) that arise from compactification all have the property that (in units $\alpha' = 2$) $p_L^2 - p_R^2$ is even for every lattice point, and, if for some vector (v_L, v_R), $v_L \cdot p_L - v_R \cdot pR$ is an integer then v is actually one of the p's (i.e. the lattice is self-dual). In fact, every even self-dual lattice corresponds to some toroidal compactification, with some background

B-field. In order to factorize, we need that the momentum lattice factorize into independent L and R lattices. The inner product $p_L^2 - p_R^2$ has signature (k,k) for k compact dimensions. In order for this to factorize we need an even self-dual lattice of signature $(k,0)$. The theorem is that this is possible only for $k = 8, 16, \ldots$.

But this is just right, we're trying to combine a 26 dimensional theory with a 10 dimensional one, so there are 16 unpaired dimensions. With 16 periodic dimensions, with the periods and angles chosen just right, there are two ways to get a zero mode spectrum that factorizes. One is at a point of enhanced gauge symmetry $SO(32) \times SO(32)$ and the other is at a point of enhanced gauge symmetry $E_8 \times E_8 \times E_8 \times E_8$. So the massless spectrum is

$$(\text{vector} + 8_{s\,\text{or}\,c}) \times (\text{vector} + [SO(32) \text{ or } E_8 \times E_8] \text{ adjoint})\,. \tag{16.35}$$

The bosonic spectrum is

$$\begin{gathered} G_{\mu\nu} + B_{\mu\nu} + \Phi + SO(32)\ A_\mu \\ \text{or} \\ G_{\mu\nu} + B_{\mu\nu} + \Phi + E_8 \times E_8\ A_\mu\,. \end{gathered} \tag{16.36}$$

By the way, we have not had a chance to discuss bosonization, but in 1+1 dimensions one can change variables between bosonic and fermionic theories. The dictionary is $\partial X \to \psi^\dagger \psi$ (for a complex Weyl fermion = 2 MW fermions) and $\psi \to e^{iX_L}$. One check on the latter relation is that both operators have weights $(1,0)$. The anticommuting property of the fermion comes about from the branch cut in the definition of X_L. Using this, one can rewrite the heterotic string theories in terms of 10 ordinary dimensions, 8 right-moving MW $\tilde{\psi}$'s as before, and 16 left-moving complex Weyl fermions. There is no left-moving world-sheet superconformal operator; the left-moving fermions are known as 'current algebra' fermions. The two different heterotic theories come from different GSO projects on the current algebra fermions.

These are all the perturbative superstring theories, we've constructed them all. Since the heterotic theories were found (by Gross and collaborators in 1984), our global understanding of string theory has increased greatly, but this set remains complete. What has changed is our interpretation: there are now seen simply as distinct vacuum (Poincaré invariant) states in a single quantum theory. Some of this is evident from T-duality, but some requires going to strong coupling, the subject of the final chapter.

Branes in superstring theories

T-duality gives us configurations of parallel branes. One might think that these would attract gravitationally, but in fact the configuration is stable due to supersymmetry. The vacuum of the Type I theory is invariant under 16 supersymmetries, so in the T-dual description the D-branes must be invariant under half of the 32 supersymmetries of the Type II theories. (Objects invariant under a part of the supersymmetry algebra are called BPS. Bogolmonyi, Prasad and Sommerfield studied stable multi-soliton solutions. They did not consider supersymmetry explicitly, but in retrospect their models were the bosonic parts of supersymmetric theories). There must be some repulsive force that balances the gravitational and dilaton attractions.

We see a variety of form fields in the massless spectra. When we look at their branes we find a simple and satisfying result: for every form field there are both electric and magnetic sources. By an electric source, I mean that the antisymmetric tensor couples

$$\int C \tag{16.37}$$

to the brane, so an antisymmetric field with q indices couples to a brane with $q-1$ spatial dimensions. To see what I mean by magnetic, note that the field strength $F = dC$ (curl) has vanishing curl and divergence, which can be written

$$dF = 0\,, \quad d{*}F = 0\,. \tag{16.38}$$

The first is the Bianchi identity, which follows from $dd = 0$. The second is a field equation. ${*}F$ is F contracted with the antisymmetric 10-form, and the vanishing of the curl of ${*}F$ is the same as the vanishing of the divergence of F. (By the way, the equations $F = dC$, $dF = d{*}F = 0$ are all at the linear level, there are additional Chern-Simons terms in general.)

Now, if C has q indices, F has $q+1$, and ${*}F$ has $9-q$.[12] The symmetry of the Bianchi identity and the equation of motion suggest that we could write ${*}F = d\tilde{C}$ where the 'magnetic' form $\tilde{C}$ has $8-q$ indices and could then couple naturally to a brane with $7-q$ spatial dimensions. So a q-form couples electrically to a $(q-1)$-brane and magnetically to a $(7-q)$-brane.

The other thing we need to know is that the R-R fields couple to D-branes. Again we can motivate this by T-duality. If we T-dual the type I string the D9-branes become lower dimensional branes, which can be separated via Wilson lines. But the vacuum energy is zero due to supersymmetry, independent of the value of the Wilson line. Something is offsetting the attractive potential due to gravity and the dilaton, and it is a repulsive force from the R-R field. So the R-R spectra of the theories correspond to the following D-branes:

$$\begin{aligned} \text{Type I}: &\ \ \text{D1}\,,\ \text{D5}\,, 32\times\text{D9} \\ \text{Type IIA}: &\ \ \text{D0}\,,\ \text{D2}\,,\ \text{D4}\,,\ \text{D6}\,,\ \text{D8(!)}\,, \\ \text{Type IIB}: &\ \ \text{D1}\,,\ \text{D3}\,,\ \text{D5}\,,\ \text{D7}\ . \end{aligned}$$

[12] Note that for $q = 4$, F and ${*}F$ both have five indices. The condition $F = \pm{*}F$ defines a self-dual or anti-self-dual field. The R-R fields have this property, noted as SD and ASD above.

There's one non-obvious thing here, which is the D8: T-duality tells us we must have it, but it's not required by any of R-R fields. It would couple to a 9-form potential, whose field strength has 10 indices. There must be such a field in the theory (it's called a Romans mass for historic reasons from supergravity), but there is no associated particle. It's exactly like an ordinary 2-index field strength in two dimensions: the number of photon polarizations is $2-2=0$, but still gives rise to a Coulomb potential. Similarly, there is a 10-form potential that couples to the D9-branes.

By the way, this is the list of stable, supersymmetric, D-branes. One can also construct non-supersymmetric D-branes of various dimensions, but these usually have tachyons.

Note that D0 and D6 couple to the same RR field, one electrically and one magnetically. Similarly D1/D5 and D2/D4. D3 couples both electrically and magnetically to the potential $C_{(4)}$: this is because it's field strength satisfies $F_{(5)} = *F_{(5)}$. Dirac showed that if there are both electric and magnetic charges, there is a quantization condition, $q_e q_m = 2\pi n$ in order that the wavefunction for a charged particle in the presence of a monopole be well-defined. Dirac's principle extends to each of these pairs of branes. One can calculate the charges via the same stringy calculation 14.18 used of these notes to get the D-brane tension, and the product of the charges works to exactly one Dirac unit. It seems to be a general, and attractive, principle in string theory that every gauge field comes with a complete set of electric and magnetic charges, one that saturates the Dirac quantization.

What if we apply our principle to $B_{\mu\nu}$? Its electric source is the string, but its magnetic source would be a 5-brane. This is known as the NS 5-brane, and it is a classical solution to the field equations. Its tension is of order g^{-2}.

The IIB theory has both fundamental strings and D1-branes, and one can form a bound state of p fundamentals and q D-strings. One can also form (p,q) bound states of the two kinds of 5-brane. And, it has a rich spectrum of 7-branes, which couple to both C and Φ (but it has only one kind of D3-brane).

The Type I theory does not have $B_{\mu\nu}$, so the string carries no charge. And it's not stable, it can break. It is the one unstable object that is particularly interesting, because we can turn off its instability by taking the coupling to zero.

In the heterotic theories, the only brane is the NS5-brane.

For the D-brane action, we have to add to the earlier bosonic result the form coupling (16.39). The full coupling is a bit more complicated: by considering T-dualities with tilted branes as before, you get

$$\int_{\mathrm{D}p} C_{(p+1)} + F_{(2)} \wedge C_{(p+1)} + \frac{1}{2} F_{(2)} \wedge F_{(2)} \wedge C_{(p-1)} + \dots \tag{16.39}$$

integrated over the world-volume of the Dp-branes. These additional 'Chern-Simons' couplings have interesting effects, in particular in terms of the bound states of D-branes: they show that if we turn on the gauge fields on a Dp-brane, it carries the R-R charges of lower dimensional D-branes.

Discussion (see sec. 13.6 of the text): an F-string can dissolve into a D-brane and it becomes electric flux. This respects the $B_{\mu\nu}$ gauge invariance. Similarly, lower dimensional D-branes can dissolve into higher dimensional D-branes and become magnetic flux.

17 Superstring theories at strong coupling

Let us think about what happens when the string coupling $g_\mathrm{s} = e^\Phi$ becomes large (earlier I called this g_c, c for closed, but it's normally called g_s). There is an overall $e^{-2\Phi}$ in front of the action, so this is getting smaller meaning that the path integral is less peaked on the classical solutions, the quantum fluctuations are getting larger. This includes the fluctuations of the metric, so you might expect that at large Φ spacetime goes wild. But something very different seems to happen.

Let us think about this another way. The gravitational action is proportional to $g_\mathrm{s}^{-2}\alpha'^4$, where the α' comes from dimensional analysis. In other words, the Planck length is

$$l_\mathrm{P} = g_\mathrm{s}^{-1/4}\alpha'^{-1/2}\,. \tag{17.40}$$

In the first lecture we argued that as we decrease the distance toward the Planck length, the fluctuations of the metric become large. But in string theory, at small g_s, we first reach the string length scale $l_\mathrm{F} = \alpha'^{-1/2}$, and the fluctuations never get large, we have a good perturbation theory. But if $g_\mathrm{s} > 1$ we get to the Planck scale first. In fact, though, something else happens first. What it is is different in each string theory, we'll start with the simplest case, IIB.

The IIB theory has a D1-brane, whose tension is proportional to $1/g_\mathrm{s}$. In fact, it is convenient to adopt a convention in which it is precisely $1/g_\mathrm{s}$ times the F1 tension $1/2\pi\alpha'$. What is not a convention is that the ratio of the F1 and D1 tensions is precisely proportional to $1/g_\mathrm{s}$, it does not get any corrections: this is a consequence of supersymmetry. So we have a new length scale $l_\mathrm{D} = g_\mathrm{s}^{-1/2}\alpha'^{-1/2}$. At weak coupling $l_\mathrm{F} > l_\mathrm{P} > l_\mathrm{D}$, but at strong coupling $l_\mathrm{D} > l_\mathrm{P} > l_\mathrm{F}$, so we get to the scale of the D-string tension before we get to the Planck scale. So we might conjecture that at strong coupling, there is a dual description in which we quantize the D-strings instead of the F-strings, and with the reciprocal coupling. This conjecture is supported by other evidence: the supergravity action has a $\Phi \to -\Phi$ symmetry, which interchanges the RR and NS-NS two forms. Also, the symmetry relates the D5 and the NS5 with the correct tensions. It carries the D3 into itself, but inverts the coupling of the D3 $U(N)$ gauge group; this would imply a purely field-theoretical duality for the $\mathcal{N} = 4$ gauge theory, which is also believed to be true.

There is no proof of this, or even of the purely field-theoretic duality just mentioned. But there is a lot of circumstantial evidence of the sort just mentioned. Most telling is that it seems to be a very general principle: whenever a theory with enough supersymmetry is taken to strong coupling, it starts to look like some other weakly coupled theory. From another point of view: in the IIB theory at strong coupling, as we increase the energy, *something* happens at the D-string scale. What could it be? The simplest possibility is that it is the same thing that happens at the F-string scale at weak coupling.

By the way, the IIB theory has a larger $SL(2, R)$ symmetry, which acts on $S = C + ie^{\Phi}$ in the same way that the modular group acts on τ.

Now, how about the Type I theory? It has a D-string too, but the D-string doesn't look like a Type I string. In particular, the D-string couples to $C_{\mu\nu}$ but the Type I theory has no $B_{\mu\nu}$. In fact, the Type I string is not a BPS object, and it is not stable, it can break. The Type I D-string is an odd object, when you quantize it you get R 1-1 strings moving one way and R 1-9 strings moving the other (the 1-9 sector was partly developed in the homework). This sounds heterotic, and we have just the right candidate to describe its strongly coupled behavior: the $SO(32)$ heterotic string. All the other circumstantial evidence falls into line as well: these two theories are weak-strong duals.

The IIA theory does not have a D-string, but it has a D0-brane whose mass, $g_{\mathrm{s}}^{-1}\alpha'^{-1/2}$, corresponds to a length-scale that is again longer than that Planck length at strong coupling. So we hit some new physics before gravity gets strong - what is it? One can show that the D0-branes have supersymmetric bound states, whose masses are precisely $ng_{\mathrm{s}}^{-1}\alpha'^{-1/2}$, so we get a whole tower of light states at strong coupling. This looks like a Kaluza-Klein spectrum, and it is: at strong coupling an eleventh dimension appears, with radius $g_{\mathrm{s}}\alpha'^{1/2}$. There is a unique eleven dimensional supergravity theory, whose Kaluza-Klein compactification gives precisely the IIA supergravity action. The $D = 11$ theory has 2-branes and 5-branes, but no strings. The various branes of the IIA theory can all be understood in terms of the $D = 11$ geometry. The term M-theory is sometimes used for the$D = 11$ theory, and sometimes for the full quantum theory that contains all the string theories.

Finally, the heterotic $E_8 \times E_8$ theory is the trickiest, it doesn't have a lot of BPS objects to follow. But it's dual can be deduced from a series of T and S dualities: it is the $D = 11$ theory again, but with the new dimension a bounded segment rather than a circle. In the bulk of the segment we have M theory again, and one E_8 gauge field lives at each end.

A remarkably intricate and unexpected pattern! (Hull and Townsend 1994, Witten 1995). All the string theories are limits of a single quantum theory, whose full form we do not yet know, though we know many limits and special cases.