which succeed in preserving the conformal symmetry at the quantum level.

Looking Ahead: Even when the conformal invariance survives in a 2d quantum theory, the vanishing trace $T^\alpha_{\ \alpha} = 0$ will only turn out to hold in flat space. We will derive this result in section 4.4.2.

The Stress-Tensor in Complex Coordinates

In complex coordinates, $z = \sigma^1 + i\sigma^2$, the vanishing of the trace $T^\alpha_{\ \alpha} = 0$ becomes

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

Meanwhile, the conservation equation $\partial_\alpha T^{\alpha\beta} = 0$ becomes $\partial T^{zz} = \bar{\partial} T^{\bar{z}\bar{z}} = 0$. Or, lowering the indices on T,

$$\bar{\partial}\, T_{zz} = 0 \quad \text{and} \quad \partial\, T_{\bar{z}\bar{z}} = 0$$

In other words, $T_{zz} = T_{zz}(z)$ is a holomorphic function while $T_{\bar{z}\bar{z}} = T_{\bar{z}\bar{z}}(\bar{z})$ is an anti-holomorphic function. We will often use the simplified notation

$$T_{zz}(z) \equiv T(z) \quad \text{and} \quad T_{\bar{z}\bar{z}}(\bar{z}) \equiv \bar{T}(\bar{z})$$

4.1.2 Noether Currents

The stress-energy tensor $T_{\alpha\beta}$ provides the Noether currents for translations. What are the currents associated to the other conformal transformations? Consider the infinitesimal change,

$$z' = z + \epsilon(z) \quad , \quad \bar{z}' = \bar{z} + \bar{\epsilon}(\bar{z})$$

where, making contact with the two examples above, constant ϵ corresponds to a translation while $\epsilon(z) \sim z$ corresponds to a rotation and dilatation. To compute the current, we'll use the same trick that we saw before: we promote the parameter ϵ to depend on the worldsheet coordinates. But it's already a function of half of the worldsheet coordinates, so this now means $\epsilon(z) \to \epsilon(z, \bar{z})$. Then we can compute the change in the action, again using the fact that we can make a compensating change in the metric,

$$\begin{aligned}
\delta S &= -\int d^2\sigma\ \frac{\partial S}{\partial g^{\alpha\beta}}\,\delta g^{\alpha\beta} \\
&= \frac{1}{2\pi}\int d^2\sigma\ T_{\alpha\beta}\,(\partial^\alpha \delta\sigma^\beta) \\
&= \frac{1}{2\pi}\int d^2 z\ \frac{1}{2}\left[T_{zz}\,(\partial^z \delta z) + T_{\bar{z}\bar{z}}\,(\partial^{\bar{z}}\delta\bar{z})\right] \\
&= \frac{1}{2\pi}\int d^2 z\ \left[T_{zz}\,\partial_{\bar{z}}\epsilon + T_{\bar{z}\bar{z}}\,\partial_z\bar{\epsilon}\right]
\end{aligned} \tag{4.6}$$

Firstly note that if ϵ is holomorphic and $\bar{\epsilon}$ is anti-holomorphic, then we immediately have $\delta S = 0$. This, of course, is the statement that we have a symmetry on our hands. (You may wonder where in the above derivation we used the fact that the theory was conformal. It lies in the transition to the third line where we needed $T_{z\bar{z}} = 0$).

At this stage, let's use the trick of treating z and $\bar{z}$ as independent variables. We look at separate currents that come from shifts in z and shifts $\bar{z}$. Let's first look at the symmetry

$$\delta z = \epsilon(z) \quad , \quad \delta\bar{z} = 0$$

We can read off the conserved current from (4.6) by using the standard trick of letting the small parameter depend on position. Since $\epsilon(z)$ already depends on position, this means promoting $\epsilon \to \epsilon(z) f(\bar{z})$ for some function f and then looking at the $\bar{\partial} f$ terms in (4.6). This gives us the current

$$J^z = 0 \quad \text{and} \quad J^{\bar{z}} = T_{zz}(z)\,\epsilon(z) \equiv T(z)\,\epsilon(z) \tag{4.7}$$

Importantly, we find that the current itself is also holomorphic. We can check that this is indeed a conserved current: it should satisfy $\partial_\alpha J^\alpha = \partial_z J^z + \partial_{\bar{z}} J^{\bar{z}} = 0$. But in fact it does so with room to spare: it satisfies the much stronger condition $\partial_{\bar{z}} J^{\bar{z}} = 0$.

Similarly, we can look at transformations $\delta\bar{z} = \bar{\epsilon}(\bar{z})$ with $\delta z = 0$. We get the anti-holomorphic current $\bar{J}$,

$$\bar{J}^z = \bar{T}(\bar{z})\,\bar{\epsilon}(\bar{z}) \quad \text{and} \quad \bar{J}^{\bar{z}} = 0 \tag{4.8}$$

4.1.3 An Example: The Free Scalar Field

Let's illustrate some of these ideas about classical conformal theories with the free scalar field,

$$S = \frac{1}{4\pi\alpha'} \int d^2\sigma\ \partial_\alpha X\, \partial^\alpha X$$

Notice that there's no overall minus sign, in contrast to our earlier action (1.30). That's because we're now working with a Euclidean worldsheet metric. The theory of a free scalar field is, of course, dead easy. We can compute anything we like in this theory. Nonetheless, it will still exhibit enough structure to provide an example of all the abstract concepts that we will come across in CFT. For this reason, the free scalar field will prove a good companion throughout this part of the lectures.

Firstly, let's just check that this free scalar field is actually conformal. In particular, we can look at rescaling $\sigma^\alpha \to \lambda\sigma^\alpha$. If we view this in the sense of an active transformation, the coordinates remain fixed but the value of the field at point σ gets moved to point $\lambda\sigma$. This means,

$$X(\sigma) \;\to\; X(\lambda^{-1}\sigma) \quad \text{and} \quad \frac{\partial X(\sigma)}{\partial\sigma^\alpha} \;\to\; \frac{\partial X(\lambda^{-1}\sigma)}{\partial\sigma^\alpha} = \frac{1}{\lambda}\frac{\partial X(\tilde{\sigma})}{\partial\tilde{\sigma}}$$

where we've defined $\tilde{\sigma} = \lambda^{-1}\sigma$. The factor of λ^{-2} coming from the two derivatives in the Lagrangian then cancels the Jacobian factor from the measure $d^2\sigma = \lambda^2\, d^2\tilde{\sigma}$, leaving the action invariant. Note that any polynomial interaction term for X would break conformal invariance.

The stress-energy tensor for this theory is defined using (4.4),

$$T_{\alpha\beta} = -\frac{1}{\alpha'}\left(\partial_\alpha X\partial_\beta X - \frac{1}{2}\,\delta_{\alpha\beta}(\partial X)^2\right)\,, \tag{4.9}$$

which indeed satisfies $T^\alpha_{\ \alpha} = 0$ as it should. The stress-energy tensor looks much simpler in complex coordinates. It is simple to check that $T_{z\bar{z}} = 0$ while

$$T = -\frac{1}{\alpha'}\,\partial X\partial X \quad \text{and} \quad \bar{T} = -\frac{1}{\alpha'}\,\bar{\partial} X\bar{\partial} X$$

The equation of motion for X is $\partial\bar{\partial}X = 0$. The general classical solution decomposes as,

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

When evaluated on this solution, T and $\bar{T}$ become holomorphic and anti-holomorphic functions respectively.

4.2 Quantum Aspects

So far our discussion has been entirely classical. We now turn to the quantum theory. The first concept that we want to discuss is actually a feature of any quantum field theory. But it really comes into its own in the context of CFT: it is the *operator product expansion*.

4.2.1 Operator Product Expansion

Let's first describe what we mean by a *local* operator in a CFT. We will also refer to these objects as *fields*. There is a slight difference in terminology between CFTs and more general quantum field theories. Usually in quantum field theory, one reserves the term "field" for the objects ϕ which sit in the action and are integrated over in the path integral. In contrast, in CFT the term "field" refers to any local expression that we can write down. This includes ϕ, but also includes derivatives $\partial^n\phi$ or composite operators such as $e^{i\phi}$. All of these are thought of as different fields in a CFT. It should be clear from this that the set of all "fields" in a CFT is always infinite even though, if you were used to working with quantum field theory, you would talk about only a finite number of fundamental objects ϕ. Obviously, this is nothing to be scared about. It's just a change of language: it doesn't mean that our theory got harder.

We now define the *operator product expansion* (OPE). It is a statement about what happens as local operators approach each other. The idea is that two local operators inserted at nearby points can be closely approximated by a string of operators at one of these points. Let's denote all the local operators of the CFT by $\mathcal{O}_i$, where i runs over the set of all operators. Then the OPE is

$$\mathcal{O}_i(z,\bar{z})\,\mathcal{O}_j(w,\bar{w}) = \sum_k C_{ij}^k(z-w,\bar{z}-\bar{w})\,\mathcal{O}_k(w,\bar{w}) \tag{4.10}$$

Here $C_{ij}^k(z-w,\bar{z}-\bar{w})$ are a set of functions which, on grounds of translational invariance, depend only on the separation between the two operators. We will write a lot of operator equations of the form (4.10) and it's important to clarify exactly what they mean: they are always to be understood as statements which hold as operator insertions inside time-ordered correlation functions,

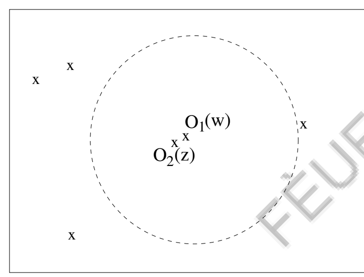


Figure 19:

$$\langle \mathcal{O}_i(z,\bar{z})\,\mathcal{O}_j(w,\bar{w})\ldots\rangle = \sum_k C_{ij}^k(z-w,\bar{z}-\bar{w})\;\langle \mathcal{O}_k(w,\bar{w})\ldots\rangle$$

where the ... can be any other operator insertions that we choose. Obviously it would be tedious to continually write $\langle\ldots\rangle$. So we don't. But it's always implicitly there. There are further caveats about the OPE that are worth stressing

- The correlation functions are always assumed to be time-ordered. (Or something similar that we will discuss in Section 4.5.1). This means that as far as the OPE is concerned, everything commutes since the ordering of operators is determined inside the correlation function anyway. So we must have $\mathcal{O}_i(z,\bar{z})\,\mathcal{O}_j(w,\bar{w}) = \mathcal{O}_j(w,\bar{w})\,\mathcal{O}_i(z,\bar{z})$. (There is a caveat here: if the operators are Grassmann objects, then they pick up an extra minus sign when commuted, even inside time-ordered products).

- The other operator insertions in the correlation function (denoted ... above) are arbitrary. *Except* they should be at a distance large compared to $|z-w|$. It turns out — rather remarkably — that in a CFT the OPEs are exact statements and have a radius of convergence equal to the distance to the nearest other insertion. We will return to this in Section 4.6. The radius of convergence is denoted in the figure by the dotted line.

- The OPEs have singular behaviour as $z \to w$. In fact, this singular behaviour will really be the only thing we care about! It will turn out to contain the same information as commutation relations, as well as telling us how operators transform under symmetries. Indeed, in many equations we will simply write the singular terms in the OPE and denote the non-singular terms as $+\ldots$.

4.2.2 Ward Identities

The spirit of Noether's theorem in quantum field theories is captured by operator equations known as *Ward Identities*. Here we derive the Ward identities associated to conformal invariance. We start by considering a general theory with a symmetry. Later we will restrict to conformal symmetries.

Games with Path Integrals

We'll take this opportunity to get comfortable with some basic techniques using path integrals. Schematically, the path integral takes the form

$$Z = \int \mathcal{D}\phi \; e^{-S[\phi]}$$

where ϕ collectively denote all the fields (in the path integral sense...not the CFT sense!). A symmetry of the quantum theory is such that an infinitesimal transformation

$$\phi' = \phi + \epsilon\delta\phi$$

leaves both the action *and* the measure invariant,

$$S[\phi'] = S[\phi] \quad \text{and} \quad \mathcal{D}\phi' = \mathcal{D}\phi$$

(In fact, we only really need the combination $\mathcal{D}\phi\, e^{-S[\phi]}$ to be invariant but this subtlety won't matter in this course). We use the same trick that we employed earlier in the classical theory and promote $\epsilon \to \epsilon(\sigma)$. Then, typically, neither the action nor the measure are invariant but, to leading order in ϵ, the change has to be proportional to

$\partial\epsilon$. We have

$$Z \longrightarrow \int \mathcal{D}\phi' \;\exp\left(-S[\phi']\right)$$
$$= \int \mathcal{D}\phi \;\exp\left(-S[\phi] - \frac{1}{2\pi}\int J^\alpha\, \partial_\alpha \epsilon\right)$$
$$= \int \mathcal{D}\phi \; e^{-S[\phi]}\left(1 - \frac{1}{2\pi}\int J^\alpha\, \partial_\alpha \epsilon\right)$$

where the factor of $1/2\pi$ is merely a convention and $\int$ is shorthand for $\int d^2\sigma\sqrt{g}$. Notice that the current J^α may now also have contributions from the measure transformation as well as the action.

Now comes the clever step. Although the integrand has changed, the actual value of the partition function can't have changed at all. After all, we just redefined a dummy integration variable ϕ. So the expression above must be equal to the original Z. Or, in other words,

$$\int \mathcal{D}\phi\, e^{-S[\phi]} \;\left(\int J^\alpha\, \partial_\alpha \epsilon\right) = 0$$

Moreover, this must hold for all ϵ. This gives us the quantum version of Noether's theorem: the vacuum expectation value of the divergence of the current vanishes:

$$\langle \partial_\alpha J^\alpha \rangle = 0 \;.$$

We can repeat these tricks of this sort to derive some stronger statements. Let's see what happens when we have other insertions in the path integral. The time-ordered correlation function is given by

$$\langle \mathcal{O}_1(\sigma_1) \ldots \mathcal{O}_n(\sigma_n)\rangle = \frac{1}{Z}\int \mathcal{D}\phi\, e^{-S[\phi]}\, \mathcal{O}_1(\sigma_1)\ldots \mathcal{O}_n(\sigma_n)$$

We can think of these as operators inserted at particular points on the plane as shown in the figure. As we described above, the operators $\mathcal{O}_i$ are any general expressions that we can form from the ϕ fields. Under the symmetry of interest, the operator will change in some way, say

$$\mathcal{O}_i \to \mathcal{O}_i + \epsilon\, \delta\mathcal{O}_i$$

We once again promote $\epsilon \to \epsilon(\sigma)$. As our first pass, let's pick a choice of $\epsilon(\sigma)$ which only has support away from the operator insertions as shown in the Figure 20. Then,

$$\delta\mathcal{O}_i(\sigma_i) = 0$$

and the above derivation goes through in exactly the same way to give

$$\langle \partial_\alpha J^\alpha(\sigma)\,\mathcal{O}_1(\sigma_1)\ldots\mathcal{O}_n(\sigma_n)\rangle = 0 \qquad \text{for } \sigma \neq \sigma_i$$

Because this holds for any operator insertions away from σ, from the discussion in Section 4.2.1 we are entitled to write the operator equation

$$\partial_\alpha J^\alpha = 0$$

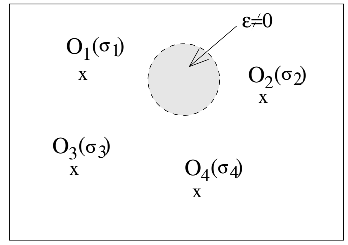


Figure 20:

But what if there are operator insertions that lie at the same point as J^α? In other words, what happens as σ approaches one of the insertion points? The resulting formulae are called Ward identities. To derive these, let's take $\epsilon(\sigma)$ to have support in some region that includes the point σ_1, but not the other points as shown in Figure 21. The simplest choice is just to take $\epsilon(\sigma)$ to be constant inside the shaded region and zero outside. Now using the same procedure as before, we find that the original correlation function is equal to,

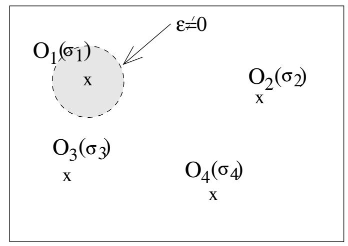


Figure 21:

$$\frac{1}{Z}\int \mathcal{D}\phi\, e^{-S[\phi]}\,\left(1 - \frac{1}{2\pi}\int J^\alpha\,\partial_\alpha\epsilon\right)\,(\mathcal{O}_1 + \epsilon\,\delta\mathcal{O}_1)\,\mathcal{O}_2\ldots\mathcal{O}_n$$

Working to leading order in ϵ, this gives

$$-\frac{1}{2\pi}\int_\epsilon \partial_\alpha\langle J^\alpha(\sigma)\,\mathcal{O}_1(\sigma_1)\ldots\rangle = \langle\delta\mathcal{O}_1(\sigma_1)\ldots\rangle \tag{4.11}$$

where the integral on the left-hand-side is only over the region of non-zero ϵ. This is the *Ward Identity.*

Ward Identities for Conformal Transformations

Ward identities (4.11) hold for any symmetries. Let's now see what they give when applied to conformal transformations. There are two further steps needed in the derivation. The first simply comes from the fact that we're working in two dimensions and we can use Stokes' theorem to convert the integral on the left-hand-side of (4.11) to a line integral around the boundary. Let $\hat{n}^\alpha$ be the unit vector normal to the boundary. For any vector J^α, we have

$$\int_\epsilon \partial_\alpha J^\alpha = \oint_{\partial\epsilon} J_\alpha\hat{n}^\alpha = \oint_{\partial\epsilon}(J_1\,d\sigma^2 - J_2\,d\sigma^1) = -i\oint_{\partial\epsilon}(J_z\,dz - J_{\bar{z}}\,d\bar{z})$$

where we have written the expression both in Cartesian coordinates σ^α and complex coordinates on the plane. As described in Section 4.0.1, the complex components of the vector with indices down are defined as $J_z = \frac{1}{2}(J_1 - iJ_2)$ and $J_{\bar{z}} = \frac{1}{2}(J_1 + iJ_2)$. So, applying this to the Ward identity (4.11), we find for two dimensional theories

$$\frac{i}{2\pi}\oint_{\partial\epsilon} dz\; \langle J_z(z,\bar{z})\,\mathcal{O}_1(\sigma_1)\ldots\rangle - \frac{i}{2\pi}\oint_{\partial\epsilon} d\bar{z}\; \langle J_{\bar{z}}(z,\bar{z})\,\mathcal{O}_1(\sigma_1)\ldots\rangle = \langle\delta\mathcal{O}_1(\sigma_1)\ldots\rangle$$

So far our derivation holds for any conserved current J in two dimensions. At this stage we specialize to the currents that arise from conformal transformations (4.7) and (4.8). Here something nice happens because J_z is holomorphic while $J_{\bar{z}}$ is anti-holomorphic. This means that the contour integral simply picks up the residue,

$$\frac{i}{2\pi}\oint_{\partial\epsilon} dz\, J_z(z)\mathcal{O}_1(\sigma_1) = -\operatorname{Res}\left[J_z\mathcal{O}_1\right]$$

where this means the residue in the OPE between the two operators,

$$J_z(z)\,\mathcal{O}_1(w,\bar{w}) = \ldots + \frac{\operatorname{Res}\left[J_z\mathcal{O}_1(w,\bar{w})\right]}{z-w} + \ldots$$

So we find a rather nice way of writing the Ward identities for conformal transformations. If we again view z and $\bar{z}$ as independent variables, the Ward identities split into two pieces. From the change $\delta z = \epsilon(z)$, we get

$$\delta\mathcal{O}_1(\sigma_1) = -\operatorname{Res}\left[J_z(z)\mathcal{O}_1(\sigma_1)\right] = -\operatorname{Res}\left[\epsilon(z)T(z)\mathcal{O}_1(\sigma_1)\right] \tag{4.12}$$

where, in the second equality, we have used the expression for the conformal current (4.7). Meanwhile, from the change $\delta\bar{z} = \bar{\epsilon}(\bar{z})$, we have

$$\delta\mathcal{O}_1(\sigma_1) = -\operatorname{Res}\left[\bar{J}_{\bar{z}}(\bar{z})\mathcal{O}_1(\sigma_1)\right] = -\operatorname{Res}\left[\bar{\epsilon}(\bar{z})\bar{T}(\bar{z})\mathcal{O}_1(\sigma_1)\right]$$

where the minus sign comes from the fact that the $\oint d\bar{z}$ boundary integral is taken in the opposite direction.

This result means that if we know the OPE between an operator and the stress-tensors $T(z)$ and $\bar{T}(\bar{z})$, then we immediately know how the operator transforms under conformal symmetry. Or, standing this on its head, if we know how an operator transforms then we know at least some part of its OPE with T and $\bar{T}$.

4.2.3 Primary Operators

The Ward identity allows us to start piecing together some OPEs by looking at how operators transform under conformal symmetries. Although we don't yet know the action of general conformal symmetries, we can start to make progress by looking at the two simplest examples.

Translations: If $\delta z = \epsilon$, a constant, then all operators transform as

$$\mathcal{O}(z-\epsilon) = \mathcal{O}(z) - \epsilon\,\partial\mathcal{O}(z) + \ldots$$

The Noether current for translations is the stress-energy tensor T. The Ward identity in the form (4.12) tells us that the OPE of T with any operator $\mathcal{O}$ must be of the form,

$$T(z)\,\mathcal{O}(w,\bar{w}) = \ldots + \frac{\partial\mathcal{O}(w,\bar{w})}{z-w} + \ldots \tag{4.13}$$

Similarly, the OPE with $\bar{T}$ is

$$\bar{T}(\bar{z})\,\mathcal{O}(w,\bar{w}) = \ldots + \frac{\bar{\partial}\mathcal{O}(w,\bar{w})}{\bar{z}-\bar{w}} + \ldots \tag{4.14}$$

Rotations and Scaling: The transformation

$$z \to z + \epsilon z \quad \text{ and } \quad \bar{z} \to \bar{z} + \bar{\epsilon}\bar{z} \tag{4.15}$$

describes rotation for ϵ purely imaginary and scaling (dilatation) for ϵ real. Not all operators have good transformation properties under these actions. This is entirely analogous to the statement in quantum mechanics that not all states transform nicely under the Hamiltonian H and angular momentum operator L. However, in quantum mechanics we know that the eigenstates of H and L can be chosen as a basis of the Hilbert space provided, of course, that $[H, L] = 0$.

The same statement holds for operators in a CFT: we can choose a basis of local operators that have good transformation properties under rotations and dilatations. In fact, we will see in Section 4.6 that the statement about local operators actually follows from the statement about states.

Definition: An operator $\mathcal{O}$ is said to have *weight* $(h, \tilde{h})$ if, under $\delta z = \epsilon z$ and $\delta\bar{z} = \bar{\epsilon}\bar{z}$, $\mathcal{O}$ transforms as

$$\delta\mathcal{O} = -\epsilon(h\mathcal{O} + z\,\partial\mathcal{O}) - \bar{\epsilon}(\tilde{h}\mathcal{O} + \bar{z}\,\bar{\partial}\mathcal{O}) \tag{4.16}$$

The terms $\partial\mathcal{O}$ in this expression would be there for any operator. They simply come from expanding $\mathcal{O}(z-\epsilon z, \bar{z}-\bar{\epsilon}\bar{z})$. The terms $h\mathcal{O}$ and $\tilde{h}\mathcal{O}$ are special to operators which are eigenstates of dilatations and rotations. Some comments:

- Both h and $\tilde{h}$ are real numbers. In a unitary CFT, all operators have $h, \tilde{h} \geq 0$. We will prove this is Section 4.5.4.

- The weights are not as unfamiliar as they appear. They simply tell us how operators transform under rotations and scalings. But we already have names for these concepts from undergraduate days. The eigenvalue under rotation is usually called the *spin*, s, and is given in terms of the weights as

$$s = h - \tilde{h}$$

  Meanwhile, the *scaling dimension* Δ of an operator is

$$\Delta = h + \tilde{h}$$

- To motivate these definitions, it's worth recalling how rotations and scale transformations act on the underlying coordinates. Rotations are implemented by the operator

$$L = -i(\sigma^1\partial_2 - \sigma^2\partial_1) = z\partial - \bar{z}\bar{\partial}$$

  while the dilation operator D which gives rise to scalings is

$$D = \sigma^\alpha\partial_\alpha = z\partial + \bar{z}\bar{\partial}$$

- The scaling dimension is nothing more than the familiar "dimension" that we usually associate to fields and operators by dimensional analysis. For example, worldsheet derivatives always increase the dimension of an operator by one: $\Delta[\partial] = +1$. The tricky part is that the naive dimension that fields have in the classical theory is not necessarily the same as the dimension in the quantum theory.

Let's compare the transformation law (4.16) with the Ward identity (4.12). The Noether current arising from rotations and scaling $\delta z = \epsilon z$ was given in (4.7): it is $J(z) = zT(z)$. This means that the residue of the $J\mathcal{O}$ OPE will determine the $1/z^2$ term in the $T\mathcal{O}$ OPE. Similar arguments hold, of course, for $\delta\bar{z} = \bar{\epsilon}\bar{z}$ and $\bar{T}$. So, the upshot of this is that, for an operator $\mathcal{O}$ with weight $(h, \tilde{h})$, the OPE with T and $\bar{T}$ takes the form

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

Primary Operators

A *primary* operator is one whose OPE with T and $\bar{T}$ truncates at order $(z-w)^{-2}$ or order $(\bar{z}-\bar{w})^{-2}$ respectively. There are no higher singularities:

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

Since we now know all singularities in the $T\mathcal{O}$ OPE, we can reconstruct the transformation under all conformal transformations. The importance of primary operators is that they have particularly simple transformation properties. Focussing on $\delta z = \epsilon(z)$, we have

$$\begin{aligned}\delta\mathcal{O}(w,\bar{w}) &= -\text{Res}\left[\epsilon(z)\,T(z)\,\mathcal{O}(w,\bar{w})\right]\\ &= -\text{Res}\left[\epsilon(z)\left(h\frac{\mathcal{O}(w,\bar{w})}{(z-w)^2} + \frac{\partial\mathcal{O}(w,\bar{w})}{z-w} + \dots\right)\right]\end{aligned}$$

We want to look at smooth conformal transformations and so require that $\epsilon(z)$ itself has no singularities at $z=w$. We can then Taylor expand

$$\epsilon(z) = \epsilon(w) + \epsilon'(w)\,(z-w) + \dots$$

We learn that the infinitesimal change of a primary operator under a general conformal transformation $\delta z = \epsilon(z)$ is

$$\delta\mathcal{O}(w,\bar{w}) = -h\epsilon'(w)\,\mathcal{O}(w,\bar{w}) - \epsilon(w)\,\partial\mathcal{O}(w,\bar{w}) \tag{4.17}$$

There is a similar expression for the anti-holomorphic transformations $\delta\bar{z} = \bar{\epsilon}(\bar{z})$.

Equation (4.17) holds for infinitesimal conformal transformations. It is a simple matter to integrate up to find how primary operators change under a finite conformal transformation,

$$z \to \tilde{z}(z) \quad \text{ and } \quad \bar{z} \to \bar{\tilde{z}}(\bar{z})$$

The general transformation of a primary operator is given by

$$\mathcal{O}(z,\bar{z}) \;\to\; \tilde{\mathcal{O}}(\tilde{z},\bar{\tilde{z}}) = \left(\frac{\partial\tilde{z}}{\partial z}\right)^{-h}\left(\frac{\partial\bar{\tilde{z}}}{\partial\bar{z}}\right)^{-\tilde{h}}\mathcal{O}(z,\bar{z}) \tag{4.18}$$

It will turn out that one of the main objects of interest in a CFT is the spectrum of weights $(h,\tilde{h})$ of primary fields. This will be equivalent to computing the particle mass spectrum in a quantum field theory. In the context of statistical mechanics, the weights of primary operators are the critical exponents.

4.3 An Example: The Free Scalar Field

Let's look at how all of this works for the free scalar field. We'll start by familiarizing ourselves with some techniques using the path integral. The action is,

$$S = \frac{1}{4\pi\alpha'}\int d^2\sigma\ \partial_\alpha X\,\partial^\alpha X \tag{4.19}$$

The classical equation of motion is $\partial^2 X = 0$. Let's start by seeing how to derive the analogous statement in the quantum theory using the path integral. The key fact that we'll need is that the integral of a total derivative vanishes in the path integral just as it does in an ordinary integral. From this we have,

$$0 = \int \mathcal{D}X\ \frac{\delta}{\delta X(\sigma)}\,e^{-S} = \int \mathcal{D}X\ e^{-S}\left[\frac{1}{2\pi\alpha'}\,\partial^2 X(\sigma)\right]$$

But this is nothing more than the Ehrenfest theorem which states that expectation values of operators obey the classical equations of motion,

$$\langle \partial^2 X(\sigma)\rangle = 0$$

4.3.1 The Propagator

The next thing that we want to do is compute the propagator for X. We could do this using canonical quantization, but it will be useful to again see how it works using the path integral. This time we look at,

$$0 = \int \mathcal{D}X\ \frac{\delta}{\delta X(\sigma)}\left[e^{-S}\,X(\sigma')\right] = \int \mathcal{D}X\ e^{-S}\left[\tfrac{1}{2\pi\alpha'}\,\partial^2 X(\sigma)\,X(\sigma') + \delta(\sigma-\sigma')\right]$$

So this time we learn that

$$\langle \partial^2 X(\sigma)\,X(\sigma')\rangle = -2\pi\alpha'\,\delta(\sigma-\sigma') \tag{4.20}$$

Note that if we'd computed this in the canonical approach, we would have found the same answer: the δ-function arises in this calculation because all correlation functions are time-ordered.

We can now treat (4.20) as a differential equation for the propagator $\langle X(\sigma)X(\sigma')\rangle$. To solve this equation, we need the following standard result

$$\partial^2 \ln(\sigma-\sigma')^2 = 4\pi\delta(\sigma-\sigma') \tag{4.21}$$

Since this is important, let's just quickly check that it's true. It's a simple application of Stokes' theorem. Set $\sigma' = 0$ and integrate over $\int d^2\sigma$. We obviously get 4π from the right-hand-side. The left-hand-side gives

$$\int d^2\sigma \; \partial^2 \ln(\sigma_1^2 + \sigma_2^2) = \int d^2\sigma \; \partial^\alpha \left(\frac{2\sigma_\alpha}{\sigma_1^2+\sigma_2^2}\right) = 2\oint \frac{(\sigma_1\, d\sigma^2 - \sigma_2 d\sigma^1)}{\sigma_1^2+\sigma_2^2}$$

Switching to polar coordinates $\sigma_1 + i\sigma_2 = re^{i\theta}$, we can rewrite this expression as

$$2\int \frac{r^2 d\theta}{r^2} = 4\pi$$

confirming (4.21). Applying this result to our equation (4.20), we get the propagator of a free scalar in two-dimensions,

$$\langle X(\sigma)X(\sigma')\rangle = -\frac{\alpha'}{2}\ln(\sigma-\sigma')^2$$

The propagator has a singularity as $\sigma \to \sigma'$. This is an ultra-violet divergence and is common to all field theories. It also has a singularity as $|\sigma - \sigma'| \to \infty$. This is telling us something important that we will mention in Section 4.3.2.

Finally, we could repeat our trick of looking at total derivatives in the path integral, now with other operator insertions $\mathcal{O}_1(\sigma_1), \ldots \mathcal{O}_n(\sigma_n)$ in the path integral. As long as $\sigma, \sigma' \neq \sigma_i$, then the whole analysis goes through as before. But this is exactly our criterion to write the operator product equation,

$$X(\sigma)X(\sigma') = -\frac{\alpha'}{2}\ln(\sigma-\sigma')^2 + \ldots \tag{4.22}$$

We can also write this in complex coordinates. The classical equation of motion $\partial\bar{\partial}X = 0$ allows us to split the operator X into left-moving and right-moving pieces,

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

We'll focus just on the left-moving piece. This has the operator product expansion,

$$X(z)X(w) = -\frac{\alpha'}{2}\ln(z-w) + \ldots$$

The logarithm means that $X(z)$ doesn't have any nice properties under the conformal transformations. For this reason, the "fundamental field" X is not really the object of interest in this theory! However, we can look at the derivative of X. This has a rather nice looking OPE,

$$\partial X(z)\,\partial X(w) = -\frac{\alpha'}{2}\frac{1}{(z-w)^2} + \text{non-singular} \tag{4.23}$$

4.3.2 An Aside: No Goldstone Bosons in Two Dimensions

The infra-red divergence in the propagator has an important physical implication. Let's start by pointing out one of the big differences between quantum mechanics and quantum field theory in $d = 3 + 1$ dimensions. Since the language used to describe these two theories is rather different, you may not even be aware that this difference exists.

Consider the quantum mechanics of a particle on a line. This is a $d = 0 + 1$ dimensional theory of a free scalar field X. Let's prepare the particle in some localized state – say a Gaussian wavefunction $\Psi(X) \sim \exp(-X^2/L^2)$. What then happens? The wavefunction starts to spread out. And the spreading doesn't stop. In fact, the would-be ground state of the system is a uniform wavefunction of infinite width, which isn't a state in the Hilbert space because it is non-normalizable.

Let's now compare this to the situation of a free scalar field X in a $d = 3 + 1$ dimensional field theory. Now we think of this as a scalar without potential. The physics is very different: the theory has an infinite number of ground states, determined by the expectation value $\langle X \rangle$. Small fluctuations around this vacuum are massless: they are Goldstone bosons for broken translational invariance $X \to X + c$.

We see that the physics is very different in field theories in $d = 0 + 1$ and $d = 3 + 1$ dimensions. The wavefunction spreads along flat directions in quantum mechanics, but not in higher dimensional field theories. But what happens in $d = 1 + 1$ and $d = 2 + 1$ dimensions? It turns out that field theories in $d = 1 + 1$ dimensions are more like quantum mechanics: the wavefunction spreads. Theories in $d = 2 + 1$ dimensions and higher exhibit the opposite behaviour: they have Goldstone bosons. The place to see this is the propagator. In d spacetime dimensions, it takes the form

$$\langle X(r)\, X(0) \rangle \sim \begin{cases} 1/r^{d-2} & d \neq 2 \\ \ln r & d = 2 \end{cases}$$

which diverges at large r only for $d = 1$ and $d = 2$. If we perturb the vacuum slightly by inserting the operator $X(0)$, this correlation function tells us how this perturbation falls off with distance. The infra-red divergence in low dimensions is telling us that the wavefunction wants to spread.

The spreading of the wavefunction in low dimensions means that there is no spontaneous symmetry breaking and no Goldstone bosons. It is usually referred to as the Coleman-Mermin-Wagner theorem. Note, however, that it certainly doesn't prohibit massless excitations in two dimensions: it only prohibits Goldstone-like massless excitations.

4.3.3 The Stress-Energy Tensor and Primary Operators

We want to compute the OPE of T with other operators. Firstly, what is T? We computed it in the classical theory in (4.9). It is,

$$T = -\frac{1}{\alpha'}\,\partial X \partial X \tag{4.24}$$

But we need to be careful about what this means in the quantum theory. It involves the product of two operators defined at the same point and this is bound to mean divergences if we just treat it naively. In canonical quantization, we would be tempted to normal order by putting all annihilation operators to the right. This guarantees that the vacuum has zero energy. Here we do something that is basically equivalent, but without reference to creation and annihilation operators. We write

$$T = -\frac{1}{\alpha'}\,:\partial X \partial X:\;\equiv\; -\frac{1}{\alpha'}\,\underset{z\to w}{\text{limit}}\left(\partial X(z)\partial X(w) - \langle \partial X(z)\partial X(w)\rangle\right) \tag{4.25}$$

which, by construction, has $\langle T\rangle = 0$.

With this definition of T, let's start to compute the OPEs to determine the primary fields in the theory.

Claim 1: ∂X is a primary field with weight $h = 1$ and $\tilde{h} = 0$.

Proof: We need to figure out how to take products of normal ordered operators

$$T(z)\,\partial X(w) = -\frac{1}{\alpha'}\,:\partial X(z)\partial X(z):\,\partial X(w)$$

The operators on the left-hand side are time-ordered (because all operator expressions of this type are taken to live inside time-ordered correlation functions). In contrast, the right-hand side is a product of normal-ordered operators. But we know how to change normal ordered products into time ordered products: this is the content of Wick's theorem. Although we have defined normal ordering in (4.25) without reference to creation and annihilation operators, Wick's theorem still holds. We must sum over all possible contractions of pairs of operators, where the term "contraction" means that we replace the pair by the propagator,

$$\overbrace{\partial X(z)\,\partial X(w)} = -\frac{\alpha'}{2}\,\frac{1}{(z-w)^2}$$

Using this, we have

$$T(z)\partial X(w) = -\frac{2}{\alpha'}\,\partial X(z)\,\left(-\frac{\alpha'}{2}\frac{1}{(z-w)^2} + \text{non-singular}\right)$$

Here the "non-singular" piece includes the totally normal ordered term $:T(z)\partial X(w):$. It is only the singular part that interests us. Continuing, we have

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

This is indeed the OPE for a primary operator of weight $h = 1$. $\square$

Note that higher derivatives $\partial^n X$ are not primary for $n > 1$. For example, $\partial^2 X$ has weight $(h, \tilde{h}) = (2, 0)$, but is not a primary operator, as we see from the OPE,

$$T(z)\,\partial^2 X(w) = \partial_w \left[\frac{\partial X(w)}{(z-w)^2} + \ldots\right] = \frac{2\partial X(w)}{(z-w)^3} + \frac{2\partial^2 X(w)}{(z-w)^2} + \ldots$$

The fact that the field $\partial^n X$ has weight $(h, \tilde{h}) = (n, 0)$ fits our natural intuition: each derivative provides spin $s = 1$ and dimension $\Delta = 1$, while the field X does not appear to be contributing, presumably reflecting the fact that it has naive, classical dimension zero. However, in the quantum theory, it is not correct to say that X has vanishing dimension: it has an ill-defined dimension due to the logarithmic behaviour of its OPE (4.22). This is responsible for the following, more surprising, result

Claim 2: The field $:e^{ikX}:$ is primary with weight $h = \tilde{h} = \alpha' k^2/4$.

This result is not what we would guess from the classical theory[5]. Indeed, it's obvious that it has a quantum origin because the weight is proportional to α', which sits outside the action in the same place that $\hbar$ would (if we hadn't set it to one). Note also that this means that the spectrum of the free scalar field is continuous. This is related to the fact that the range of X is non-compact. Generally, CFTs will have a discrete spectrum.

Proof: Let's first compute the OPE with ∂X. We have

$$\begin{aligned}\partial X(z)\ :e^{ikX(w)}: &= \sum_{n=0}^{\infty} \frac{(ik)^n}{n!}\,\partial X(z)\ :X(w)^n: \\ &= \sum_{n=1}^{\infty} \frac{(ik)^n}{(n-1)!}\ :X(w)^{n-1}:\ \left(-\frac{\alpha'}{2}\,\frac{1}{z-w}\right) + \ldots \\ &= -\frac{i\alpha' k}{2}\,\frac{:e^{ikX(w)}:}{z-w} + \ldots \end{aligned} \tag{4.26}$$

[5]We could, however, guess it with a little knowledge of renormalisation. Indeed, we previously derived this result in the lectures on Statistical Field Theory where we computed RG flows in the Sine-Gordon model; see Section 4.4.3 of those lectures.

From this, we can compute the OPE with T.

$$\begin{aligned} T(z)\, :e^{ikX(w)}: &= -\frac{1}{\alpha'}\, :\partial X(z)\partial X(z):\, :e^{ikX(w)}: \\ &= \frac{\alpha' k^2}{4}\frac{:e^{ikX(w)}:}{(z-w)^2} + ik\,\frac{:\partial X(z) e^{ikX(w)}:}{z-w} + \ldots \end{aligned}$$

where the first term comes from two contractions, while the second term comes from a single contraction. Replacing ∂_z by ∂_w in the final term we get

$$T(z)\, :e^{ikX(w)}: = \frac{\alpha' k^2}{4}\frac{:e^{ikX(w)}:}{(z-w)^2} + \frac{\partial_w :e^{ikX(w)}:}{z-w} + \ldots \tag{4.27}$$

showing that $:e^{ikX(w)}:$ is indeed primary. We will encounter this operator frequently later, but will choose to simplify notation and drop the normal ordering colons. Normal ordering will just be assumed from now on. $\square$.

Finally, lets check to see the OPE of T with itself. This is again just an exercise in Wick contractions.

$$\begin{aligned} T(z)\,T(w) &= \frac{1}{\alpha'^2}\, :\partial X(z)\,\partial X(z):\, :\partial X(w)\,\partial X(w): \\ &= \frac{2}{\alpha'^2}\left(-\frac{\alpha'}{2}\frac{1}{(z-w)^2}\right)^2 - \frac{4}{\alpha'^2}\frac{\alpha'}{2}\frac{:\partial X(z)\,\partial X(w):}{(z-w)^2} + \ldots \end{aligned}$$

The factor of 2 in front of the first term comes from the two ways of performing two contractions; the factor of 4 in the second term comes from the number of ways of performing a single contraction. Continuing,

$$\begin{aligned} T(z)\,T(w) &= \frac{1/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} - \frac{2}{\alpha'}\frac{\partial^2 X(w)\,\partial X(w)}{z-w} + \ldots \\ &= \frac{1/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w} + \ldots \end{aligned} \tag{4.28}$$

We learn that T is *not* a primary operator in the theory of a single free scalar field. It is an operator of weight $(h, \tilde{h}) = (2, 0)$, but it fails the primary test on account of the $(z-w)^{-4}$ term. In fact, this property of the stress energy tensor a general feature of all CFTs which we now explore in more detail.

4.4 The Central Charge

In any CFT, the most prominent example of an operator which is not primary is the stress-energy tensor itself.

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

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

and similar for $\bar{T}\bar{T}$. What other terms could we have in this expansion? Since each term has dimension $\Delta = 4$, any operators that appear on the right-hand-side must be of the form

$$\frac{\mathcal{O}_n}{(z-w)^n} \tag{4.29}$$

where $\Delta[\mathcal{O}_n] = 4 - n$. But, in a unitary CFT there are no operators with $h,\ \tilde{h} < 0$. (We will prove this shortly). So the most singular term that we can have is of order $(z-w)^{-4}$. Such a term must be multiplied by a constant. We write,

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

and, similarly,

$$\bar{T}(\bar{z})\,\bar{T}(\bar{w}) = \frac{\tilde{c}/2}{(\bar{z}-\bar{w})^4} + \frac{2\bar{T}(\bar{w})}{(\bar{z}-\bar{w})^2} + \frac{\bar{\partial}\bar{T}(\bar{w})}{\bar{z}-\bar{w}} + \ldots$$

The constants c and $\tilde{c}$ are called the *central charges*. (Sometimes they are referred to as left-moving and right-moving central charges). They are perhaps the most important numbers characterizing the CFT. We can already get some intuition for the information contained in these two numbers. Looking back at the free scalar field (4.28) we see that it has $c = \tilde{c} = 1$. If we instead considered D non-interacting free scalar fields, we would get $c = \tilde{c} = D$. This gives us a hint: c and $\tilde{c}$ are somehow measuring the number of degrees of freedom in the CFT. This is true in a deep sense! However, be warned: c is not necessarily an integer.

Before moving on, it's worth pausing to explain why we didn't include a $(z-w)^{-3}$ term in the TT OPE. The reason is that the OPE must obey $T(z)T(w) = T(w)T(z)$ because, as explained previously, these operator equations are all taken to hold inside time-ordered correlation functions. So the quick answer is that a $(z-w)^{-3}$ term would not be invariant under $z \leftrightarrow w$. However, you may wonder how the $(z-w)^{-1}$ term manages to satisfy this property. Let's see how this works:

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

Now we can Taylor expand $T(z) = T(w) + (z-w)\partial T(w) + \ldots$ and $\partial T(z) = \partial T(w) + \ldots$. Using this in the above expression, we find

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

This trick of Taylor expanding saves the $(z-w)^{-1}$ term. It wouldn't work for the $(z-w)^{-3}$ term.

The Transformation of Energy

So T is not primary unless $c = 0$. And we will see shortly that all theories have $c > 0$. What does this mean for the transformation of T?

$$\begin{aligned}\delta T(w) &= -\text{Res}\,[\epsilon(z)\,T(z)\,T(w)] \\ &= -\text{Res}\left[\epsilon(z)\left(\frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w} + \ldots\right)\right]\end{aligned}$$

If $\epsilon(z)$ contains no singular terms, we can expand

$$\epsilon(z) = \epsilon(w) + \epsilon'(w)(z-w) + \frac{1}{2}\epsilon''(z-w)^2 + \frac{1}{6}\epsilon'''(w)(z-w)^3 + \ldots$$

from which we find

$$\delta T(w) = -\epsilon(w)\,\partial T(w) - 2\epsilon'(w)\,T(w) - \frac{c}{12}\epsilon'''(w) \tag{4.30}$$

This is the infinitesimal version. We would like to know what becomes of T under the finite conformal transformation $z \to \tilde{z}(z)$. The answer turns out to be

$$\tilde{T}(\tilde{z}) = \left(\frac{\partial \tilde{z}}{\partial z}\right)^{-2}\left[T(z) - \frac{c}{12}\,S(\tilde{z}, z)\right] \tag{4.31}$$

where $S(\tilde{z}, z)$ is known as the *Schwarzian* and is defined by

$$S(\tilde{z}, z) = \left(\frac{\partial^3 \tilde{z}}{\partial z^3}\right)\left(\frac{\partial \tilde{z}}{\partial z}\right)^{-1} - \frac{3}{2}\left(\frac{\partial^2 \tilde{z}}{\partial z^2}\right)^2\left(\frac{\partial \tilde{z}}{\partial z}\right)^{-2} \tag{4.32}$$

It is simple to check that the Schwarzian has the right infinitesimal form to give (4.30). Its key property is that it preserves the group structure of successive conformal transformations.

4.4.1 c is for Casimir

Note that the extra term in the transformation (4.31) of T does not depend on T itself. In particular, it will be the same evaluated on all states. It only affects the constant term — or zero mode — in the energy. In other words, it is the Casimir energy of the system.

Let's look at an example that will prove to be useful later for the string. Consider the Euclidean cylinder, parameterized by

$$w = \sigma + i\tau \qquad , \qquad \sigma \in [0, 2\pi)$$

We can make a conformal transformation from the cylinder to the complex plane by

$$z = e^{-iw}$$

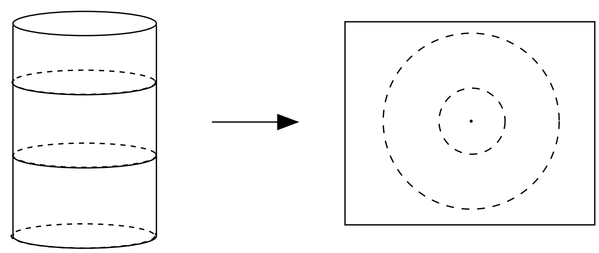


Figure 22:

The fact that the cylinder and the plane are related by a conformal map means that if we understand a given CFT on the cylinder, then we immediately understand it on the plane. And vice-versa. Notice that constant time slices on the cylinder are mapped to circles of constant radius. The origin, $z = 0$, is the distant past, $\tau \to -\infty$.

What becomes of T under this transformation? The Schwarzian can be easily calculated to be $S(z, w) = 1/2$. So we find,

$$T_{\text{cylinder}}(w) = -z^2\, T_{\text{plane}}(z) + \frac{c}{24} \tag{4.33}$$

Suppose that the ground state energy vanishes when the theory is defined on the plane: $\langle T_{\text{plane}} \rangle = 0$. What happens on the cylinder? We want to look at the Hamiltonian, which is defined by

$$H \equiv \int d\sigma\ T_{\tau\tau} = -\int d\sigma\, (T_{ww} + \bar{T}_{\bar{w}\bar{w}})$$

The conformal transformation then tells us that the ground state energy on the cylinder is

$$E = -\frac{2\pi(c + \tilde{c})}{24} \tag{4.34}$$

This is indeed the (negative) Casimir energy on a cylinder. For a free scalar field, we have $c = \tilde{c} = 1$ and the energy density $E/2\pi = -1/12$. This is the same result that we got in Section 2.2.2, but this time with no funny business where we throw out infinities.

An Application: The Lüscher Term

If we're looking at a physical system, the cylinder will have a radius L. In this case, the Casimir energy is given by $E = -2\pi(c+\tilde{c})/24L$. There is an application of this to QCD-like theories. Consider two quarks in a confining theory, separated by a distance L. If the tension of the confining flux tube is T, then the string will be stable as long as $TL \lesssim m$, the mass of the lightest quark. The energy of the stretched string as a function of L is given by

$$E(L) = TL + a - \frac{\pi c}{24L} + \dots$$

Here a is an undetermined constant, while c counts the number of degrees of freedom of the QCD flux tube. (There is no analog of $\tilde{c}$ here because of the reflecting boundary conditions at the end of the string). If the string has no internal degrees of freedom, then $c = 2$ for the two transverse fluctuations. This contribution to the string energy is known as the *Lüscher term*.

4.4.2 The Weyl Anomaly

There is another way in which the central charge affects the stress-energy tensor. Recall that in the classical theory, one of the defining features of a CFT was the vanishing of the trace of the stress tensor,

$$T^\alpha_{\ \alpha} = 0$$

However, things are more subtle in the quantum theory. While $\langle T^\alpha_{\ \alpha}\rangle$ indeed vanishes in flat space, it will not longer be true if we place the theory on a curved background. The purpose of this section is to show that

$$\langle T^\alpha_{\ \alpha}\rangle = -\frac{c}{12}R \tag{4.35}$$

where R is the Ricci scalar of the 2d worldsheet. Before we derive this formula, some quick comments:

- Equation (4.35) holds for any state in the theory — not just the vacuum. This reflects the fact that it comes from regulating short distant divergences in the theory. But, at short distances all finite energy states look basically the same.

- Because $\langle T^\alpha_{\ \alpha}\rangle$ is the same for any state it must be equal to something that depends only on the background metric. This something should be local and must be dimension 2. The only candidate is the Ricci scalar R. For this reason, the formula $\langle T^\alpha_{\ \alpha}\rangle \sim R$ is the most general possibility. The only question is: what is the coefficient. And, in particular, is it non-zero?

- By a suitable choice of coordinates, we can always put any 2d metric in the form $g_{\alpha\beta} = e^{2\omega}\delta_{\alpha\beta}$. In these coordinates, the Ricci scalar is given by

$$R = -2e^{-2\omega}\partial^2\omega \tag{4.36}$$

which depends explicitly on the function ω. Equation (4.35) is then telling us that any conformal theory with $c \neq 0$ has at least one physical observable, $\langle T^\alpha_{\ \alpha}\rangle$, which takes different values on backgrounds related by a Weyl transformation ω. This result is referred to as the *Weyl anomaly*, or sometimes as the trace anomaly.

- There is also a Weyl anomaly for conformal field theories in higher dimensions. For example, $4d$ CFTs are characterized by two numbers, a and c, which appear as coefficients in the Weyl anomaly,

$$\langle T^\mu_{\ \mu}\rangle_{4d} = \frac{c}{16\pi^2}\, C_{\rho\sigma\kappa\lambda}C^{\rho\sigma\kappa\lambda} - \frac{a}{16\pi^2}\tilde{R}_{\rho\sigma\kappa\lambda}\tilde{R}^{\rho\sigma\kappa\lambda}$$

where C is the Weyl tensor and $\tilde{R}$ is the dual of the Riemann tensor.

- Equation (4.35) involves only the left-moving central charge c. You might wonder what's special about the left-moving sector. The answer, of course, is nothing. We also have

$$\langle \bar{T}^\alpha_{\ \alpha}\rangle = -\frac{\tilde{c}}{12}\, R$$

In flat space, conformal field theories with different c and $\tilde{c}$ are perfectly acceptable. However, if we wish these theories to be consistent in fixed, curved backgrounds, then we require $c = \tilde{c}$. This is an example of a *gravitational anomaly*.

- The fact that Weyl invariance requires $c = 0$ will prove crucial in string theory. We shall return to this in Chapter 5.

We will now prove the Weyl anomaly formula (4.35). Firstly, we need to derive an intermediate formula: the $T_{z\bar{z}}\,T_{w\bar{w}}$ OPE. Of course, in the classical theory we found that conformal invariance requires $T_{z\bar{z}} = 0$. We will now show that it's a little more subtle in the quantum theory.

Our starting point is the equation for energy conservation,

$$\partial T_{z\bar{z}} = -\bar{\partial}\, T_{zz}$$

Using this, we can express our desired OPE in terms of the familiar TT OPE,

$$\partial_z T_{z\bar{z}}(z,\bar{z})\ \partial_w T_{w\bar{w}}(w,\bar{w}) = \bar{\partial}_{\bar{z}} T_{zz}(z,\bar{z})\ \bar{\partial}_{\bar{w}} T_{ww}(w,\bar{w}) = \bar{\partial}_{\bar{z}}\bar{\partial}_{\bar{w}}\left[\frac{c/2}{(z-w)^4} + \ldots\right] \tag{4.37}$$

Now you might think that the right-hand-side just vanishes: after all, it is an anti-holomorphic derivative $\bar{\partial}$ of a holomorphic quantity. But we shouldn't be so cavalier because there is a singularity at $z = w$. For example, consider the following equation,

$$\bar{\partial}_{\bar{z}}\partial_z \ln|z-w|^2 = \bar{\partial}_{\bar{z}} \frac{1}{z-w} = 2\pi\delta(z-w, \bar{z}-\bar{w}) \tag{4.38}$$

We proved this statement after equation (4.21). (The factor of 2 difference from (4.21) can be traced to the conventions we defined for complex coordinates in Section 4.0.1). Looking at the intermediate step in (4.38), we again have an anti-holomorphic derivative of a holomorphic function and you might be tempted to say that this also vanishes. But you'd be wrong: subtle things happen because of the singularity and equation (4.38) tells us that the function $1/z$ secretly depends on $\bar{z}$. (This should really be understood as a statement about distributions, with the delta function integrated against arbitrary test functions). Using this result, we can write

$$\bar{\partial}_{\bar{z}}\bar{\partial}_{\bar{w}} \frac{1}{(z-w)^4} = \frac{1}{6}\bar{\partial}_{\bar{z}}\bar{\partial}_{\bar{w}}\left(\partial_z^2\,\partial_w \frac{1}{z-w}\right) = \frac{\pi}{3}\,\partial_z^2\,\partial_w\bar{\partial}_{\bar{w}}\,\delta(z-w,\bar{z}-\bar{w})$$

Inserting this into the correlation function (4.37) and stripping off the $\partial_z\partial_w$ derivatives on both sides, we end up with what we want,

$$T_{z\bar{z}}(z,\bar{z})\; T_{w\bar{w}}(w,\bar{w}) = \frac{c\pi}{6}\partial_z\bar{\partial}_{\bar{w}}\,\delta(z-w,\bar{z}-\bar{w}) \tag{4.39}$$

So the OPE of $T_{z\bar{z}}$ and $T_{w\bar{w}}$ almost vanishes, but there's some strange singular behaviour going on as $z \to w$. This is usually referred to as a contact term between operators and, as we have shown, it is needed to ensure the conservation of energy-momentum. We will now see that this contact term is responsible for the Weyl anomaly.

We assume that $\langle T^\alpha_{\ \alpha}\rangle = 0$ in flat space. Our goal is to derive an expression for $\langle T^\alpha_{\ \alpha}\rangle$ close to flat space. Firstly, consider the change of $\langle T^\alpha_{\ \alpha}\rangle$ under a general shift of the metric $\delta g_{\alpha\beta}$. Using the definition of the energy-momentum tensor (4.4), we have

$$\begin{aligned}\delta\,\langle T^\alpha_{\ \alpha}(\sigma)\rangle &= \delta\int \mathcal{D}\phi\; e^{-S}\,T^\alpha_{\ \alpha}(\sigma)\\ &= \frac{1}{4\pi}\int \mathcal{D}\phi\; e^{-S}\left(T^\alpha_{\ \alpha}(\sigma)\int d^2\sigma'\sqrt{g}\;\delta g^{\beta\gamma}\,T_{\beta\gamma}(\sigma')\right)\end{aligned}$$

If we now restrict to a Weyl transformation, the change to a flat metric is $\delta g_{\alpha\beta} = 2\omega\delta_{\alpha\beta}$, so the change in the inverse metric is $\delta g^{\alpha\beta} = -2\omega\delta^{\alpha\beta}$. This gives

$$\delta\,\langle T^\alpha_{\ \alpha}(\sigma)\rangle = -\frac{1}{2\pi}\int \mathcal{D}\phi\; e^{-S}\left(T^\alpha_{\ \alpha}(\sigma)\int d^2\sigma'\;\omega(\sigma')\,T^\beta_{\ \beta}(\sigma')\right) \tag{4.40}$$

Now we see why the OPE (4.39) determines the Weyl anomaly. We need to change between complex coordinates and Cartesian coordinates, keeping track of factors of 2. We have

$$T^{\alpha}_{\ \alpha}(\sigma)\, T^{\beta}_{\ \beta}(\sigma') = 16\, T_{z\bar{z}}(z,\bar{z})\ T_{w\bar{w}}(w,\bar{w})$$

Meanwhile, using the conventions laid down in 4.0.1, we have $8\partial_z\bar{\partial}_{\bar{w}}\delta(z-w,\bar{z}-\bar{w}) = -\partial^2\,\delta(\sigma-\sigma')$. This gives us the OPE in Cartesian coordinates

$$T^{\alpha}_{\ \alpha}(\sigma)\, T^{\beta}_{\ \beta}(\sigma') = -\frac{c\pi}{3}\,\partial^2\,\delta(\sigma-\sigma')$$

We now plug this into (4.40) and integrate by parts to move the two derivatives onto the conformal factor ω. We're left with,

$$\delta\,\langle T^{\alpha}_{\ \alpha}\rangle = \frac{c}{6}\,\partial^2\omega \quad\Rightarrow\quad \langle T^{\alpha}_{\ \alpha}\rangle = -\frac{c}{12}R$$

where, to get to the final step, we've used (4.36) and, since we're working infinitesimally, we can replace $e^{-2\omega}\approx 1$. This completes the proof of the Weyl anomaly, at least for spaces infinitesimally close to flat space. The fact that R remains on the right-hand-side for general 2d surfaces follows simply from the comments after equation (4.35), most pertinently the need for the expression to be reparameterization invariant.

4.4.3 c is for Cardy

The Casimir effect and the Weyl anomaly have a similar smell. In both, the central charge provides an extra contribution to the energy. We now demonstrate a different avatar of the central charge: it tells us the density of high energy states.

We will study conformal field theory on a Euclidean torus. We'll keep our normalization $\sigma\in[0,2\pi)$, but now we also take τ to be periodic, lying in the range

$$\tau\in[0,\beta)$$

The partition function of a theory with periodic Euclidean time has a very natural interpretation: it is related to the free energy of the theory at temperature $T=1/\beta$.

$$Z[\beta] = \mathrm{Tr}\, e^{-\beta H} = e^{-\beta F} \tag{4.41}$$

At very low temperatures, $\beta\to\infty$, the free energy is dominated by the lowest energy state. All other states are exponentially suppressed. But we saw in 4.4.1 that the vacuum state on the cylinder has Casimir energy $H=-c/12$. In the limit of low temperature, the partition function is therefore approximated by

$$Z \ \rightarrow\ e^{c\beta/12} \quad \text{as } \beta\to\infty \tag{4.42}$$

Now comes the trick. In Euclidean space, both directions of the torus are on equal footing. We're perfectly at liberty to decide that σ is "time" and τ is "space". This can't change the value of the partition function. So let's make the swap. To compare to our original partition function, we want the spatial direction to have range $[0, 2\pi)$. Happily, due to the conformal nature of our theory, we arrange this through the scaling

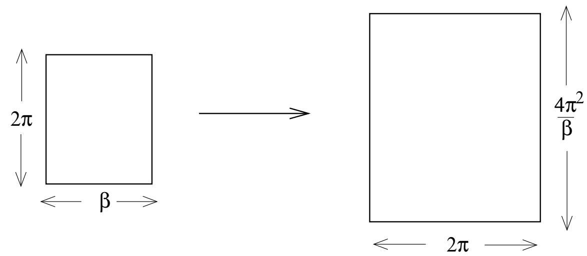


Figure 23:

$$\tau \to \frac{2\pi}{\beta}\,\tau \quad , \quad \sigma \to \frac{2\pi}{\beta}\,\sigma$$

Now we're back where we started, but with the temporal direction taking values in $\sigma \in [0, 4\pi^2/\beta)$. This tells us that the high-temperature and low-temperature partition functions are related,

$$Z[4\pi^2/\beta] = Z[\beta]$$

This is called modular invariance. We'll come across it again in Section 6.4. Writing $\beta' = 4\pi^2/\beta$, this tells us the very high temperature behaviour of the partition function

$$Z[\beta'] \to e^{c\pi^2/3\beta'} \quad \text{as} \quad \beta' \to 0$$

But the very high temperature limit of the partition function is sampling all states in the theory. On entropic grounds, this sampling is dominated by the high energy states. So this computation is telling us how many high energy states there are.

To see this more explicitly, let's do some elementary manipulations in statistical mechanics. Any system has a density of states $\rho(E) = e^{S(E)}$, where $S(E)$ is the entropy. The free energy is given by

$$e^{-\beta F} = \int dE\ \rho(E)\, e^{-\beta E} = \int dE\ e^{S(E)-\beta E}$$

In two dimensions, all systems have an entropy which scales at large energy as

$$S(E) \to N\sqrt{E} \tag{4.43}$$

The coefficient N counts the number of degrees of freedom. The fact that $S \sim \sqrt{E}$ is equivalent to the fact that $F \sim T^2$, as befits an energy density in a theory with one spatial dimension. To see this, we need only approximate the integral by the saddle point $S'(E_\star) = \beta$. From (4.43), this gives us the free energy

$$F \sim N^2 T^2$$

We can now make the statement about the central charge more explicit. In a conformal field theory, the entropy of high energy states is given by

$$S(E) \sim \sqrt{cE}$$

This is *Cardy's formula*. A more careful analysis of the coefficients shows that the high energy density of states scales as

$$S(E) \to 2\pi \sqrt{\frac{c}{6}\left(ER - \frac{c}{24}\right)} \tag{4.44}$$

where the offset is the Casimir energy (4.34) that we derived previously. This is the contribution from left-movers. There is a similar contribution from right-movers, depending on $\tilde{c}$.

4.4.4 c has a Theorem

The connection between the central charge and the degrees of freedom in a theory is given further weight by a result of Zamalodchikov, known as the *c-theorem*. The idea of the c-theorem is to stand back and look at the space of all theories and the renormalization group (RG) flows between them.

Conformal field theories are special. They are the fixed points of the renormalization group, looking the same at all length scales. One can consider perturbing a conformal field theory by adding an extra term to the action,

$$S \to S + \alpha \int d^2\sigma \; \mathcal{O}(\sigma)$$

Here $\mathcal{O}$ is a local operator of the theory, while α is some coefficient. These perturbations fall into three classes, depending on the dimension Δ of $\mathcal{O}$.

- $\Delta < 2$: In this case, α has positive dimension: $[\alpha] = 2 - \delta$. Such deformations are called *relevant* because they are important in the infra-red. RG flow takes us away from our original CFT. We only stop flowing when we hit a new CFT (which could be trivial with $c = 0$).
- $\Delta = 2$: The constant α is dimensionless. Such deformations are called *marginal*. The deformed theory defines a new CFT.

- $\Delta > 2$: The constant α has negative dimension. These deformations are irrelevant. The infra-red physics is still described by the original CFT. But the ultra-violet physics is altered.

We expect information is lost as we flow from an ultra-violet theory to the infra-red. The c-theorem makes this intuition precise. The theorem exhibits a function c on the space of all theories which monotonically decreases along RG flows. At the fixed points, c coincides with the central charge of the CFT.

A Thermodynamic Proof of the c-Theorem

There are a number of different proofs of the c-theorem. Here we give one that is particularly physical. The basic idea is to heat up the system to a finite temperature T and compute the speed of sound. The c-theorem follows from the requirement that the speed of sound does not exceed the speed of light (which, in our conventions, is simply 1). I should warn you that the style of argument in this section is somewhat different from the rest of these lectures. But, if nothing else, it reminds you that just because you're learning string theory, you shouldn't neglect basic physics!

Let's first start with a CFT. For simplicity, we assume that $c = \tilde{c}$. Then, from (4.44), we have the asymptotic behaviour

$$S(E) \rightarrow 4\pi\sqrt{\frac{cER}{6}}$$

where we have dropped the $c/24$ offset, and the overall coefficient is 4π rather than 2π because we are including both left- and right-moving sectors. To compare with familiar, thermodynamic formulae we write this in terms of the spatial volume $V = 2\pi R$, so

$$S(E) \rightarrow 4\pi\sqrt{\frac{\pi cEV}{3}}$$

Now, the temperature is defined to be

$$\frac{1}{T} = \frac{\partial S}{\partial E} = 2\pi\sqrt{\frac{\pi cV}{3E}} \quad \Rightarrow \quad \sqrt{E} = 2\pi T\sqrt{\frac{\pi cV}{3}}$$

From this, we can compute the entropy of a CFT as a function of temperature, rather than as a function of energy

$$S(T) = \frac{8\pi^3 cVT}{3} \quad \Rightarrow \quad s(T) = \frac{8\pi^3 c}{3}T \tag{4.45}$$

where $s = S/V$ is the entropy density.

Now we'll consider a more general situation. We'll flow from some CFT in the UV with central charge c_{UV} to another CFT in the IR with central charge c_{IR}. It may be that the final theory is gapped – meaning that everything is massless – in which case $c_{IR} = 0$. Our goal is to prove that, regardless of the flow, we always have $c_{UV} \geq c_{IR}$ (with equality if there is no flow at all). To achieve this, we need to play around with some thermodynamic identities. In particular, we need to following result

Claim:

$$s = \left.\frac{\partial P}{\partial T}\right|_V \tag{4.46}$$

with P the pressure.

Proof: Given the energy $E = E(S, V)$, the first law of thermodynamics tells us

$$dE = TdS - PdV$$

The free energy is then defined as $F(T, V) = E - TS$ and obeys

$$dF = -SdT - PdV \tag{4.47}$$

But the free energy is extensive and this means that it must, in fact, be proportional to V since this is the only extensive quantity that it can depend on. So

$$F(T, V) = -P(T)\, V$$

From this we learn that

$$dF = -\frac{\partial P}{\partial T} V dT - PdV$$

Comparing to (4.47) gives us the claimed result (4.46). □

Finally, we recall that the speed of sound in a system is given by (see, for example, the lectures on Fluid Mechanics

$$c_s^2 = \frac{dP}{d\epsilon}$$

where $\epsilon = E/V$ is the energy density. At fixed volume, we have

$$dE = TdS \quad \Rightarrow \quad d\epsilon = Tds$$

All of which means that we can express the speed of sound as

$$c_s^2 = \frac{1}{T}\frac{dP}{ds} = \frac{1}{T}\frac{dP}{dT}\frac{dT}{ds} = \frac{s}{T}\frac{dT}{ds} = \frac{d\,\log T}{d\,\log s}$$

This is the key result that we need. Now we define a thermal *c-function*

$$\chi = \frac{s}{T}$$

As we've seen in (4.45), when we have a CFT the function χ is proportional to the central charge: $\chi = 8\pi^3 c/3$. If we flow from a CFT in the UV, with central charge c_{UV}, to a different CFT in the IR with central charge c_{IR}, then χ will interpolate between these two values (multiplied by $8\pi^3/3$) as we vary the temperature. To prove the c-theorem, we need to show that as we decrease the temperature, and so excite lower energy degrees of freedom, the function χ necessarily decreases. We do this by relating χ to the speed of sound,

$$\frac{1}{c_s^2} = \frac{d\,\log s}{d\,\log T} = \frac{d\,\log(\chi T)}{d\,\log T} = 1 + \frac{d\,\log \chi}{d\,\log T}$$

By causality, we must have $c_s^2 \leq 1$ (with equality when we have a CFT) and so

$$\frac{d\,\log \chi}{d\,\log T} \geq 0 \quad \Rightarrow \quad \frac{d\chi}{dT} \geq 0$$

But this is what we wanted. We learn that we necessarily have $c_{UV} \geq c_{IR}$. This is the c-theorem.

4.5 The Virasoro Algebra

So far our discussion has been limited to the operators of the CFT. We haven't said anything about states. We now remedy this. We start by taking a closer look at the map between the cylinder and the plane.

4.5.1 Radial Quantization

To discuss states in a quantum field theory we need to think about where they live and how they evolve. For example, consider a two dimensional quantum field theory defined on the plane. Traditionally, when quantizing this theory, we parameterize the plane by Cartesian coordinates (t, x) which we'll call "time" and "space". The states live on spatial slices. The Hamiltonian generates time translations and hence governs the evolution of states.

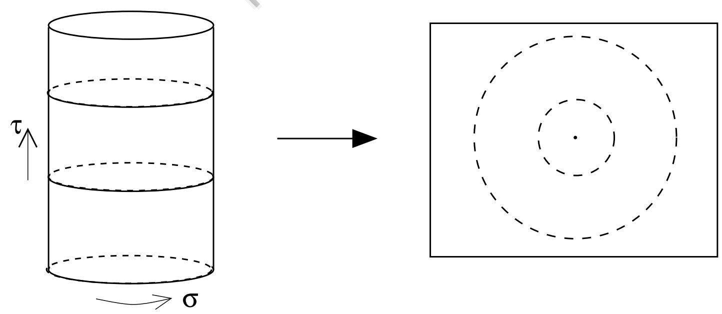


Figure 25: The map from the cylinder to the plane.

However, the map between the cylinder and the plane suggests a different way to quantize a CFT on the plane. The complex coordinate on the cylinder is taken to be ω, while the coordinate on the plane is z. They are related by,

$$\omega = \sigma + i\tau \quad , \quad z = e^{-i\omega}$$

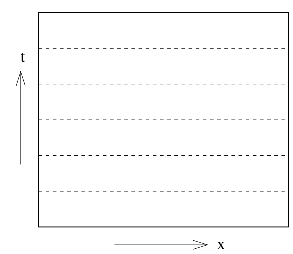


Figure 24:

On the cylinder, states live on spatial slices of constant σ and evolve by the Hamiltonian,

$$H = \partial_\tau$$

After the map to the plane, the Hamiltonian becomes the dilatation operator

$$D = z\partial + \bar{z}\bar{\partial}$$

If we want the states on the plane to remember their cylindrical roots, they should live on circles of constant radius. Their evolution is governed by the dilatation operator D. This approach to a theory is known as *radial quantization*.

Usually in a quantum field theory, we're interested in time-ordered correlation functions. Time ordering on the cylinder becomes radial ordering on the plane. Operators in correlation functions are ordered so that those inserted at larger radial distance are moved to the left.

Virasoro Generators

Let's look at what becomes of the stress tensor $T(z)$ evaluated on the plane. On the cylinder, we would decompose T in a Fourier expansion.

$$T_{\text{cylinder}}(w) = -\sum_{m=-\infty}^{\infty} L_m e^{imw} + \frac{c}{24}$$

After the transformation (4.33) to the plane, this becomes the Laurent expansion

$$T(z) = \sum_{m=-\infty}^{\infty} \frac{L_m}{z^{m+2}}$$

As always, a similar statement holds for the right-moving sector

$$\bar{T}(\bar{z}) = \sum_{m=-\infty}^{\infty} \frac{\tilde{L}_m}{\bar{z}^{m+2}}$$

We can invert these expressions to get L_m in terms of $T(z)$. We need to take a suitable contour integral

$$L_n = \frac{1}{2\pi i}\oint dz\; z^{n+1}\, T(z) \quad , \quad \tilde{L}_n = \frac{1}{2\pi i}\oint d\bar{z}\; \bar{z}^{n+1}\, \bar{T}(\bar{z}) \tag{4.48}$$

where, if we just want L_n or $\tilde{L}_n$, we must make sure that there are no other insertions inside the contour.

In radial quantization, L_n is the conserved charge associated to the conformal transformation $\delta z = z^{n+1}$. To see this, recall that the corresponding Noether current, given in (4.7), is $J(z) = z^{n+1}T(z)$. Moreover, the contour integral $\oint dz$ maps to the integral around spatial slices on the cylinder. This tells us that L_n is the conserved charge where "conserved" means that it is constant under time evolution on the cylinder, or under radial evolution on the plane. Similarly, $\tilde{L}_n$ is the conserved charge associated to the conformal transformation $\delta\bar{z} = \bar{z}^{n+1}$.

When we go to the quantum theory, conserved charges become generators for the transformation. Thus the operators L_n and $\tilde{L}_n$ generate the conformal transformations $\delta z = z^{n+1}$ and $\delta\bar{z} = \bar{z}^{n+1}$. They are known as the *Virasoro* generators. In particular, our two favorite conformal transformations are

- L_{-1} and $\tilde{L}_{-1}$ generate translations in the plane.
- L_0 and $\tilde{L}_0$ generate scaling and rotations.

The Hamiltonian of the system — which measures the energy of states on the cylinder — is mapped into the dilatation operator on the plane. When acting on states of the theory, this operator is represented as

$$D = L_0 + \tilde{L}_0$$

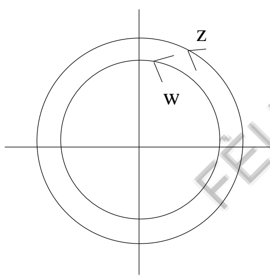
4.5.2 The Virasoro Algebra

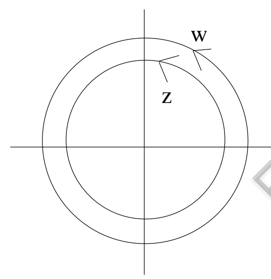
If we have some number of conserved charges, the first thing that we should do is compute their algebra. Representations of this algebra then classify the states of the theory. (For example, think angular momentum in the hydrogen atom). For conformal symmetry, we want to determine the algebra obeyed by the L_n generators. It's a nice fact that the commutation relations are actually encoded TT OPE. Let's see how this works.

We want to compute $[L_m, L_n]$. Let's write L_m as a contour integral over $\oint dz$ and L_n as a contour integral over $\oint dw$. (Note: both z and w denote coordinates on the complex plane now). The commutator is

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

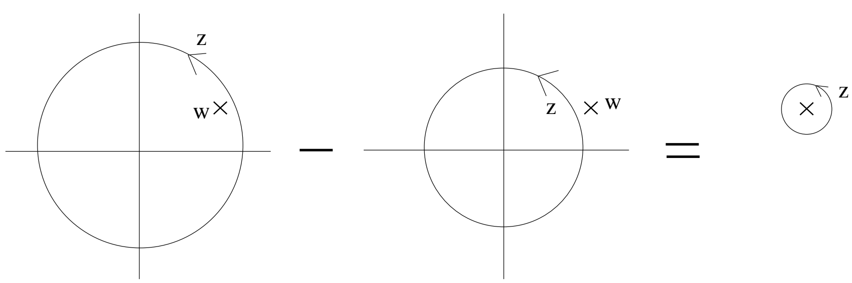
What does this actually mean?! We need to remember that all operator equations are to be viewed as living inside time-ordered correlation functions. Except, now we're working on the z-plane, this statement has transmuted into radially ordered correlation functions: outies to the left, innies to the right.

So $L_m L_n$ means 

while $L_n L_m$ means 

.

The trick to computing the commutator is to first fix w and do the $\oint dz$ integrations. The resulting contour is,



In other words, we do the z-integration around a fixed point w, to get

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

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

To compute the residue at $z = w$, we first need to Taylor expand z^{m+1} about the point w,

$$z^{m+1} = w^{m+1} + (m+1)w^m(z-w) + \frac{1}{2}m(m+1)w^{m-1}(z-w)^2$$
$$+\frac{1}{6}m(m^2-1)w^{m-2}(z-w)^3 + \dots$$

The residue then picks up a contribution from each of the three terms,

$$[L_m, L_n] = \oint \frac{dw}{2\pi i}\ w^{n+1} \left[w^{m+1}\partial T(w) + 2(m+1)w^m\, T(w) + \frac{c}{12}m(m^2-1)w^{m-2} \right]$$

To proceed, it is simplest to integrate the first term by parts. Then we do the w-integral. But for both the first two terms, the resulting integral is of the form (4.48) and gives us L_{m+n}. For the third term, we pick up the pole. The end result is

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

This is the *Virasoro algebra*. It's quite famous. The $\tilde{L}_n$'s satisfy exactly the same algebra, but with c replaced by $\tilde{c}$. Of course, $[L_n, \tilde{L}_m] = 0$. The appearance of c as an extra term in the Virasoro algebra is the reason it is called the "central charge". In general, a central charge is an extra term in an algebra that commutes with everything else.

Conformal = Diffeo + Weyl

We can build some intuition for the Virasoro algebra. We know that the L_n's generate conformal transformations $\delta z = z^{n+1}$. Let's consider something closely related: a coordinate transformation $\delta z = z^{n+1}$. These are generated by the vector fields

$$l_n = z^{n+1}\partial_z \tag{4.49}$$

But it's a simple matter to compute their commutation relations:

$$[l_n, l_m] = (m-n)l_{m+n}$$

So this is giving us the first part of the Virasoro algebra. But what about the central term? The key point to remember is that, as we stressed at the beginning of this chapter, a conformal transformation is not just a reparameterization of the coordinates: it is a reparameterization, followed by a compensating Weyl rescaling. The central term in the Virasoro algebra is due to the Weyl rescaling.

4.5.3 Representations of the Virasoro Algebra

With the algebra of conserved charges at hand, we can now start to see how the conformal symmetry classifies the states into representations.

Suppose that we have some state $|\psi\rangle$ that is an eigenstate of L_0 and $\tilde{L}_0$.

$$L_0\,|\psi\rangle = h\,|\psi\rangle \quad , \quad \tilde{L}_0\,|\psi\rangle = \tilde{h}\,|\psi\rangle$$

Back on the cylinder, this corresponds to some state with energy

$$\frac{E}{2\pi} = h + \tilde{h} - \frac{c+\tilde{c}}{24}$$

For this reason, we'll refer to the eigenvalues h and $\tilde{h}$ as the energy of the state. By acting with the L_n operators, we can get further states with eigenvalues

$$L_0 L_n\,|\psi\rangle = (L_n L_0 - nL_n)\,|\psi\rangle = (h-n)L_n\,|\psi\rangle$$

This tells us that L_n are raising and lowering operators depending on the sign of n. When $n > 0$, L_n lowers the energy of the state and L_{-n} raises the energy of the state. If the spectrum is to be bounded below, there must be some states which are annihilated by all L_n and $\tilde{L}_n$ for $n > 0$. Such states are called *primary*. They obey

$$L_n\,|\psi\rangle = \tilde{L}_n\,|\psi\rangle = 0 \quad \text{for all } n > 0$$

In the language of representation theory, they are also called highest weight states. They are the states of lowest energy.

Representations of the Virasoro algebra can now be built by acting on the primary states with raising operators L_{-n} with $n > 0$. Obviously this results in an infinite tower of states. All states obtained in this way are called *descendants*. From an initial primary state $|\psi\rangle$, the tower fans out...

$$\begin{gathered}
|\psi\rangle \\
L_{-1}\,|\psi\rangle \\
L_{-1}^2\,|\psi\rangle \;,\; L_{-2}\,|\psi\rangle \\
L_{-1}^3\,|\psi\rangle \;,\; L_{-1}L_{-2}\,|\psi\rangle \;,\; L_{-3}\,|\psi\rangle
\end{gathered}$$

The whole set of states is called a *Verma* module. They are the irreducible representations of the Virasoro algebra. This means that if we know the spectrum of primary states, then we know the spectrum of the whole theory.

Some comments:

- The vacuum state $|0\rangle$ has $h = 0$. This state obeys

$$L_n\,|0\rangle = 0 \quad \text{for all } n \geq -1 \tag{4.50}$$

Note that this state preserves the maximum number of symmetries: like all primary states, it is annihilated by L_n with $n > 0$, but it is also annihilated by L_0 and L_{-1}. This fits with our intuition that the vacuum state should be invariant under as many symmetries as possible. You might think that we could go further and require that the vacuum state obeys $L_n\,|\,0\rangle = 0$ for all n. But that isn't consistent with the central charge term in Virasoro algebra. The requirements (4.50) are the best we can do.

- This discussion should be ringing bells. We saw something very similar in the covariant quantization of the string, where we imposed conditions (2.6) as constraints. We will see the connection between the primary states and the spectrum of the string in Section 5.

- There's a subtlety that you should be aware of: the states in the Verma module are not necessarily all independent. It could be that some linear combination of the states vanishes. This linear combination is known as a null state. The existence of null states depends on the values of h and c. For example, suppose that we are in a theory in which the central charge is $c = 2h(5 - 8h)/(2h + 1)$, where h is the energy of a primary state $|\psi\rangle$. Then it is simple to check that the following combination has vanishing norm:

$$L_{-2}\,|\psi\rangle - \frac{3}{2(2h+1)}\,L_{-1}^2\,|\psi\rangle \tag{4.51}$$

- There is a close relationship between the primary states and the primary operators defined in Section 4.2.3. In fact, the energies h and $\tilde{h}$ of primary states will turn out to be exactly the weights of primary operators in the theory. This connection will be described in Section 4.6.

4.5.4 Consequences of Unitarity

There is one physical requirement that a theory must obey which we have so far neglected to mention: *unitarity*. This is the statement that probabilities are conserved when we are in Minkowski signature spacetime. Unitarity follows immediately if we have a Hermitian Hamiltonian which governs time evolution. But so far our discussion has been somewhat algebraic and we've not enforced this condition. Let's do so now.

We retrace our footsteps back to the Euclidean cylinder and then back again to the Minkowski cylinder where we can ask questions about time evolution. Here the Hamiltonian density takes the form

$$\mathcal{H} = T_{ww} + T_{\bar{w}\bar{w}} = \sum_n L_n e^{-in\sigma^+} + \tilde{L}_n e^{-in\sigma^-}$$

So for the Hamiltonian to be Hermitian, we require

$$L_n = L_{-n}^\dagger$$

This requirement imposes some strong constraints on the structure of CFTs. Here we look at a couple of trivial, but important, constraints that arise due to unitarity and the requirement that the physical Hilbert space does not contain negative norm states.

- $h \geq 0$: This fact follows from looking at the norm,

$$|L_{-1}\,|\psi\rangle|^2 = \langle\psi|\,L_{+1}L_{-1}\,|\psi\rangle = \langle\psi|\,[L_{+1}, L_{-1}]\,|\psi\rangle = 2h\,\langle\psi|\,\psi\rangle \geq 0$$

  The only state with $h = 0$ is the vacuum state $|\,0\rangle$.

- $c > 0$: To see this, we can look at

$$|L_{-n}\,|0\rangle|^2 = \langle 0|\,[L_n, L_{-n}]\,|0\rangle = \frac{c}{12}n(n^2-1) \geq 0 \qquad (4.52)$$

  So $c \geq 0$. If $c = 0$, the only state in the vacuum module is the vacuum itself. It turns out that, in fact, the only state in the whole theory is the vacuum itself. Any non-trivial CFT has $c > 0$.

There are many more requirements of this kind that constrain the theory. In fact, it turns out that for CFTs with $c < 1$ these requirements are enough to classify and solve all theories.

4.6 The State-Operator Map

In this section we describe one particularly important aspect of conformal field theories: a map between states and local operators.

Firstly, let's get some perspective. In a typical quantum field theory, the states and local operators are very different objects. While local operators live at a point in spacetime, the states live over an entire spatial slice. This is most clear if we write down a Schrödinger-style wavefunction. In field theory, this object is actually a wave-functional, $\Psi[\phi(\sigma)]$, describing the probability for every field configuration $\phi(\sigma)$ at each point σ in space (but at a fixed time).

Given that states and local operators are such very different beasts, it's a little surprising that in a CFT there is an isomorphism between them: it's called the state-operator map. The key point is that the distant past in the cylinder gets mapped to a single point $z = 0$ in the complex plane. So specifying a state on the cylinder in the far past is equivalent to specifying a local disturbance at the origin.

To make this precise, we need to recall how to write down wavefunctions using path integrals. Different states are computed by putting different boundary conditions on the functional integral. Let's start by returning to quantum mechanics and reviewing a few simple facts. The propagator for a particle to move from position x_i at time τ_i to position x_f at time τ_f is given by

$$G(x_f, x_i) = \int_{x(\tau_i)=x_i}^{x(\tau_f)=x_f} \mathcal{D}x \; e^{iS}$$

This means that if our system starts off in some state described by the wavefunction $\psi_i(x_i)$ at time τ_i then (ignoring the overall normalization) it evolves to the state

$$\psi_f(x_f, \tau_f) = \int dx_i \, G(x_f, x_i) \, \psi_i(x_i, \tau_i)$$

There are two lessons to take from this. Firstly, to determine the value of the wavefunction at a given point x_f, we evaluate the path integral restricting to paths which satisfy $x(\tau_f) = x_f$. Secondly, the initial state $\psi(x_i)$ acts as a weighting factor for the integral over initial boundary conditions.

Let's now write down the same formula in a field theory, where we're dealing with wavefunctionals. We'll work with the Euclidean path integral on the cylinder. If we start with some state $\Psi_i[\phi_i(\sigma)]$ at time τ_i, then it will evolve to the state

$$\Psi_f[\phi_f(\sigma), \tau_f] = \int \mathcal{D}\phi_i \int_{\phi(\tau_i)=\phi_i}^{\phi(\tau_f)=\phi_f} \mathcal{D}\phi \; e^{-S[\phi]} \; \Psi_i[\phi_i(\sigma), \tau_i]$$

How do we write a similar expression for states after the map to the complex plane? Now the states are defined on circles of constant radius, say $|z| = r$, and evolution is governed by the dilatation operator. Suppose the initial state is defined at $|z| = r_i$. In the path integral, we integrate over all fields with fixed boundary conditions $\phi(r_i) = \phi_i$ and $\phi(r_f) = \phi_f$ on the two edges of the annulus shown in the figure,

$$\Psi_f[\phi_f(\sigma), r_f] = \int \mathcal{D}\phi_i \int_{\phi(r_i)=\phi_i}^{\phi(r_f)=\phi_f} \mathcal{D}\phi \; e^{-S[\phi]} \; \Psi_i[\phi_i(\sigma), r_i]$$

This is the traditional way to define a state in field theory, albeit with a slight twist because we're working in radial quantization. We see that the effect of the initial state is to change the weighting of the path integral over the inner ring at $|z| = r_i$.

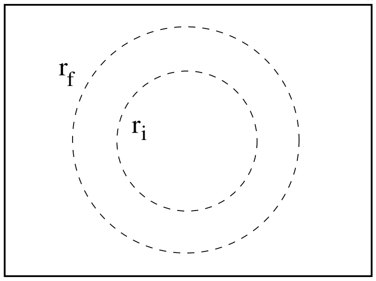


Figure 26:

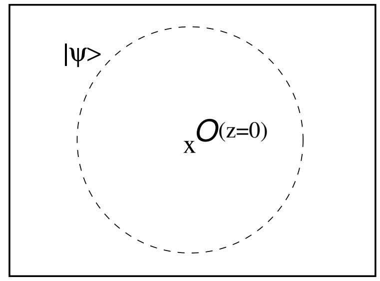


Figure 27:

Let's now see what happens as we take the initial state back to the far past and, ultimately, to $z = 0$? We must now integrate over the whole disc $|z| \leq r_f$, rather than the annulus. The only effect of the initial state is now to change the weighting of the path integral at the point $z = 0$. But that's exactly what we mean by a local operator inserted at that point. This means that each local operator $\mathcal{O}(z = 0)$ defines a different state in the theory,

$$\Psi[\phi_f; r] = \int^{\phi(r)=\phi_f} \mathcal{D}\phi \; e^{-S[\phi]} \, \mathcal{O}(z = 0)$$

We're now integrating over all field configurations within the disc, including all possible values of the field at $z = 0$, which is analogous to integrating over the boundary conditions $\int \mathcal{D}\phi_i$ on the inner circle.

- The state-operator map is only true in conformal field theories where we can map the cylinder to the plane. It also holds in conformal field theories in higher dimensions (where $\mathbf{R} \times \mathbf{S}^{D-1}$ can be mapped to the plane $\mathbf{R}^D$). In non-conformal field theories, a typical local operator creates many different states.

- The state-operator map does not say that the number of states in the theory is equal to the number of operators: this is never true. It does say that the states are in one-to-one correspondence with the *local* operators.

- You might think that you've seen something like this before. In the canonical quantization of free fields, we create states in a Fock space by acting with creation operators. That's *not* what's going on here! The creation operators are just about as far from local operators as you can get. They are the Fourier transforms of local operators.

- There's a special state that we can create this way: the vacuum. This arises by inserting the identity operator **1** into the path integral. Back in the cylinder picture, this just means that we propagate the state back to time $\tau = -\infty$ which is a standard trick used in the Euclidean path integral to project out all but the ground state. For this reason the vacuum is sometimes referred to, in operator notation, as $|\mathbf{1}\rangle$.

4.6.1 Some Simple Consequences

Let's use the state-operator map to wrap up a few loose ends that have arisen in our study of conformal field theory.

Firstly, we've defined two objects that we've called "primary": states and operators. The state-operator map relates the two. Consider the state $|\mathcal{O}\rangle$, built from inserting a primary operator $\mathcal{O}$ into the path integral at $z = 0$. We can look at,

$$\begin{aligned} L_n\,|\mathcal{O}\rangle &= \oint \frac{dz}{2\pi i}\; z^{n+1}\, T(z)\, \mathcal{O}(z=0) \\ &= \oint \frac{dz}{2\pi i}\; z^{n+1} \left(\frac{h\mathcal{O}}{z^2} + \frac{\partial \mathcal{O}}{z} + \ldots\right) \end{aligned} \tag{4.53}$$

You may wonder what became of the path integral $\int \mathcal{D}\phi\, e^{-S[\phi]}$ in this expression. The answer is that it's still implicitly there. Remember that operator expressions such as (4.48) are always taken to hold inside correlation functions. But putting an operator in the correlation function is the same thing as putting it in the path integral, weighted with $e^{-S[\phi]}$.

From (4.53) we can see the effect of various generators on states

- $L_{-1}\,|\mathcal{O}\rangle = |\partial\mathcal{O}\rangle$: In fact, this is true for all operators, not just primary ones. It is expected since L_{-1} is the translation generator.
- $L_0\,|\mathcal{O}\rangle = h\,|\mathcal{O}\rangle$: This is true of any operator with well defined transformation under scaling.
- $L_n\,|\mathcal{O}\rangle = 0$ for all $n > 0$. This is true only of primary operators $\mathcal{O}$. Moreover, it is our requirement for $|\mathcal{O}\rangle$ to be a primary state.

This has an important consequence. We stated earlier that one of the most important things to compute in a CFT is the spectrum of weights of primary operators. This seems like a slightly obscure thing to do. But now we see that it has a much more direct, physical meaning. It is the spectrum of energy and angular momentum of states of the theory defined on the cylinder.

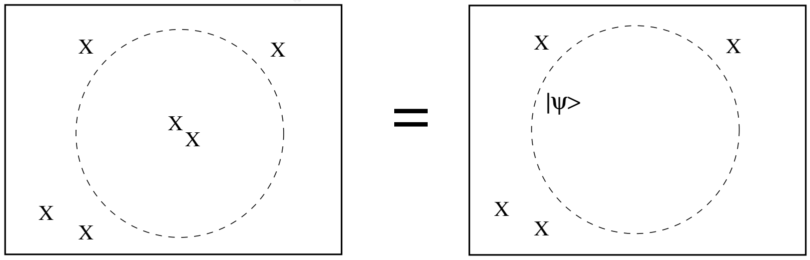


Figure 28:

Another loose end: when defining operators which carry specific weight, we made the statement that we could always work in a basis of operators which have specified eigenvalues under D and L. This follows immediately from the statement that we can always find a basis of eigenstates of H and L on the cylinder.

Finally, we can use this idea of the state-operator map to understand why the OPE works so well in conformal field theories. Suppose that we're interested in some correlation function, with operator insertions as shown in the figure. The statement of the OPE is that we can replace the two inner operators by a sum of operators at $z = 0$, *independent* of what's going on outside of the dotted line. As an operator statement, that sounds rather surprising. But this follows by computing the path integral up to the dotted line, by which point the only effect of the two operators is to determine what state we have. This provides us a way of understanding why the OPE is exact in CFTs, with a radius of convergence equal to the next-nearest insertion.

4.6.2 Our Favourite Example: The Free Scalar Field

Let's illustrate the state-operator map by returning yet again to the free scalar field. On a Euclidean cylinder, we have the mode expansion

$$X(w,\bar{w}) = x + \alpha' p\,\tau + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\,\left(\alpha_n\,e^{inw} + \tilde{\alpha}_n\,e^{in\bar{w}}\right)$$

where we retain the requirement of reality in Minkowski space, which gave us $\alpha_n^\star = \alpha_{-n}$ and $\tilde{\alpha}_n^\star = \tilde{\alpha}_{-n}$. We saw in Section 4.3 that X does not have good conformal properties. Before transforming to the $z = e^{-iw}$ plane, we should work with the primary field on the cylinder,

$$\partial_w X(w,\bar{w}) = -\sqrt{\frac{\alpha'}{2}}\sum_n \alpha_n\,e^{inw} \qquad \text{with } \alpha_0 \equiv i\sqrt{\frac{\alpha'}{2}}p$$

Since ∂X is a primary field of weight $h = 1$, its transformation to the plane is given by (4.18) and reads

$$\partial_z X(z) = \left(\frac{\partial z}{\partial w}\right)^{-1} \partial_w X(w) = -i\sqrt{\frac{\alpha'}{2}} \sum_n \frac{\alpha_n}{z^{n+1}}$$

and similar for $\bar{\partial} X$. Inverting this gives an equation for α_n as a contour integral,

$$\alpha_n = i\sqrt{\frac{2}{\alpha'}} \oint \frac{dz}{2\pi i}\, z^n\, \partial X(z) \tag{4.54}$$

Just as the TT OPE allowed us to determine the $[L_m, L_n]$ commutation relations in the previous section, so the $\partial X \partial X$ OPE contains the information about the $[\alpha_m, \alpha_n]$ commutation relations. The calculation is straightforward,

$$\begin{aligned}
[\alpha_m, \alpha_n] &= -\frac{2}{\alpha'}\left(\oint \frac{dz}{2\pi i}\oint \frac{dw}{2\pi i} - \oint \frac{dw}{2\pi i}\oint \frac{dz}{2\pi i}\right) z^m w^n\, \partial X(z)\, \partial X(w) \\
&= -\frac{2}{\alpha'}\oint \frac{dw}{2\pi i}\, \text{Res}_{\,z=w}\left[z^m w^n \left(\frac{-\alpha'/2}{(z-w)^2} + \ldots\right)\right] \\
&= m \oint \frac{dw}{2\pi i}\, w^{m+n-1} = m\delta_{m+n,0}
\end{aligned}$$

where, in going from the second to third line, we have Taylor expanded z around w. Hearteningly, the final result agrees with the commutation relation (2.2) that we derived in string theory using canonical quantization.

The State-Operator Map for the Free Scalar Field

Let's now look at the map between states and local operators. We know from canonical quantization that the Fock space is defined by acting with creation operators α_{-m} with $m > 0$ on the vacuum $|0\rangle$. The vacuum state itself obeys $\alpha_m|0\rangle = 0$ for $m > 0$. Finally, there is also the zero mode $\alpha_0 \sim p$ which provides all states with another quantum number. A general state is given by

$$\prod_{m=1}^{\infty} \alpha_{-m}^{k_m}\, |0; p\rangle$$

Let's try and recover these states by inserting operators into the path integral. Our first task is to check whether the vacuum state is indeed equivalent to the insertion of the identity operator. In other words, is the ground state wavefunctional of the theory on the circle $|z| = r$ really given by

$$\Psi_0[X_f] = \int^{X_f(r)} \mathcal{D}X\, e^{-S[X]} \qquad ? \tag{4.55}$$

We want to check that this satisfies the definition of the vacuum state, namely $\alpha_m|0\rangle = 0$ for $m > 0$. How do we act on the wavefunctional with an operator? We should still integrate over all field configurations $X(z,\bar{z})$, subject to the boundary conditions at $X(|z| = r) = X_f$. But now we should insert the contour integral (4.54) at some $|w| < r$ (because, after all, the state is only going to vanish after we've hit it with α_m, not before!). So we look at

$$\alpha_m \Psi_0[X_f] = \int^{X_f} \mathcal{D}X\ e^{-S[X]} \oint \frac{dw}{2\pi i}\, w^m \partial X(w)$$

The path integral is weighted by the action (4.19) for a free scalar field. If a given configuration diverges somewhere inside the disc $|z| < r$, then the action also diverges. This ensures that only smooth functions $\partial X(z)$, which have no singularity inside the disc, contribute. But for such functions we have

$$\oint \frac{dw}{2\pi i}\, w^m \partial X(w) = 0 \qquad \text{for all } m \geq 0$$

So the state (4.55) is indeed the vacuum state. In fact, since α_0 also annihilates this state, it is identified as the vacuum state with vanishing momentum.

What about the excited states of the theory?

Claim: $\alpha_{-m}|0\rangle = |\partial^m X\rangle$. By which we mean that the state $\alpha_{-m}|0\rangle$ can be built from the path integral,

$$\alpha_{-m}|0\rangle = \int \mathcal{D}X\ e^{-S[X]}\ \partial^m X(z=0) \tag{4.56}$$

Proof: We can check this by acting on $|\partial^m X\rangle$ with the annihilation operators α_n.

$$\alpha_n\, |\partial^m X\rangle \sim \int^{X_f(r)} \mathcal{D}X\ e^{-S[X]}\ \oint \frac{dw}{2\pi i}\ w^n\, \partial X(w)\, \partial^m X(z=0)$$

We can focus on the operator insertions and use the OPE (4.23). We drop the path integral and just focus on the operator equation (because, after all, operator equations only make sense in correlation functions which is the same thing as in path integrals). We have

$$\oint \frac{dw}{2\pi i}\ w^n\, \partial_z^{m-1} \frac{1}{(w-z)^2}\bigg|_{z=0} = m! \oint \frac{dw}{2\pi i}\ w^{n-m-1} = 0 \qquad \text{unless } m = n$$

This confirms that the state (4.56) has the right properties. □

Finally, we should worry about the zero mode, or momentum $\alpha_0 \sim p$. It is simple to show using the techniques above (together with the OPE (4.26)) that the momentum of a state arises by the insertion of the primary operator e^{ipX}. For example,

$$|0;p\rangle \sim \int \mathcal{D}X \; e^{-S[X]} \; e^{ipX(z=0)} \; .$$

4.7 Brief Comments on Conformal Field Theories with Boundaries

The open string lives on the infinite strip with spatial coordinate $\sigma \in [0, \pi]$. Here we make just a few brief comments on the corresponding conformal field theories.

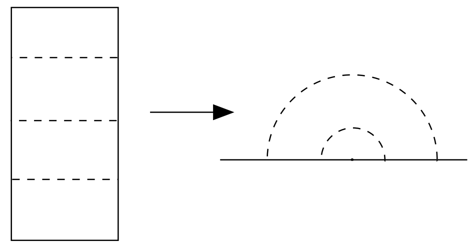


Figure 29:

As before, we can define the complex coordinate $w = \sigma + i\tau$ and make the conformal map

$$z = e^{-iw}$$

This time the map takes us to the upper-half plane: $\text{Im}z \geq 0$. The end points of the string are mapped to the real axis, $\text{Im}z = 0$.

Much of our previous discussion goes through as before. But now we need to take care of boundary conditions at $\text{Im}z = 0$. Let's first look at $T_{\alpha\beta}$. Recall that the stress-energy tensor exists because of translational invariance. We still have translational invariance in the direction parallel to the boundary — let's call the associated tangent vector t^α. But translational invariance is broken perpendicular to the boundary — we call the normal vector n^α. The upshot of this is that $T_{\alpha\beta}t^\beta$ remains a conserved current.

To implement Neumann boundary conditions, we insist that none of the current flows out of the boundary. The condition is

$$T_{\alpha\beta} n^\alpha t^\beta = 0 \qquad \text{at Im}z = 0$$

In complex coordinates, this becomes

$$T_{zz} = T_{\bar{z}\bar{z}} \qquad \text{at Im}z = 0$$

There's a simple way to implement this: we extend the definition of T_{zz} from the upper-half plane to the whole complex plane by defining

$$T_{zz}(z) = T_{\bar{z}\bar{z}}(\bar{z})$$

For the closed string we had both functions T and $\bar{T}$ in the whole plane. But for the open string, we have just one of these – say, T, — in the whole plane. This contains the same information as both T and $\bar{T}$ in the upper-half plane. It's simpler to work in the whole plane and focus just on T. Correspondingly, we now have just a single set of Virasoro generators,

$$L_n = \oint \frac{dz}{2\pi i} \, z^{n+1} \, T_{zz}(z)$$

There is no independent $\tilde{L}_n$ for the open string.

A similar doubling trick works when computing the propagator for the free scalar field. The scalar field $X(z, \bar{z})$ is only defined in the upper-half plane. Suppose we want to implement Neumann boundary conditions. Then the propagator is defined by

$$\langle X(z, \bar{z}) \, X(w, \bar{w}) \rangle = G(z, \bar{z}; w, \bar{w})$$

which obeys $\partial^2 G = -2\pi\alpha' \, \delta(z - w, \bar{z} - \bar{w})$ subject to the boundary condition

$$\partial_\sigma \, G(z, \bar{z}; w, \bar{w})\big|_{\sigma=0} = 0$$

But we solve problems like this in our electrodynamics courses. A useful way of proceeding is to introduce an "image charge" in the lower-half plane. We now let $X(z, \bar{z})$ vary over the whole complex plane with its dynamics governed by the propagator

$$G(z, \bar{z}; w, \bar{w}) = -\frac{\alpha'}{2} \ln |z - w|^2 - \frac{\alpha'}{2} \ln |z - \bar{w}|^2 \tag{4.57}$$

Much of the remaining discussion of CFTs carries forward with only minor differences. However, there is one point that is simple but worth stressing because it will be of importance later. This concerns the state-operator map. Recall the logic that leads us to this idea: we consider a state at fixed time on the strip and propagate it back to past infinity $\tau \to -\infty$. After the map to the half-plane, past infinity is again the origin. But now the origin lies on the boundary. We learn that the state-operator map relates states to local operators defined on the boundary.

This fact ensures that theories on a strip have fewer states than those on the cylinder. For example, for a free scalar field, Neumann boundary conditions require $\partial X = \bar{\partial} X$ at $\text{Im} z = 0$. (This follows from the requirement that $\partial_\sigma X = 0$ at $\sigma = 0, \pi$ on the strip). On the cylinder, the operators ∂X and $\bar{\partial} X$ give rise to different states; on the strip they give rise to the same state. This, of course, mirrors what we've seen for the quantization of the open string where boundary conditions mean that we have only half the oscillator modes to play with.

5. The Polyakov Path Integral and Ghosts

At the beginning of the last chapter, we stressed that there are two very different interpretations of conformal symmetry depending on whether we're thinking of a fixed 2d background or a dynamical 2d background. In applications to statistical physics, the background is fixed and conformal symmetry is a global symmetry. In contrast, in string theory the background is dynamical. Conformal symmetry is a gauge symmetry, a remnant of diffeomorphism invariance and Weyl invariance.

But gauge symmetries are not symmetries at all. They are redundancies in our description of the system. As such, we can't afford to lose them and it is imperative that they don't suffer an anomaly in the quantum theory. At worst, theories with gauge anomalies make no sense. (For example, Yang-Mills theory coupled to only left-handed fundamental fermions is a nonsensical theory for this reason). At best, it may be possible to recover the quantum theory, but it almost certainly has nothing to do with the theory that you started with.

Piecing together some results from the previous chapter, it looks like we're in trouble. We saw that the Weyl symmetry is anomalous since the expectation value of the stress-energy tensor takes different values on backgrounds related by a Weyl symmetry:

$$\langle T^\alpha_{\ \alpha} \rangle = -\frac{c}{12}\, R$$

On fixed backgrounds, that's merely interesting. On dynamical backgrounds, it's fatal. What can we do? It seems that the only way out is to ensure that our theory has $c = 0$. But we've already seen that $c > 0$ for all non-trivial, unitary CFTs. We seem to have reached an impasse. In this section we will discover the loophole. It turns out that we do indeed require $c = 0$, but there's a way to achieve this that makes sense.

5.1 The Path Integral

In Euclidean space the Polyakov action is given by,

$$S_{\text{Poly}} = \frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{g}\; g^{\alpha\beta}\, \partial_\alpha X^\mu\, \partial_\beta X^\nu\; \delta_{\mu\nu}$$

From now on, our analysis of the string will be in terms of the path integral[6]. We integrate over all embedding coordinates X^μ and all worldsheet metrics $g_{\alpha\beta}$. Schematically,

[6] The analysis of the string path integral was first performed by Polyakov in "*Quantum geometry of bosonic strings*,", Phys. Lett. B **103**, 207 (1981). The paper weighs in at a whopping 4 pages. As a follow-up, he took another 2.5 pages to analyze the superstring in "*Quantum geometry of fermionic strings*," Phys. Lett. B **103**, 211 (1981).

the path integral is given by,

$$Z = \frac{1}{\mathrm{Vol}} \int \mathcal{D}g\mathcal{D}X \; e^{-S_{\mathrm{Poly}}[X,g]}$$

The "Vol" term is all-important. It refers to the fact that we shouldn't be integrating over all field configurations, but only those physically distinct configurations not related by diffeomorphisms and Weyl symmetries. Since the path integral, as written, sums over all fields, the "Vol" term means that we need to divide out by the volume of the gauge action on field space.

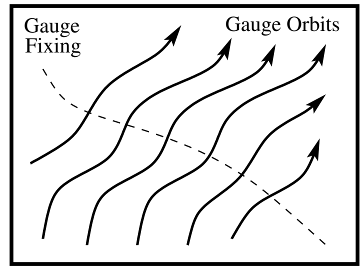


Figure 30:

To make the situation more explicit, we need to split the integration over all field configurations into two pieces: those corresponding to physically distinct configurations — schematically depicted as the dotted line in the figure — and those corresponding to gauge transformations — which are shown as solid lines. Dividing by "Vol" simply removes the piece of the partition function which comes from integrating along the solid-line gauge orbits.

In an ordinary integral, if we change coordinates then we pick up a Jacobian factor for our troubles. The path integral is no different. We want to decompose our integration variables into physical fields and gauge orbits. The tricky part is to figure out what Jacobian we get. Thankfully, there is a standard method to determine the Jacobian, first introduced by Faddeev and Popov. This method works for all gauge symmetries, including Yang-Mills and you will also learn about it in the "Advanced Quantum Field Theory" course.

5.1.1 The Faddeev-Popov Method

We have two gauge symmetries: diffeomorphisms and Weyl transformations. We will schematically denote both of these by ζ. The change of the metric under a general gauge transformation is $g \to g^{\zeta}$. This is shorthand for,

$$g_{\alpha\beta}(\sigma) \;\longrightarrow\; g^{\zeta}_{\alpha\beta}(\sigma') = e^{2\omega(\sigma)} \frac{\partial\sigma^{\gamma}}{\partial\sigma'^{\alpha}} \frac{\partial\sigma^{\delta}}{\partial\sigma'^{\beta}} \, g_{\gamma\delta}(\sigma)$$

In two dimensions these gauge symmetries allow us to put the metric into any form that we like — say, $\hat{g}$. This is called the fiducial metric and will represent our choice of gauge fixing. Two caveats:

- Firstly, it's not true that we can put any 2d metric into the form $\hat{g}$ of our choosing. This is only true locally. Globally, it remains true if the worldsheet has the topology of a cylinder or a sphere, but not for higher genus surfaces. We'll revisit this issue in Section 6.

- Secondly, fixing the metric locally to $\hat{g}$ does not fix all the gauge symmetries. We still have the conformal symmetries to deal with. We'll revisit this in the Section 6 as well.

Our goal is to only integrate over physically inequivalent configurations. To achieve this, first consider the integral over the gauge orbit of $\hat{g}$. For some value of the gauge transformation ζ, the configuration g^ζ will coincide with our original metric g. We can put a delta-function in the integral to get

$$\int \mathcal{D}\zeta \; \delta(g - \hat{g}^\zeta) = \Delta_{FP}^{-1}[g] \tag{5.1}$$

This integral isn't equal to one because we need to take into account the Jacobian factor. This is analogous to the statement that $\int dx\, \delta(f(x)) = 1/|f'|$, evaluated at points where $f(x) = 0$. In the above equation, we have written this Jacobian factor as Δ_{FP}^{-1}. The inverse of this, namely Δ_{FP}, is called the *Faddeev-Popov determinant*. We will evaluate it explicitly shortly. Some comments:

- This whole procedure is rather formal and runs into the usual difficulties with trying to define the path integral. Just as for Yang-Mills theory, we will find that it results in sensible answers.

- We will assume that our gauge fixing is good, meaning that the dotted line in the previous figure cuts through each physically distinct configuration exactly once. Equivalently, the integral over gauge transformations $\mathcal{D}\zeta$ clicks exactly once with the delta-function and we don't have to worry about discrete ambiguities (known as Gribov copies in QCD).

- The measure is taken to be the analogue of the Haar measure for Lie groups, invariant under left and right actions

$$\mathcal{D}\zeta = \mathcal{D}(\zeta'\zeta) = \mathcal{D}(\zeta\zeta')$$

When gauge fixing in Yang-Mills theory, the first thing we do is prove that the Faddeev-Popov determinant Δ_{FP} is gauge invariant. However, our route here is a little more subtle. As we've stressed above, the Weyl anomaly means that our original theory actually fails to be gauge invariant. We will see that the Faddeev-Popov determinant also fails but can, in certain circumstances, cancel the original failure leaving behind a well-defined theory.

The Faddeev-Popov procedure starts by inserting a factor of unity into the path integral, in the guise of

$$1 = \Delta_{FP}[g] \int \mathcal{D}\zeta \; \delta(g - \hat{g}^{\zeta})$$

We'll call the resulting path integral expression $Z[\hat{g}]$ since it depends on the choice of fiducial metric $\hat{g}$. The first thing we do is use the $\delta(g - \hat{g}^{\zeta})$ delta-function to do the integral over metrics,

$$\begin{aligned} Z[\hat{g}] &= \frac{1}{\text{Vol}} \int \mathcal{D}\zeta \mathcal{D}X \mathcal{D}g \; \Delta_{FP}[g] \, \delta(g - \hat{g}^{\zeta}) \, e^{-S_{\text{Poly}}[X,g]} \\ &= \frac{1}{\text{Vol}} \int \mathcal{D}\zeta \mathcal{D}X \; \Delta_{FP}[\hat{g}^{\zeta}] \, e^{-S_{\text{Poly}}[X,\hat{g}^{\zeta}]} \end{aligned} \tag{5.2}$$

At this stage the integrand depends on $\hat{g}^{\zeta}$, where ζ is shorthand for a diffeoemorphism and Weyl transformation. Everything in the equation is invariant under diffeomorphisms, but Weyl transformations are another matter. We know that quantum theory $\int \mathcal{D}X \, e^{-S_{\text{Poly}}}$ suffers a Weyl anomaly. The action S_{Poly} is invariant under Weyl rescalings, so the subtlety must come from the measure. Meanwhile, anticipating what's to come, we will find a similar issue with the Faddeev-Popov determinant Δ_{FP}.

If, however, we find ourselves in the fortunate situation where the problems cancel then things would work out nicely. In that situation, everything on the right-hand side of (5.2) would be conspire to be invariant under both diffeomorphisms and Weyl transformations and we could write

$$Z[\hat{g}] = \frac{1}{\text{Vol}} \int \mathcal{D}\zeta \mathcal{D}X \; \Delta_{FP}[\hat{g}] \, e^{-S_{\text{Poly}}[X,\hat{g}]}$$

But now, nothing depends on the gauge transformation ζ. Indeed, this is precisely the integration over the gauge orbits that we wanted to isolate and it cancels the "Vol" factor sitting outside. We're left with

$$Z[\hat{g}] = \int \mathcal{D}X \; \Delta_{FP}[\hat{g}] \, e^{-S_{\text{Poly}}[X,\hat{g}]} \tag{5.3}$$

This is the integral over physically distinct configurations — the dotted line in the previous figure. We see that the Faddeev-Popov determinant is precisely the Jacobian factor that we need.

Clearly the above discussion only flies if we find ourselves in a situation in which the theory (5.2) is genuinely Weyl invariant. Our next task is to understand when this happens which means that we need to figure out what becomes of Δ_{FP} when we do a Weyl transformation.

5.1.2 The Faddeev-Popov Determinant

We still need to compute $\Delta_{FP}[\hat{g}]$. It's defined in (5.1). Let's look at gauge transformations ζ which are close to the identity. In this case, the delta-function $\delta(g - \hat{g}^\zeta)$ is going to be non-zero when the metric g is close to the fiducial metric $\hat{g}$. In fact, it will be sufficient to look at the delta-function $\delta(\hat{g} - \hat{g}^\zeta)$, which is only non-zero when $\zeta = 0$. We take an infinitesimal Weyl transformation parameterized by $\omega(\sigma)$ and an infinitesimal diffeomorphism $\delta\sigma^\alpha = v^\alpha(\sigma)$. The change in the metric is

$$\delta\hat{g}_{\alpha\beta} = 2\omega\hat{g}_{\alpha\beta} + \nabla_\alpha v_\beta + \nabla_\beta v_\alpha$$

Plugging this into the delta-function, the expression for the Faddeev-Popov determinant becomes

$$\Delta_{FP}^{-1}[\hat{g}] = \int \mathcal{D}\omega\mathcal{D}v\ \delta(2\omega\hat{g}_{\alpha\beta} + \nabla_\alpha v_\beta + \nabla_\beta v_\alpha) \tag{5.4}$$

where we've replaced the integral $\mathcal{D}\zeta$ over the gauge group with the integral $\mathcal{D}\omega\mathcal{D}v$ over the Lie algebra of group since we're near the identity. (We also suppress the subscript on v_α in the measure factor to keep things looking tidy).

At this stage it's useful to represent the delta-function in its integral, Fourier form. For a single delta-function, this is $\delta(x) = \int dp\ \exp(2\pi i p\, x)$. But the delta-function in (5.4) is actually a delta-functional: it restricts a whole function. Correspondingly, the integral representation is in terms of a functional integral,

$$\Delta_{FP}^{-1}[\hat{g}] = \int \mathcal{D}\omega\mathcal{D}v\mathcal{D}\beta\ \exp\left(2\pi i \int d^2\sigma\ \sqrt{\hat{g}}\ \beta^{\alpha\beta}[2\omega\hat{g}_{\alpha\beta} + \nabla_\alpha v_\beta + \nabla_\beta v_\alpha]\right)$$

where $\beta^{\alpha\beta}$ is a symmetric 2-tensor on the worldsheet.

We now simply do the $\int \mathcal{D}\omega$ integral. It doesn't come with any derivatives, so it merely acts as a Lagrange multiplier, setting

$$\beta^{\alpha\beta}\hat{g}_{\alpha\beta} = 0$$

In other words, after performing the ω integral, $\beta^{\alpha\beta}$ is symmetric and traceless. We'll take this to be the definition of $\beta^{\alpha\beta}$ from now on. So, finally we have

$$\Delta_{FP}^{-1}[\hat{g}] = \int \mathcal{D}v\mathcal{D}\beta\ \exp\left(4\pi i \int d^2\sigma\ \sqrt{\hat{g}}\ \beta^{\alpha\beta}\nabla_\alpha v_\beta\right)$$

5.1.3 Ghosts

The previous manipulations give us an expression for Δ_{FP}^{-1}. But we want to invert it to get Δ_{FP}. Thankfully, there's a simple way to achieve this. Because the integrand is quadratic in v and β, we know that the integral computes the inverse determinant of the operator ∇_α. (Strictly speaking, it computes the inverse determinant of the projection of ∇_α onto symmetric, traceless tensors. This observation is important because it means the relevant operator is a square matrix which is necessary to talk about a determinant). But we also know how to write down an expression for the determinant Δ_{FP}, instead of its inverse, in terms of path integrals: we simply need to replace the commuting integration variables with anti-commuting fields,

$$\begin{aligned}\beta_{\alpha\beta} &\longrightarrow b_{\alpha\beta}\\ v^\alpha &\longrightarrow c^\alpha\end{aligned}$$

where b and c are both Grassmann-valued fields (i.e. anti-commuting). They are known as *ghost fields*. This gives us our final expression for the Faddeev-Popov determinant,

$$\Delta_{FP}[g] = \int \mathcal{D}b\mathcal{D}c\ \exp[iS_{\text{ghost}}]$$

where the ghost action is defined to be

$$S_{\text{ghost}} = \frac{1}{2\pi}\int d^2\sigma\sqrt{g}\ b_{\alpha\beta}\nabla^\alpha c^\beta \tag{5.5}$$

and we have chosen to rescale the b and c fields at this last step to get a factor of $1/2\pi$ sitting in front of the action. (This only changes the normalization of the partition function which doesn't matter). Rotating back to Euclidean space, the factor of i disappears. The expression for the full partition function (5.3) is

$$Z[\hat{g}] = \int \mathcal{D}X\mathcal{D}b\mathcal{D}c\ \exp\left(-S_{\text{Poly}}[X,\hat{g}] - S_{\text{ghost}}[b,c,\hat{g}]\right)$$

Something lovely has happened. Although the ghost fields were introduced as some auxiliary constructs, they now appear on the same footing as the dynamical fields X. We learn that gauge fixing comes with a price: our theory has extra ghost fields.

The role of these ghost fields is to cancel the unphysical gauge degrees of freedom, leaving only the $D-2$ transverse modes of X^μ. Unlike lightcone quantization, they achieve this in a way which preserves Lorentz invariance.

Simplifying the Ghost Action

The ghost action (5.5) looks fairly simple. But it looks even simpler if we work in conformal gauge,

$$\hat{g}_{\alpha\beta} = e^{2\omega}\delta_{\alpha\beta}$$

The determinant is $\sqrt{\hat{g}} = e^{2\omega}$. Recall that in complex coordinates, the measure is $d^2\sigma = \frac{1}{2}d^2z$, while we can lower the index on the covariant derivative using $\nabla^z = g^{z\bar{z}}\nabla_{\bar{z}} = 2e^{-2\omega}\nabla_{\bar{z}}$. We have

$$S_{\text{ghost}} = \frac{1}{2\pi}\int d^2z\ \left(b_{zz}\nabla_{\bar{z}}c^z + b_{\bar{z}\bar{z}}\nabla_z c^{\bar{z}}\right)$$

In deriving this, remember that there is no field $b_{z\bar{z}}$ because $b_{\alpha\beta}$ is traceless. Now comes the nice part: the covariant derivatives are actually just ordinary derivatives. To see why this is the case, look at

$$\nabla_{\bar{z}}c^z = \partial_{\bar{z}}c^z + \Gamma^z_{\bar{z}\alpha}c^\alpha$$

But the Christoffel symbols are given by

$$\Gamma^z_{\bar{z}\alpha} = \frac{1}{2}g^{z\bar{z}}\left(\partial_{\bar{z}}g_{\alpha\bar{z}} + \partial_\alpha g_{\bar{z}\bar{z}} - \partial_{\bar{z}}g_{\bar{z}\alpha}\right) = 0 \qquad \text{for } \alpha = z, \bar{z}$$

So in conformal gauge, the ghost action factorizes into two free theories,

$$S_{\text{ghost}} = \frac{1}{2\pi}\int d^2z\ b_{zz}\,\partial_{\bar{z}}c^z + b_{\bar{z}\bar{z}}\,\partial_z c^{\bar{z}}$$

The action doesn't depend on the conformal factor ω. In other words, it is Weyl invariant without any need to change b and c: these are therefore both neutral under Weyl transformations.

(It's worth pointing out that $b_{\alpha\beta}$ and c^α are neutral under Weyl transformations. But if we raise or lower these indices, then the fields pick up factors of the metric. So $b^{\alpha\beta}$ and c_α would not be neutral under Weyl transformations).

5.2 The Ghost CFT

Fixing the Weyl and diffeomorphism gauge symmetries has left us with two new dynamical ghost fields, b and c. Both are Grassmann (i.e. anti-commuting) variables. Their dynamics is governed by a CFT. Define

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

The ghost action is given by

$$S_{\text{ghost}} = \frac{1}{2\pi}\int d^2z\ \left(b\,\bar{\partial}c + \bar{b}\,\partial\bar{c}\right)$$

Which gives the equations of motion

$$\bar{\partial}b = \partial\bar{b} = \bar{\partial}c = \partial\bar{c} = 0$$

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

Before moving onto quantization, there's one last bit of information we need from the classical theory: the stress tensor for the bc ghosts. The calculation is a little bit fiddly. We use the general definition of the stress tensor (4.4), which requires us to return to the theory (5.5) on a general background and vary the metric $g^{\alpha\beta}$. The complications are twofold. Firstly, we pick up a contribution from the Christoffel symbol that is lurking inside the covariant derivative ∇^α. Secondly, we must also remember that $b_{\alpha\beta}$ is traceless. But this is a condition which itself depends on the metric: $b_{\alpha\beta}g^{\alpha\beta} = 0$. To account for this we should add a Lagrange multiplier to the action imposing tracelessness. After correctly varying the metric, we may safely retreat back to flat space where the end result is rather simple. We have $T_{z\bar{z}} = 0$, as we must for any conformal theory. Meanwhile, the holomorphic and anti-holomorphic parts of the stress tensor are given by,

$$T = 2(\partial c)\,b + c\,\partial b \quad , \quad \bar{T} = 2(\bar{\partial}\bar{c})\,\bar{b} + \bar{c}\,\bar{\partial}\bar{b}. \tag{5.6}$$

Operator Product Expansions

We can compute the OPEs of these fields using the standard path integral techniques that we employed in the last chapter. In what follows, we'll just focus on the holomorphic piece of the CFT. We have, for example,

$$0 = \int \mathcal{D}b\mathcal{D}c\ \frac{\delta}{\delta b(\sigma)}\ \left[e^{-S_{\text{ghost}}}\,b(\sigma')\right] = \int \mathcal{D}b\mathcal{D}c\ e^{-S_{\text{ghost}}}\left[-\frac{1}{2\pi}\,\bar{\partial}c(\sigma)\,b(\sigma') + \delta(\sigma-\sigma')\right]$$

which tells us that

$$\bar{\partial}c(\sigma)\,b(\sigma') = 2\pi\,\delta(\sigma-\sigma')$$

Similarly, looking at $\delta/\delta c(\sigma)$ gives

$$\bar{\partial}b(\sigma)\,c(\sigma') = 2\pi\,\delta(\sigma-\sigma')$$

We can integrate both of these equations using our favorite formula $\bar{\partial}(1/z) = 2\pi\delta(z,\bar{z})$. We learn that the OPEs between fields are given by

$$b(z)\,c(w) = \frac{1}{z-w} + \ldots$$
$$c(w)\,b(z) = \frac{1}{w-z} + \ldots$$

In fact the second equation follows from the first equation and Fermi statistics. The OPEs of $b(z)\,b(w)$ and $c(z)\,c(w)$ have no singular parts. They vanish as $z \to w$.

Finally, we need the stress tensor of the theory. After normal ordering, it is given by

$$T(z) = 2 : \partial c(z)\,b(z) : + : c(z)\partial b(z) :$$

We will shortly see that with this choice, b and c carry appropriate weights for tensor fields which are neutral under Weyl rescaling.

Primary Fields

We will now show that both b and c are primary fields, with weights $h = 2$ and $h = -1$ respectively. Let's start by looking at c. The OPE with the stress tensor is

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

confirming that c has weight -1. When taking the OPE with b, we need to be a little more careful with minus signs. We get

$$\begin{aligned} T(z)\,b(w) &= 2 : \partial c(z)\,b(z) : b(w) + : c(z)\,\partial b(z) : b(w) \\ &= -2b(z)\left(\frac{-1}{(z-w)^2}\right) - \frac{\partial b(z)}{z-w} = \frac{2b(w)}{(z-w)^2} + \frac{\partial b(w)}{z-w} + \ldots \end{aligned}$$

showing that b has weight 2. As we've pointed out a number of times, conformal = diffeo + Weyl. We mentioned earlier that the fields b and c are neutral under Weyl transformations. This is reflected in their weights, which are due solely to diffeomorphisms as dictated by their index structure: b_{zz} and c^z.

The Central Charge

Finally, we can compute the TT OPE to determine the central charge of the bc ghost system.

$$\begin{aligned} T(z)\,T(w) = {} & 4 : \partial c(z)b(z) :: \partial c(w)b(w) : + 2 : \partial c(z)b(z) :: c(w)\partial b(w) : \\ & + 2 : c(z)\partial b(z) :: \partial c(w)b(w) : + : c(z)\partial b(z) :: c(w)\partial b(w) : \end{aligned}$$

For each of these terms, making two contractions gives a $(z-w)^{-4}$ contribution to the OPE. There are also two ways to make a single contraction. These give $(z-w)^{-1}$ or $(z-w)^{-2}$ or $(z-w)^{-3}$ contributions depending on what the derivatives hit. The end result is

$$\begin{aligned} T(z)\,T(w) = & \frac{-4}{(z-w)^4} + \frac{4:\partial c(z)b(w):}{(z-w)^2} - \frac{4:b(z)\partial c(w):}{(z-w)^2} \\ & - \frac{4}{(z-w)^4} + \frac{2:\partial c(z)\,\partial b(w):}{z-w} - \frac{4:b(z)c(w):}{(z-w)^3} \\ & - \frac{4}{(z-w)^4} - \frac{4:c(z)b(w):}{(z-w)^3} + \frac{2:\partial b(z)\partial c(w):}{z-w} \\ & - \frac{1}{(z-w)^4} - \frac{:c(z)\partial b(w):}{(z-w)^2} + \frac{\partial b(z)c(w):}{(z-w)^2} + \ldots \end{aligned}$$

After some Taylor expansions to turn $f(z)$ functions into $f(w)$ functions, together with a little collecting of terms, this can be written as,

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

The first thing to notice is that it indeed has the form expected of TT OPE. The second, and most important, thing to notice is the central charge of the bc ghost system: it is

$$c = -26$$

5.3 The Critical "Dimension" of String Theory

Let's put the pieces together. We've learnt that gauge fixing the diffeomorphisms and Weyl gauge symmetries results in the introduction of ghosts which contribute central charge $c = -26$. We've also learnt that the Weyl symmetry is anomalous unless $c = 0$. Since the Weyl symmetry is a gauge symmetry, it's crucial that we keep it. We're forced to add exactly the right degrees of freedom to the string to cancel the contribution from the ghosts.

The simplest possibility is to add D free scalar fields. Each of these contributes $c = 1$ to the central charge, so the whole procedure is only consistent if we pick

$$D = 26$$

This agrees with the result we found in Chapter 2: it is the critical dimension of string theory.

However, there's no reason that we have to work with free scalar fields. The consistency requirement is merely that the degrees of freedom of the string are described by a CFT with $c = 26$. Any CFT will do. Each such CFT describes a different background in which a string can propagate. If you like, the space of CFTs with $c = 26$ can be thought of as the space of classical solutions of string theory.

We learn that the "critical dimension" of string theory is something of a misnomer: it is really a "critical central charge". Only for rather special CFTs can this central charge be thought of as a spacetime dimension.

For example, if we wish to describe strings moving in 4d Minkowski space, we can take $D = 4$ free scalars (one of which will be timelike) together with some other $c = 22$ CFT. This CFT may have a geometrical interpretation, or it may be something more abstract. The CFT with $c = 22$ is sometimes called the "internal sector" of the theory. It is what we really mean when we talk about the "extra hidden dimensions of string theory". We'll see some examples of CFTs describing curved spaces in Section 7.

There's one final subtlety: we need to be careful with the transition back to Minkowski space. After all, we want one of the directions of the CFT, X^0, to have the wrong sign kinetic term. One safe way to do this is to keep X^0 as a free scalar field, with the remaining degrees of freedom described by some $c = 25$ CFT. This doesn't seem quite satisfactory though since it doesn't allow for spacetimes which evolve in time — and, of course, these are certainly necessary if we wish to understand early universe cosmology. There are still some technical obstacles to understanding the worldsheet of the string in time-dependent backgrounds. To make progress, and discuss string cosmology, we usually bi-pass this issue by working with the low-energy effective action which we will derive in Section 7.

5.3.1 The Usual Nod to the Superstring

The superstring has another gauge symmetry on the worldsheet: supersymmetry. This gives rise to more ghosts, the so-called $\beta\gamma$ system, which turns out to have central charge +11. Consistency then requires that the degrees of freedom of the string have central charge $c = 26 - 11 = 15$.

However, now the CFTs must themselves be invariant under supersymmetry, which means that bosons come matched with fermions. If we add D bosons, then we also need to add D fermions. A free boson has $c = 1$, while a free fermion has $c = 1/2$. So, the total number of free bosons that we should add is $D(1 + 1/2) = 15$, giving us the critical dimension of the superstring:

$$D = 10$$

5.3.2 An Aside: Non-Critical Strings

Although it's a slight departure from the our main narrative, it's worth pausing to mention what Polyakov actually did in his four page paper. His main focus was not critical strings, with $D = 26$, but rather *non-critical* strings with $D \neq 26$. From the discussion above, we know that these suffer from a Weyl anomaly. But it turns out that there is a way to make sense of the situation.

The starting point is to abandon Weyl invariance from the beginning. We start with D free scalar fields coupled to a dynamical worldsheet metric $g_{\alpha\beta}$. (More generally, we could have any CFT). We still want to keep reparameterization invariance, but now we ignore the constraints of Weyl invariance. Of course, it seems likely that this isn't going to have too much to do with the Nambu-Goto string, but let's proceed anyway. Without Weyl invariance, there is one extra term that it is natural to add to the 2d theory: a worldsheet cosmological constant μ,

$$S_{\rm non-critical} = \frac{1}{4\pi\alpha'}\int d^2\sigma \sqrt{g}\left(g^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X_\mu + \mu\right)$$

Our goal will be to understand how the partition function changes under a Weyl rescaling. There will be two contributions: one from the explicit μ dependence and one from the Weyl anomaly. Consider two metrics related by a Weyl transformation

$$\hat{g}_{\alpha\beta} = e^{2\omega} g_{\alpha\beta}$$

As we vary ω, the partition function $Z[\hat{g}]$ changes as

$$\begin{aligned}
\frac{1}{Z}\frac{\partial Z}{\partial \omega} &= \frac{1}{Z}\int \mathcal{D}X\ e^{-S}\left(-\frac{\partial S}{\partial \hat{g}_{\alpha\beta}}\frac{\partial \hat{g}_{\alpha\beta}}{\partial \omega}\right) \\
&= \frac{1}{Z}\int \mathcal{D}X\ e^{-S}\left(-\frac{1}{2\pi}\sqrt{\hat{g}}\, T^\alpha_{\ \alpha}\right) \\
&= \frac{c}{24\pi}\sqrt{\hat{g}}\,\hat{R} - \frac{1}{2\pi\alpha'}\,\mu e^{2\omega} \\
&= \frac{c}{24\pi}\sqrt{g}(R - 2\nabla^2\omega) - \frac{1}{2\pi\alpha'}\,\mu e^{2\omega}
\end{aligned}$$

where, in the last two lines, we used the Weyl anomaly (4.35) and the relationship between Ricci curvatures (1.29). The central charge appearing in these formulae includes the contribution from the ghosts,

$$c = D - 26$$

We can now just treat this as a differential equation for the partition function Z and solve. This allows us to express the partition function $Z[\hat{g}]$, defined on one worldsheet metric, in terms of $Z[g]$, defined on another. The relationship is,

$$Z[\hat{g}] = Z[g] \exp\left[-\frac{1}{4\pi\alpha'}\int d^2\sigma \,\sqrt{g}\left(2\mu e^{2\omega} - \frac{c\alpha'}{6}\left(g^{\alpha\beta}\,\partial_\alpha\omega\,\partial^\beta\omega + R\omega\right)\right)\right]$$

We see that the scaling mode ω inherits a kinetic term. It now appears as a new dynamical scalar field in the theory. It is often called the Liouville field on account of the exponential potential term multiplying μ. Solving this theory is quite hard[7]. Notice also that our new scalar field ω appears in the final term multiplying the Ricci scalar R. We will describe the significance of this in Section 7.2.1. We'll also see another derivation of this kind of Lagrangian in Section 7.4.4.

5.4 States and Vertex Operators

In Chapter 2 we determined the spectrum of the string in flat space. What is the spectrum for a general string background? The theory consists of the b and c ghosts, together with a $c = 26$ CFT. At first glance, it seems that we have a greatly enlarged Hilbert space since we can act with creation operators from all fields, including the ghosts. However, as you might expect, not all of these states will be physical. After correctly accounting for the gauge symmetry, only some subset survives.

The elegant method to determine the physical Hilbert space in a gauge fixed action with ghosts is known as *BRST quantization*. You will learn about it in the "Advanced Quantum Field Theory" course where you will apply it to Yang-Mills theory. Although a correct construction of the string spectrum employs the BRST method, we won't describe it here for lack of time. A very clear description of the general method and its application to the string can be found in Section 4.2 of Polchinski's book.

Instead, we will make do with a poor man's attempt to determine the spectrum of the string. Our strategy is to simply pretend that the ghosts aren't there and focus on the states created by the fields of the matter CFT (i.e. the X^μ fields if we're talking about flat space). As we'll explain in the next section, if we're only interested in tree-level scattering amplitudes then this will suffice.

To illustrate how to compute the spectrum of the string, let's go back to flat $D = 26$ dimensional Minkowski space and the discussion of covariant quantization in Section

[7] A good review can be found Seiberg's article "*Notes on Quantum Liouville Theory and Quantum Gravity*", Prog. Theor. Phys. Supl. 102 (1990) 319.

2.1. We found that physical states $|\Psi\rangle$ are subject to the Virasoro constraints (2.6) and (2.7) which read

$$
\begin{aligned}
L_n\,|\Psi\rangle &= 0 \qquad\qquad \text{for } n>0\\
L_0\,|\Psi\rangle &= a\,|\Psi\rangle
\end{aligned}
$$

and similar for $\tilde{L}_n$,

$$
\begin{aligned}
\tilde{L}_n\,|\Psi\rangle &= 0 \qquad\qquad \text{for } n>0\\
\tilde{L}_0\,|\Psi\rangle &= \tilde{a}\,|\Psi\rangle
\end{aligned}
$$

where we have, just briefly, allowed for the possibility of different normal ordering coefficients a and $\tilde{a}$ for the left- and right-moving sectors. But there's a name for states in a conformal field theory obeying these requirements: they are primary states of weight $(a,\tilde{a})$.

So how do we fix the normal ordering ambiguities a and $\tilde{a}$? A simple way is to first replace the states with operator insertions on the worldsheet using the state-operator map: $|\Psi\rangle \to \mathcal{O}$. But we have a further requirement on the operators $\mathcal{O}$: gauge invariance. There are two gauge symmetries: reparameterization invariance and Weyl symmetry. Both restrict the possible states.

Let's start by considering reparameterization invariance. In the last section, we happily placed operators at specific points on the worldsheet. But in a theory with a dynamical metric, this doesn't give rise to a diffeomorphism invariant operator. To make an object that is invariant under reparameterizations of the worldsheet coordinates, we should integrate over the whole worldsheet. Our operator insertions (in conformal gauge) are therefore of the form,

$$V \sim \int d^2z\ \mathcal{O} \tag{5.7}$$

Here the $\sim$ sign reflects the fact that we've dropped an overall normalization constant which we'll return to in the next section.

Integrating over the worldsheet takes care of diffeomorphisms. But what about Weyl symmetries? The measure d^2z has weight $(-1,-1)$ under rescaling. To compensate, the operator $\mathcal{O}$ must have weight $(+1,+1)$. This is how we fix the normal ordering ambiguity: we require $a=\tilde{a}=1$. Note that this agrees with the normal ordering coefficient $a=1$ that we derived in lightcone quantization in Chapter 2.

This, then, is the rather rough derivation of the string spectrum. The physical states are the primary states of the CFT with weight $(+1,+1)$. The operators (5.7) associated to these states are called *vertex operators*.

5.4.1 An Example: Closed Strings in Flat Space

Let's use this new language to rederive the spectrum of the closed string in flat space. We start with the ground state of the string, which was previously identified as a tachyon. As we saw in Section 4, the vacuum of a CFT is associated to the identity operator. But we also have the zero modes. We can give the string momentum p^μ by acting with the operator $e^{ip\cdot X}$. The vertex operator associated to the ground state of the string is therefore

$$V_{\text{tachyon}} \sim \int d^2z \ :e^{ip\cdot X}: \tag{5.8}$$

In Section 4.3.3, we showed that the operator $e^{ip\cdot X}$ is primary with weight $h = \tilde{h} = \alpha' p^2/4$. But Weyl invariance requires that the operator has weight $(+1,+1)$. This is only true if the mass of the state is

$$M^2 \equiv -p^2 = -\frac{4}{\alpha'}$$

This is precisely the mass of the tachyon that we saw in Section 2.

Let's now look at the first excited states. In covariant quantization, these are of the form $\zeta_{\mu\nu}\, \alpha^\mu_{-1}\, \tilde{\alpha}^\nu_{-1}\, |0;p\rangle$, where $\zeta_{\mu\nu}$ is a constant tensor that determines the type of state, together with its polarization. (Recall: traceless symmetric $\zeta_{\mu\nu}$ corresponds to the graviton, anti-symmetric $\zeta_{\mu\nu}$ corresponds to the $B_{\mu\nu}$ field and the trace of $\zeta_{\mu\nu}$ corresponds to the scalar known as the dilaton). From (4.56), the vertex operator associated to this state is,

$$V_{\text{excited}} \sim \int d^2z \ :e^{ip\cdot X}\, \partial X^\mu \bar{\partial} X^\nu :\ \zeta_{\mu\nu} \tag{5.9}$$

where ∂X^μ gives us a α^μ_{-1} excitation, while $\bar{\partial} X^\mu$ gives a $\tilde{\alpha}^\mu_{-1}$ excitation. It's easy to check that the weight of this operator is $h = \tilde{h} = 1 + \alpha' p^2/4$. Weyl invariance therefore requires that

$$p^2 = 0$$

confirming that the first excited states of the string are indeed massless. However, we still need to check that the operator in (5.9) is actually primary. We know that ∂X is primary and we know that $e^{ip\cdot X}$ is primary, but now we want to consider them both sitting together inside the normal ordering. This means that there are extra terms in the Wick contraction which give rise to $1/(z-w)^3$ terms in the OPE, potentially ruining the primacy of our operator. One such term arises from a double contraction, one of which includes the $e^{ip\cdot X}$ operator. This gives rise to an offending term proportional to $p^\mu \zeta_{\mu\nu}$. The same kind of contraction with $\bar{T}$ gives rise to a term proportional to $p^\nu \zeta_{\nu\mu}$. In order for these terms to vanish, the polarization tensor must satisfy

$$p^\mu \zeta_{\mu\nu} = p^\nu \zeta_{\mu\nu} = 0$$

which is precisely the transverse polarization condition expected for a massless particle.

5.4.2 An Example: Open Strings in Flat Space

As explained in Section 4.7, vertex operators for the open-string are inserted on the boundary $\partial\mathcal{M}$ of the worldsheet. We still need to ensure that these operators are diffeomorphism invariant which is achieved by integrating over $\partial\mathcal{M}$. The vertex operator for the open string tachyon is

$$V_{\text{tachyon}} \sim \int_{\partial\mathcal{M}} ds \ :e^{ip\cdot X}:$$

We need to figure out the dimension of the boundary operator $:e^{ip\cdot X}:$. It's not the same as for the closed string. The reason is due to presence of the image charge in the propagator (4.57) for a free scalar field on a space with boundary. This propagator appears in the Wick contractions in the OPEs and affects the weights. Let's see why this is the case. Firstly, we look at a single scalar field X,

$$\begin{aligned}
\partial X(z) \ :e^{ipX(w,\bar{w})}: &= \sum_{n=1}^{\infty} \frac{(ip)^n}{(n-1)!} :X(w,\bar{w})^{n-1}: \left(-\frac{\alpha'}{2}\frac{1}{z-w} - \frac{\alpha'}{2}\frac{1}{z-\bar{w}}\right) + \ldots \\
&= -\frac{i\alpha' p}{2} \ :e^{ipX(w,\bar{w})}: \left(\frac{1}{z-w} + \frac{1}{z-\bar{w}}\right) + \ldots
\end{aligned}$$

With this result, we can now compute the OPE with T,

$$T(z) \ :e^{ipX(w,\bar{w})}: = \frac{\alpha' p^2}{4} \ :e^{ipX}: \left(\frac{1}{z-w} + \frac{1}{z-\bar{w}}\right)^2 + \ldots$$

When the operator $:e^{ipX(w,\bar{w})}:$ is placed on the boundary $w = \bar{w}$, this becomes

$$T(z) :e^{ipX(w,\bar{w})}: = \frac{\alpha' p^2 \ :e^{ipX(w,\bar{w})}:}{(z-w)^2} + \ldots$$

This tells us that the boundary operator $:e^{ip\cdot X}:$ is indeed primary, with weight $\alpha' p^2$.

For the open string, Weyl invariance requires that operators have weight +1 in order to cancel the scaling dimension of −1 coming from the boundary integral $\int ds$. So the mass of the open string ground state is

$$M^2 \equiv -p^2 = -\frac{1}{\alpha'}$$

in agreement with the mass of the open string tachyon computed in Section 3.

The vertex operator for the photon is

$$V_{\text{photon}} \sim \int_{\partial \mathcal{M}} ds\ \zeta_a\ :\partial X^a\, e^{ip\cdot X}: \tag{5.10}$$

where the index $a = 0, \ldots, p$ now runs only over those directions with Neumann boundary conditions that lie parallel to the brane worldvolume. The requirement that this is a primary operator gives $p^a \zeta_a = 0$, while Weyl invariance tells us that $p^2 = 0$. This is the expected behaviour for the momentum and polarization of a photon.

5.4.3 More General CFTs

Let's now consider a string propagating in four-dimensional Minkowski space $\mathcal{M}_4$, together with some internal CFT with $c = 22$. Then any primary operator of the internal CFT with weight (h, h) can be assigned momentum p^μ, for $\mu = 0, 1, 2, 3$ by dressing the operator with $e^{ip\cdot X}$. In order to get a primary operator of weight $(+1, +1)$ as required, we must have

$$\frac{\alpha' p^2}{4} = 1 - h$$

We see that the mass spectrum of closed string states is given by

$$M^2 = \frac{4}{\alpha'}(h - 1)$$

where h runs over the spectrum of primary operators of the internal CFT. Some comments:

- Relevant operators in the internal CFT have $h < 1$ and give rise to tachyons in the spectrum. Marginal operators, with $h = 1$, give massless particles. And irrelevant operators result in massive states.
- Notice that requiring the vertex operators to be Weyl invariant determines the mass formula for the state. We say that the vertex operators are "on-shell", in the same sense that external legs of Feynman diagrams are on-shell. We will have more to say about this in the next section.

6. String Interactions

So far, despite considerable effort, we've only discussed the free string. We now wish to consider interactions. If we take the analogy with quantum field theory as our guide, then we might be led to think that interactions require us to add various non-linear terms to the action. However, this isn't the case. Any attempt to add extra non-linear terms for the string won't be consistent with our precious gauge symmetries. Instead, rather remarkably, all the information about interacting strings is already contained in the free theory described by the Polyakov action. (Actually, this statement is almost true).

To see that this is at least feasible, try to draw a cartoon picture of two strings interacting. It looks something like the worldsheet shown in the figure. The worldsheet is smooth. In Feynman diagrams in quantum field theory, information about interactions is inserted at vertices, where different lines meet. Here there are no such points. Locally, every part of the diagram looks like a free propagating string. Only globally do we see that the diagram describes interactions.

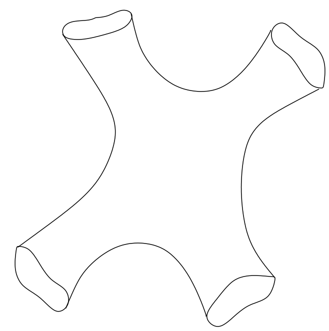


Figure 31:

6.1 What to Compute?

If the information about string interactions is already contained in the Polyakov action, let's go ahead and compute something! But what should we compute? One obvious thing to try is the probability for a particular configuration of strings at an early time to evolve into a new configuration at some later time. For example, we could try to compute the amplitude associated to the diagram above, stipulating fixed curves for the string ends.

No one knows how to do this. Moreover, there are words that we can drape around this failure that suggests this isn't really a sensible thing to compute. I'll now try to explain these words. Let's start by returning to the familiar framework of quantum field theory in a fixed background. There the basic objects that we can compute are correlation functions,

$$\langle \phi(x_1) \dots \phi(x_n) \rangle \tag{6.1}$$

After a Fourier transform, these describe Feynman diagrams in which the external legs carry arbitrary momenta. For this reason, they are referred to as *off-shell*. To get the scattering amplitudes, we simply need to put the external legs on-shell (and perform a few other little tricks captured in the LSZ reduction formula).

The discussion above needs amendment if we turn on gravity. Gravity is a gauge theory and the gauge symmetries are diffeomorphisms. In a gauge theory, only gauge invariant observables make sense. But the correlation function (6.1) is not gauge invariant because its value changes under a diffeomorphism which maps the points x_i to another point. This emphasizes an important fact: there are no local off-shell gauge invariant observables in a theory of gravity.

There is another way to say this. We know, by causality, that space-like separated operators should commute in a quantum field theory. But in gravity the question of whether operators are space-like separated becomes a dynamical issue and the causal structure can fluctuate due to quantum effects. This provides another reason why we are unable to define local gauge invariant observables in any theory of quantum gravity.

Let's now return to string theory. Computing the evolution of string configurations for a finite time is analogous to computing off-shell correlation functions in QFT. But string theory is a theory of gravity so such things probably don't make sense. For this reason, we retreat from attempting to compute correlation functions, back to the S-matrix.

The String S-Matrix

The object that we can compute in string theory is the S-matrix. This is obtained by taking the points in the correlation function to infinity: $x_i \to \infty$. This is acceptable because, just like in the case of QED, the redundancy of the system consists of those gauge transformations which die off asymptotically. Said another way, points on the boundary don't fluctuate in quantum gravity. (Such fluctuations would be over an infinite volume of space and are suppressed due to their infinite action).

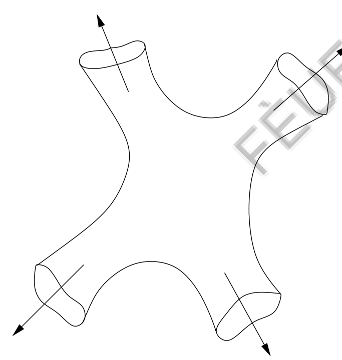


Figure 32:

So what we're really going to calculate is a diagram of the type shown in the figure, where all external legs are taken to infinity. Each of these legs can be placed in a different state of the free string and assigned some spacetime momentum p_i. The resulting expression is the string *S-matrix*.

Using the state-operator map, we know that each of these states at infinity is equivalent to the insertion of an appropriate vertex operator on the worldsheet. Therefore, to compute this S-matrix element we use a conformal transformation to bring each of these infinite legs to a finite distance. The end result is a worldsheet with the topology of the sphere, dotted with vertex operators where the legs used to be. However, we already saw in the previous section that the constraint of Weyl invariance meant that vertex operators are necessarily on-shell. Technically, this is the reason that we can only compute on-shell correlation functions in string theory.

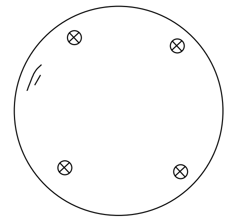
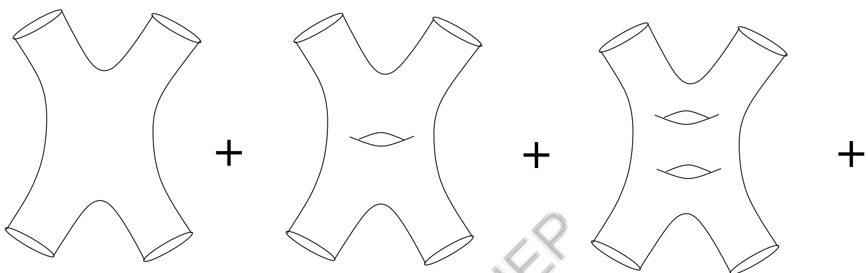


Figure 33:

6.1.1 Summing Over Topologies

The Polyakov path integral instructs us to sum over all metrics. But what about worldsheets of different topologies? In fact, we should also sum over these. It is this sum that gives the perturbative expansion of string theory. The scattering of two strings receives contributions from worldsheets of the form



(6.2)

The only thing that we need to know is how to weight these different worldsheets. Thankfully, there is a very natural coupling on the string that we have yet to consider and this will do the job. We augment the Polyakov action by

$$S_{\text{string}} = S_{\text{Poly}} + \lambda\chi \tag{6.3}$$

Here λ is simply a real number, while χ is given by an integral over the (Euclidean) worldsheet

$$\chi = \frac{1}{4\pi}\int d^2\sigma\ \sqrt{g}R \tag{6.4}$$

where R is the Ricci scalar of the worldsheet metric. This looks like the Einstein-Hilbert term for gravity on the worldsheet. It is simple to check that it is invariant under reparameterizations and Weyl transformations.

In four-dimensions, the Einstein-Hilbert term makes gravity dynamical. But life is very different in 2d. Indeed, we've already seen that all the components of the metric can be gauged away so there are no propagating degrees of freedom associated to $g_{\alpha\beta}$. So, in two-dimensions, the term (6.4) doesn't make gravity dynamical: in fact, classically, it doesn't do anything at all!

The reason for this is that χ is a topological invariant. This means that it doesn't actually depend on the metric $g_{\alpha\beta}$ at all – it depends only on the topology of the worldsheet. (More precisely, χ only depends on those global properties of the metric which themselves depend on the topology of the worldsheet). This is the content of the Gauss-Bonnet theorem: the integral of the Ricci scalar R over the worldsheet gives an integer, χ, known as the Euler number of the worldsheet. For a worldsheet without boundary (i.e. for the closed string) χ counts the number of handles h on the worldsheet. It is given by,

$$\chi = 2 - 2h = 2(1-g) \tag{6.5}$$

where g is called the *genus* of the surface. The simplest examples are shown in the figure. The sphere has $g = 0$ and $\chi = 2$; the torus has $g = 1$ and $\chi = 0$. For higher $g > 1$, the Euler character χ is negative.

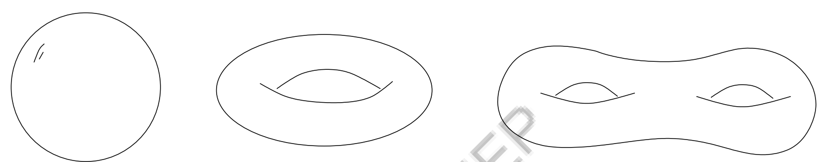


Figure 34: Examples of increasingly poorly drawn Riemann surfaces with $\chi = 2, 0$ and -2.

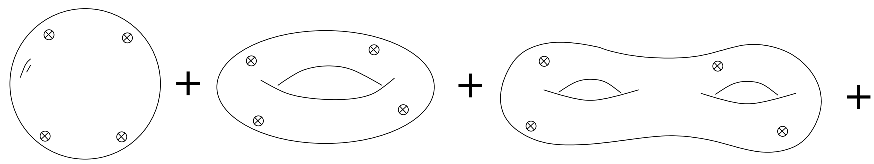
Now we see that the number λ — or, more precisely, e^λ — plays the role of the string coupling. The integral over worldsheets is weighted by,

$$\sum_{\substack{\text{topologies} \\ \text{metrics}}} e^{-S_{\text{string}}} \sim \sum_{\text{topologies}} e^{-2\lambda(1-g)} \int \mathcal{D}X\mathcal{D}g \; e^{-S_{\text{Poly}}}$$

For $e^\lambda \ll 1$, we have a good perturbative expansion in which we sum over all topologies. (In fact, it is an asymptotic expansion, just as in quantum field theory). It is standard to define the string coupling constant as

$$g_s = e^\lambda$$

After a conformal map, tree-level scattering corresponds to a worldsheet with the topology of a sphere: the amplitudes are proportional to $1/g_s^2$. One-loop scattering corresponds to toroidal worldsheets and, with our normalization, have no power of g_s. (Although, obviously, these are suppressed by g_s^2 relative to tree-level processes). The end result is that the sum over worldsheets in (6.2) becomes a sum over Riemann surfaces of increasing genus, with vertex operators inserted for the initial and final states,



The Riemann surface of genus g is weighted by

$$(g_s^2)^{g-1}$$

While it may look like we've introduced a new parameter g_s into the theory and added the coupling (6.3) by hand, we will later see why this coupling is a necessary part of the theory and provide an interpretation for g_s.

Scattering Amplitudes

We now have all the information that we need to explain how to compute string scattering amplitudes. Suppose that we want to compute the S-matrix for m states: we will label them as Λ_i and assign them spacetime momenta p_i. Each has a corresponding vertex operator $V_{\Lambda_i}(p_i)$. The S-matrix element is then computed by evaluating the correlation function in the 2d conformal field theory, with insertions of the vertex operators.

$$\mathcal{A}^{(m)}(\Lambda_i, p_i) = \sum_{\text{topologies}} g_s^{-\chi} \frac{1}{\text{Vol}} \int \mathcal{D}X\mathcal{D}g \; e^{-S_{\text{Poly}}} \prod_{i=1}^{m} V_{\Lambda_i}(p_i)$$

This is a rather peculiar equation. We are interpreting the correlation functions of a two-dimensional theory as the S-matrix for a theory in $D = 26$ dimensions!

To properly compute the correlation function, we should introduce the b and c ghosts that we saw in the last chapter and treat them carefully. However, if we're only interested in tree-level amplitudes, then we can proceed naively and ignore the ghosts. The reason can be seen in the ghost action (5.5) where we see that the ghosts couple only to the worldsheet metric, not to the other worldsheet fields. This means that if our gauge fixing procedure fixes the worldsheet metric completely — which it does for worldsheets with the topology of a sphere — then we can forget about the ghosts. (At least, we can forget about them as soon as we've made sure that the Weyl anomaly cancels). However, as we'll explain in 6.4, for higher genus worldsheets, the gauge fixing does not fix the metric completely and there are residual dynamical modes of the metric, known as moduli, which couple the ghosts and matter fields. This is analogous to the statement in field theory that we only need to worry about ghosts running in loops.

6.2 Closed String Amplitudes at Tree Level

The tree-level scattering amplitude is given by the correlation function of the 2d theory, evaluated on the sphere,

$$\mathcal{A}^{(m)} = \frac{1}{g_s^2}\frac{1}{\text{Vol}}\int \mathcal{D}X\mathcal{D}g\; e^{-S_{\text{Poly}}}\prod_{i=1}^{m} V_{\Lambda_i}(p_i)$$

where $V_{\Lambda_i}(p_i)$ are the vertex operators associated to the states.

We want to integrate over all metrics on the sphere. At first glance that sounds rather daunting but, of course, we have the gauge symmetries of diffeomorphisms and Weyl transformations at our disposal. Any metric on the sphere is conformally equivalent to the flat metric on the plane. For example, the round metric on the sphere of radius R can be written as

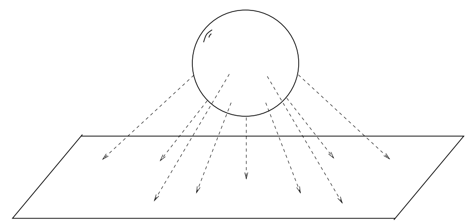


Figure 35:

$$ds^2 = \frac{4R^2}{(1+|z|^2)^2}\, dz d\bar{z}$$

which is manifestly conformally equivalent to the plane, supplemented by the point at infinity. The conformal map from the sphere to the plane is the stereographic projection depicted in the diagram. The south pole of the sphere is mapped to the origin; the north pole is mapped to the point at infinity. Therefore, instead of integrating over all metrics, we may gauge fix diffeomorphisms and Weyl transformations to leave ourselves with the seemingly easier task of computing correlation functions on the plane.

6.2.1 Remnant Gauge Symmetry: SL(2,C)

There's a subtlety. And it's a subtlety that we've seen before: there is a residual gauge symmetry. It is the conformal group, arising from diffeomorphisms which can be undone by Weyl transformations. As we saw in Section 4, there are an infinite number of such conformal transformations. It looks like we have a whole lot of gauge fixing still to do.

However, global issues actually mean that there's less remnant gauge symmetry than you might think. In Section 4, we only looked at infinitesimal conformal transformations, generated by the Virasoro operators L_n, $n \in \mathbf{Z}$. We did not examine whether these transformations are well-defined and invertible over all of space. Let's take a look at this. Recall that the coordinate changes associated to L_n are generated by the vector fields (4.49),

$$l_n = z^{n+1}\partial_z$$

which result in the shift $\delta z = \epsilon z^{n+1}$. This is non-singular at $z = 0$ only for $n \geq -1$. If we restrict to smooth maps, that gets rid of half the transformations right away. But, since we're ultimately interested in the sphere, we now also need to worry about the point at $z = \infty$ which, in stereographic projection, is just the north pole of the sphere. To do this, it's useful to work with the coordinate

$$u = \frac{1}{z}$$

The generators of coordinate transformations for the u coordinate are

$$l_n = z^{n+1}\partial_z = \frac{1}{u^{n+1}}\frac{\partial u}{\partial z}\partial_u = -u^{1-n}\partial_u$$

which is non-singular at $u = 0$ only for $n \leq 1$.

Combining these two results, the only generators of the conformal group that are non-singular over the whole Riemann sphere are l_{-1}, l_0 and l_1 which act infinitesimally as

$$\begin{aligned} l_{-1}: &\quad z \rightarrow z + \epsilon \\ l_0: &\quad z \rightarrow (1+\epsilon)z \\ l_1: &\quad z \rightarrow (1+\epsilon z)z \end{aligned}$$

The global version of these transformations is

$$\begin{aligned} l_{-1}: &\quad z \rightarrow z + \alpha \\ l_0: &\quad z \rightarrow \lambda z \\ l_1: &\quad z \rightarrow \frac{z}{1-\beta z} \end{aligned}$$

which can be combined to give the general transformation

$$z \rightarrow \frac{az+b}{cz+d} \tag{6.6}$$

with a, b, c and $d \in \mathbf{C}$. We have four complex parameters, but we've only got three transformations. What happened? Well, one transformation is fake because an overall scaling of the parameters doesn't change z. By such a rescaling, we can always insist that the parameters obey

$$ad - bc = 1$$

The transformations (6.6) subject to this constraint have the group structure $SL(2;\mathbf{C})$, which is the group of 2×2 complex matrices with unit determinant. In fact, since the transformation is blind to a flip in sign of all the parameters, the actual group of global conformal transformations is $SL(2;\mathbf{C})/\mathbf{Z}_2$, which is sometimes written as $PSL(2;\mathbf{C})$. (This $\mathbf{Z}_2$ subtlety won't be important for us in what follows).

The remnant global transformations on the sphere are known as *conformal Killing vectors* and the group $SL(2;\mathbf{C})/\mathbf{Z}_2$ is the *conformal Killing group*. This group allows us to take any three points on the plane and move them to three other points of our choosing. We will shortly make use of this fact to gauge fix, but for now we leave the $SL(2;\mathbf{C})$ symmetry intact.

6.2.2 The Virasoro-Shapiro Amplitude

We will now compute the S-matrix for closed string tachyons. You might think that this is the least interesting thing to compute: after all, we're ultimately interested in the superstring which doesn't have tachyons. This is true, but it turns out that tachyon scattering is much simpler than everything else, mainly because we don't have a plethora of extra indices on the states to worry about. Moreover, the lessons that we will learn from tachyon scattering hold for the scattering of other states as well.

The m-point tachyon scattering amplitude is given by the flat space correlation function

$$\mathcal{A}^{(m)}(p_1,\ldots,p_m) = \frac{1}{g_s^2}\frac{1}{\mathrm{Vol}(SL(2;\mathbf{C}))}\int \mathcal{D}X\; e^{-S_{\rm Poly}}\;\prod_{i=1}^{m} V(p_i)$$

where the tachyon vertex operator is given by,

$$V(p_i) = g_s\int d^2z\; e^{ip_i\cdot X} \equiv g_s\int d^2z\;\hat{V}(z,p_i) \tag{6.7}$$

Note that, in contrast to (5.8), we've added an appropriate normalization factor to the vertex operator. Heuristically, this reflects the fact that the operator is associated to the addition of a closed string mode. A rigorous derivation of this normalization can be found in Polchinski.

The amplitude can therefore be written as,

$$\mathcal{A}^{(m)}(p_1,\dots,p_m)=\frac{g_s^{m-2}}{\mathrm{Vol}(SL(2;\mathbf{C}))}\int\prod_{i=1}^{m}d^2z_i\ \langle\hat{V}(z_1,p_1)\dots\hat{V}(z_m,p_m)\rangle$$

where the expectation value $\langle\dots\rangle$ is computed using the gauge fixed Polyakov action. But the gauge fixed Polyakov action is simply a free theory and our correlation function is something eminently computable: a Gaussian integral,

$$\langle\hat{V}(z_1,p_1)\dots\hat{V}(z_m,p_m)\rangle=\int\mathcal{D}X\exp\left(-\frac{1}{2\pi\alpha'}\int d^2z\ \partial X\cdot\bar{\partial}X\right)\ \exp\left(i\sum_{i=1}^{m}p_i\cdot X(z_i,\bar{z}_i)\right)$$

The normalization in front of the Polyakov action is now $1/2\pi\alpha'$ instead of $1/4\pi\alpha'$ because we're working with complex coordinates and we need to remember that $\partial_\alpha\partial^\alpha=4\partial\bar{\partial}$ and $d^2z=2d^2\sigma$.

The Gaussian Integral

We certainly know how to compute Gaussian integrals. Let's go slow. Consider the following general integral,

$$\int\mathcal{D}X\ \exp\left(\int d^2z\ \frac{1}{2\pi\alpha'}X\cdot\partial\bar{\partial}X+iJ\cdot X\right)\sim\exp\left(\frac{\pi\alpha'}{2}\int d^2zd^2z'\ J(z,\bar{z})\,\frac{1}{\partial\bar{\partial}}\,J(z',\bar{z}')\right)$$

Here the $\sim$ symbol reflects the fact that we've dropped a whole lot of irrelevant normalization terms, including $\det^{-1/2}(-\partial\bar{\partial})$. The inverse operator $1/\partial\bar{\partial}$ on the right-hand-side of this equation is shorthand for the propagator $G(z,z')$ which solves

$$\partial\bar{\partial}G(z,\bar{z};z',\bar{z}')=\delta(z-z',\bar{z}-\bar{z}')$$

As we've seen several times before, in two dimensions this propagator is given by

$$G(z,\bar{z};z',\bar{z}')=\frac{1}{2\pi}\ln|z-z'|^2$$

Back to the Scattering Amplitude

Comparing our scattering amplitude with this general expression, we need to take the source J to be

$$J(z,\bar{z})=\sum_{i=1}^{m}p_i\ \delta(z-z_i,\bar{z}-\bar{z}_i)$$

Inserting this into the Gaussian integral gives us an expression for the amplitude

$$\mathcal{A}^{(m)} \sim \frac{g_s^{m-2}}{\mathrm{Vol}(SL(2;\mathbf{C}))} \int \prod_{i=1}^{m} d^2 z_i \;\exp\left(\frac{\alpha'}{2}\sum_{j,l} p_j \cdot p_l \;\ln|z_j - z_l|\right)$$

The terms with $j = l$ seem to be problematic. In fact, they should just be left out. This follows from correctly implementing normal ordering and leaves us with

$$\mathcal{A}^{(m)} \sim \frac{g_s^{m-2}}{\mathrm{Vol}(SL(2;\mathbf{C}))} \int \prod_{i=1}^{m} d^2 z_i \prod_{j<l} |z_j - z_l|^{\alpha' p_j \cdot p_l} \tag{6.8}$$

Actually, there's something that we missed. (Isn't there always!). We certainly expect scattering in flat space to obey momentum conservation, so there should be a $\delta^{(26)}(\sum_{i=1}^{m} p_i)$ in the amplitude. But where is it? We missed it because we were a little too quick in computing the Gaussian integral. The operator $\partial\bar{\partial}$ annihilates the zero mode, x^μ, in the mode expansion. This means that its inverse, $1/\partial\bar{\partial}$, is not well-defined. But it's easy to deal with this by treating the zero mode separately. The derivatives ∂^2 don't see x^μ, but the source J does. Integrating over the zero mode in the path integral gives us our delta function

$$\int dx \;\exp(i\sum_{i=1}^{m} p_i \cdot x) \sim \delta^{26}(\sum_{i=1}^{m} p_i)$$

So, our final result for the amplitude is

$$\mathcal{A}^{(m)} \sim \frac{g_s^{m-2}}{\mathrm{Vol}(SL(2;\mathbf{C}))}\, \delta^{26}(\sum_i p_i) \int \prod_{i=1}^{m} d^2 z_i \prod_{j<l} |z_j - z_l|^{\alpha' p_j \cdot p_l} \tag{6.9}$$

The Four-Point Amplitude

We will compute only the four-point amplitude for two-to-two scattering of tachyons. The $\mathrm{Vol}(SL(2;\mathbf{C}))$ factor is there to remind us that we still have a remnant gauge symmetry floating around. Let's now fix this. As we mentioned before, it provides enough freedom for us to take any three points on the plane and move them to any other three points. We will make use of this to set

$$z_1 = \infty \quad , \quad z_2 = 0 \quad , \quad z_3 = z \quad , \quad z_4 = 1$$

Inserting this into the amplitude (6.9), we find ourselves with just a single integral to evaluate,

$$\mathcal{A}^{(4)} \sim g_s^2\; \delta^{26}(\sum_i p_i) \int d^2 z\; |z|^{\alpha' p_2 \cdot p_3}\, |1-z|^{\alpha' p_3 \cdot p_4} \tag{6.10}$$

(There is also an overall factor of $|z_1|^4$, but this just gets absorbed into an overall normalization constant). We still need to do the integral. It can be evaluated exactly in terms of gamma functions. We relegate the proof to Appendix 6.5, where we show that

$$\int d^2z \; |z|^{2a-2}|1-z|^{2b-2} = \frac{2\pi\Gamma(a)\Gamma(b)\Gamma(c)}{\Gamma(1-a)\Gamma(1-b)\Gamma(1-c)} \tag{6.11}$$

where $a+b+c=1$.

Four-point scattering amplitudes are typically expressed in terms of Mandelstam variables. We choose p_1 and p_2 to be incoming momenta and p_3 and p_4 to be outgoing momenta, as shown in the figure. We then define

$$s = -(p_1+p_2)^2 \quad , \quad t = -(p_1+p_3)^2 \quad , \quad u = -(p_1+p_4)^2$$

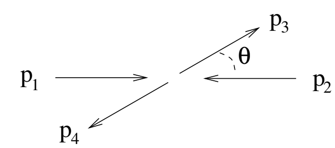


Figure 36:

These obey

$$s+t+u = -\sum_i p_i^2 = \sum_i M_i^2 = -\frac{16}{\alpha'}$$

where, in the last equality, we've inserted the value of the tachyon mass (2.27). Writing the scattering amplitude (6.10) in terms of Mandelstam variables, we have our final answer

$$\mathcal{A}^{(4)} \sim g_s^2\; \delta^{26}(\sum_i p_i)\; \frac{\Gamma(-1-\alpha's/4)\Gamma(-1-\alpha't/4)\Gamma(-1-\alpha'u/4)}{\Gamma(2+\alpha's/4)\Gamma(2+\alpha't/4)\Gamma(2+\alpha'u/4)} \tag{6.12}$$

This is the *Virasoro-Shapiro amplitude* governing tachyon scattering in the closed bosonic string.

Remarkably, the Virasoro-Shapiro amplitude was almost the first equation of string theory! (That honour actually goes to the Veneziano amplitude which is the analogous expression for open string tachyons and will be derived in Section 6.3.1). These amplitudes were written down long before people knew that they had anything to do with strings: they simply exhibited some interesting and surprising properties. It took several years of work to realise that they actually describe the scattering of strings. We will now start to tease apart the Virasoro-Shapiro amplitude to see some of the properties that got people hooked many years ago.

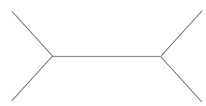
6.2.3 Lessons to Learn

So what's the physics lying behind the scattering amplitude (6.12)? Obviously it is symmetric in s, t and u. That is already surprising and we'll return to it shortly. But we'll start by fixing t and looking at the properties of the amplitude as we vary s.

The first thing to notice is that $\mathcal{A}^{(4)}$ has poles. Lots of poles. They come from the factor of $\Gamma(-1-\alpha' s/4)$ in the numerator. The first of these poles appears when

$$-1-\frac{\alpha' s}{4}=0 \quad \Rightarrow \quad s=-\frac{4}{\alpha'}$$

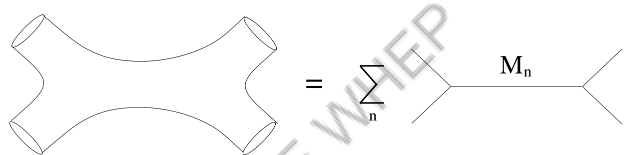
But that's the mass of the tachyon! It means that, for s close to $-4/\alpha'$, the amplitude has the form of a familiar scattering amplitude in quantum field theory with a cubic vertex,



$$\sim \frac{1}{s-M^2}$$

where M is the mass of the exchanged particle, in this case the tachyon.

Other poles in the amplitude occur at $s=4(n-1)/\alpha'$ with $n \in \mathbf{Z}^+$. This is precisely the mass formula for the higher states of the closed string. What we're learning is that the string amplitude is summing up an infinite number of tree-level field theory diagrams,



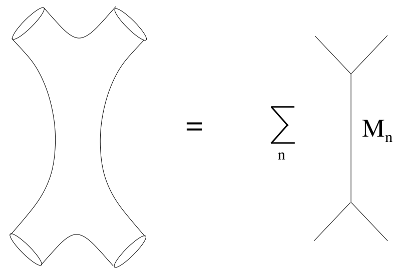
where the exchanged particles are all the different states of the free string.

In fact, there's more information about the spectrum of states hidden within these amplitudes. We can look at the residues of the poles at $s=4(n-1)/\alpha'$, for $n=0,1,\ldots$. These residues are rather complicated functions of t, but the highest power of momentum that appears for each pole is

$$\mathcal{A}^{(4)} \sim \sum_{n=0}^{\infty} \frac{t^{2n}}{s-M_n^2} \tag{6.13}$$

The power of the momentum is telling us the highest spin of the particle states at level n. To see why this is, consider a field corresponding to a spin J particle. It has a whole bunch of Lorentz indices, $\chi_{\mu_1\ldots\mu_J}$. In a cubic interaction, each of these must be soaked up by derivatives. So we have J derivatives at each vertex, contributing powers of $(\text{momentum})^{2J}$ to the numerator of the Feynman diagram. Comparing with the string scattering amplitude, we see that the highest spin particle at level n has $J=2n$. This is indeed the result that we saw from the canonical quantization of the string in Section 2.

Finally, the amplitude (6.12) has a property that is very different from amplitudes in field theory. Above, we framed our discussion by keeping t fixed and expanding in s. We could just have well done the opposite: fix s and look at poles in t. Now the string amplitude has the interpretation of an infinite number of t-channel scattering amplitudes, one for each state of the string



Usually in field theory, we sum up both s-channel and t-channel scattering amplitudes. Not so in string theory. The sum over an infinite number of s-channel amplitudes can be reinterpreted as an infinite sum of t-channel amplitudes. We don't include both: that would be overcounting. (Similar statements hold for u). The fact that the same amplitude can be written as a sum over s-channel poles *or* a sum over t-channel poles is sometimes referred to as "duality". (A much overused word). In the early days, before it was known that string theory was a theory of strings, the subject inherited its name from this duality property of amplitudes: it was called the *dual resonance model*.

High Energy Scattering

Let's use this amplitude to see what happens when we collide strings at high energies. There are different regimes that we could look at. The most illuminating is s, $t \to \infty$, with s/t held fixed. In this limit, all the exchanged momenta become large. It corresponds to high-energy scattering with the angle θ between incoming and outgoing particles kept fixed. To see this consider, for example, massless particles (our amplitude is really for tachyons, but the same considerations hold). We take the incoming and outgoing momenta to be

$$p_1 = \frac{\sqrt{s}}{2}(1, 1, 0, \ldots) \quad , \quad p_2 = \frac{\sqrt{s}}{2}(1, -1, 0, \ldots)$$
$$p_3 = \frac{\sqrt{s}}{2}(1, \cos\theta, \sin\theta, \ldots) \quad , \quad p_4 = \frac{\sqrt{s}}{2}(1, -\cos\theta, -\sin\theta, \ldots)$$

Then we see explicitly that $s \to \infty$ and $t \to \infty$ with the ratio s/t fixed also keeps the scattering angle θ fixed.

We can evaluate the scattering amplitude $\mathcal{A}^{(4)}$ in this limit by using $\Gamma(x) \sim \exp(x \ln x)$. We send $s \to \infty$ avoiding the poles. (We can achieve this by sending $s \to \infty$ in a slightly imaginary direction. Ultimately this is valid because all the higher string states are actually unstable in the interacting theory which will shift their poles off the real axis once taken into account). It is simple to check that the amplitude drops off exponentially quickly at high energies,

$$\mathcal{A}^{(4)} \sim g_s^2\, \delta^{26}(\sum_i p_i)\, \exp\left(-\frac{\alpha'}{2}(s \ln s + t \ln t + u \ln u)\right) \qquad \text{as } s \to \infty \tag{6.14}$$

The exponential fall-off seen in (6.14) is much faster than the amplitude of any field theory which, at best, fall off with power-law decay at high energies and, at worse, diverge. For example, consider the individual terms (6.13) corresponding to the amplitude for s-channel processes involving the exchange of particles with spin $2n$. We see that the exchange of a spin 2 particle results in a divergence in this limit. This is reflecting something you already know about gravity: the dimensionless coupling is $G_N E^2$ (in four-dimensions) which becomes large for large energies. The exchange of higher spin particles gives rise to even worse divergences. If we were to truncate the infinite sum (6.13) at any finite n, the whole thing would diverge. But infinite sums can do things that finite sums can't and the final behaviour of the amplitude (6.14) is much softer than any of the individual terms. The infinite number of particles in string theory conspire to render finite any divergence arising from an individual particle species.

Phrased in terms of the s-channel exchange of particles, the high-energy behaviour of string theory seems somewhat miraculous. But there is another viewpoint where it's all very obvious. The power-law behaviour of scattering amplitudes is characteristic of point-like charges. But, of course, the string isn't a point-like object. It is extended and fuzzy at length scales comparable to $\sqrt{\alpha'}$. This is the reason the amplitude has such soft high-energy behaviour. Indeed, this idea that smooth extended objects give rise to scattering amplitudes that decay exponentially at high energies is something that you've seen before in non-relativistic quantum mechanics. Consider, for example, the scattering of a particle off a Gaussian potential. In the Born approximation, the differential cross-section is just given by the Fourier transform which is again a Gaussian, now decaying exponentially for large momentum.

It's often said that theories of quantum gravity should have a "minimum length", sometimes taken to be the Planck scale. This is roughly true in string theory, although not in any crude simple manner. Rather, the minimum length reveals itself in different ways depending on which question is being asked. The above discussion highlights one example of this: strings can't probe distance scales shorter than $l_s = \sqrt{\alpha'}$ simply because they are themselves fuzzy at this scale. It turns out that D-branes are much better probes of sub-stringy physics and provide a different view on the short distance structure of spacetime. We will also see another manifestation of the minimal length scale of string theory in Section 8.3.

Graviton Scattering

Although we've derived the result (6.14) for tachyons, all tree-level amplitudes have this soft fall-off at high-energies. Most notably, this includes graviton scattering. As we noted above, this is in sharp contrast to general relativity for which tree-level scattering amplitudes diverge at high-energies. This is the first place to see that UV problems of general relativity might have a good chance of being cured in string theory.

Using the techniques described in this section, one can compute m-point tree-level amplitudes for graviton scattering. If we restrict attention to low-energies (i.e. much smaller than $1/\sqrt{\alpha'}$), one can show that these coincide with the amplitudes derived from the Einstein-Hilbert action in $D = 26$ dimensions

$$S = \frac{1}{2\kappa^2}\int d^{26}X\sqrt{-G}\,\mathcal{R}$$

where $\mathcal{R}$ is the $D = 26$ Ricci scalar (not to be confused with the worldsheet Ricci scalar which we call R). The gravitational coupling, κ^2 is related to Newton's constant in 26 dimensions. It plays no role for pure gravity, but is important when we couple to matter. We'll see shortly that it's given by

$$\kappa^2 \approx g_s^2(\alpha')^{12}$$

We won't explicitly compute graviton scattering amplitudes in this course, partly because they're fairly messy and partly because building up the Einstein-Hilbert action from m-particle scattering is hardly the best way to look at general relativity. Instead, we shall derive the Einstein-Hilbert action in a much better fashion in Section 7.

6.3 Open String Scattering

So far our discussion has been entirely about closed strings. There is a very similar story for open strings. We again compute S-matrix elements. Conformal symmetry now maps tree-level scattering to the disc, with vertex operators inserted on the boundary of the disc.

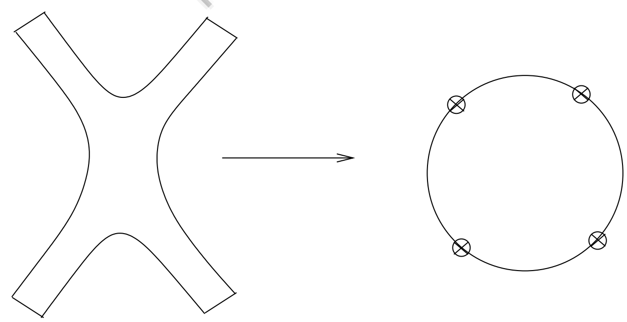


Figure 37: The conformal map from the open string worldsheet to the disc.

For the open string, the string coupling constant that we add to the Polyakov action requires the addition of a boundary term to make it well defined,

$$\chi = \frac{1}{4\pi}\int_{\mathcal{M}} d^2\sigma\ \sqrt{g}R + \frac{1}{2\pi}\int_{\partial\mathcal{M}} ds\, k \tag{6.15}$$

where k is the geodesic curvature of the boundary. To define it, we introduce two unit vectors on the worldsheet: t^α is tangential to the boundary, while n^α is normal and points outward from the boundary. The geodesic curvature is defined as

$$k = -t^\alpha n_\beta \nabla_\alpha t^\beta$$

Boundary terms of the type seen in (6.15) are also needed in general relativity for manifolds with boundaries: in that context, they are referred to as Gibbons-Hawking terms.

The Gauss-Bonnet theorem has an extension to surfaces with boundary. For surfaces with h handles and b boundaries, the Euler character is given by

$$\chi = 2 - 2h - b$$

Some examples are shown in Figure 38. The expansion for open-string scattering consists of adding consecutive boundaries to the worldsheet. The disc is weighted by $1/g_s$; the annulus has no factor of g_s and so on. We see that the open string coupling is related to the closed string coupling by

$$g^2_{\text{open}} = g_s \tag{6.16}$$

One of the key steps in computing closed string scattering amplitudes was the implementation of the conformal Killing group, which was defined as the surviving gauge

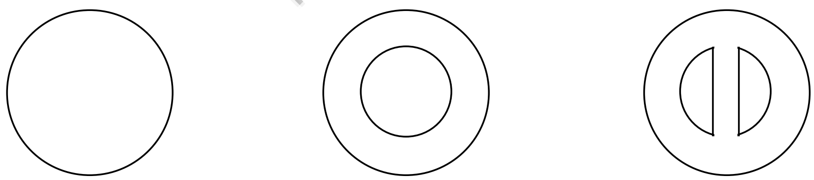


Figure 38: Riemann surfaces with boundary with $\chi = 1, 0$ and -1.

symmetry with a global action on the sphere. For the open string, there is again a residual gauge symmetry. If we think in terms of the upper-half plane, the boundary is Im$z = 0$. The conformal Killing group is composed of transformations

$$z \to \frac{az+b}{cz+d}$$

again with the requirement that $ad - bc = 1$. This time there is one further condition: the boundary Im$z = 0$ must be mapped onto itself. This requires $a, b, c, d \in \mathbf{R}$. The resulting conformal Killing group is $SL(2; \mathbf{R})/\mathbf{Z}_2$.

6.3.1 The Veneziano Amplitude

Since vertex operators now live on the boundary, they have a fixed ordering. In computing a scattering amplitude, we must sum over all orderings. Let's look again at the 4-point amplitude for tachyon scattering. The vertex operator is

$$V(p_i) = \sqrt{g_s} \int dx\; e^{ip_i \cdot X}$$

where the integral $\int dx$ is now over the boundary and $p^2 = 1/\alpha'$ is the on-shell condition for an open-string tachyon. The normalization $\sqrt{g_s}$ is that appropriate for the insertion of an open-string mode, reflecting (6.16).

Going through the same steps as for the closed string, we find that the amplitude is given by

$$\mathcal{A}^{(4)} \sim \frac{g_s}{\mathrm{Vol}(SL(2;\mathbf{R}))}\, \delta^{26}(\sum_i p_i) \int \prod_{i=1}^{4} dx_i \prod_{j<l} |\, x_j - x_l|^{2\alpha' p_j \cdot p_l} \tag{6.17}$$

Note that there's a factor of 2 in the exponent, differing from the closed string expression (6.8). This comes about because the boundary propagator (4.57) has an extra factor of 2 due to the image charge.

We now use the $SL(2;\mathbf{R})$ residual gauge symmetry to fix three points on the boundary. We choose a particular ordering and set $x_1 = 0$, $x_2 = x$, $x_3 = 1$ and $x_4 \to \infty$. The only free insertion point is $x_2 = x$ but, because of the restriction of operator ordering, this must lie in the interval $x \in [0,1]$. The interesting part of the integral is then given by

$$\mathcal{A}^{(4)} \sim g_s \int_0^1 dx\ |x|^{2\alpha' p_1\cdot p_2}\, |1-x|^{2\alpha' p_2 \cdot p_3}$$

This integral is well known: as shown in Appendix 6.5, it is the Euler beta function

$$B(a,b) = \int_0^1 dx\ x^{a-1}(1-x)^{b-1} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

After summing over the different orderings of vertex operators, the end result for the amplitude for open string tachyon scattering is,

$$\mathcal{A}^{(4)} \sim g_s\, [B(-\alpha' s-1, -\alpha' t-1) + B(-\alpha' s-1, -\alpha' u-1) + B(-\alpha' t-1, -\alpha' u-1)]$$

This is the famous *Veneziano Amplitude*, first postulated in 1968 to capture some observed features of the strong interactions. This was before the advent of QCD and before it was realised that the amplitude arises from a string.

The open string scattering amplitude contains the same features that we saw for the closed string. For example, it has poles at

$$s = \frac{n-1}{\alpha'} \qquad n = 0,1,2,\ldots$$

which we recognize as the spectrum of the open string.

6.3.2 The Tension of D-Branes

Recall that we introduced D-branes as surfaces in space on which strings can end. At the time, I promised that we would eventually discover that these D-branes are dynamical objects in their own right. We'll look at this more closely in the next section, but for now we can do a simple computation to determine the tension of D-branes.

The tension T_p of a Dp-brane is defined as the energy per spatial volume. It has dimension $[T_p] = p+1$. The tension is telling us the magnitude of the coupling between the brane and gravity. Or, in our new language, the strength of the interaction between a closed string state and an open string. The simplest such diagram is shown in the figure, with a graviton vertex operator inserted. Although we won't compute this diagram completely, we can figure out its most important property just by looking at it: it has the topology of a disc, so is proportional to $1/g_s$. Adding powers of α' to get the dimension right, the tension of a Dp-brane must scale as

$$T_p \sim \frac{1}{l_s^{p+1}} \frac{1}{g_s} \tag{6.18}$$

where the string length is defined as $l_s = \sqrt{\alpha'}$. The $1/g_s$ scaling of the tension is one of the key characteristic features of a D-brane.

I should confess that there's a lot swept under the carpet in the above discussion, not least the question of the correct normalization of the vertex operators and the difference between the string frame and the Einstein frame (which we will discuss shortly). Nonetheless, the end result (6.18) is correct. For a fuller discussion, see Section 8.7 of Polchinski.

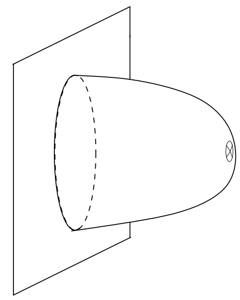


Figure 39:

6.4 One-Loop Amplitudes

We now return to the closed string to discuss one-loop effects. As we saw above, this corresponds to a worldsheet with the topology of a torus. We need to integrate over all metrics on the torus.

For tree-level processes, we used diffeomorphisms and Weyl transformations to map an arbitrary metric on the sphere to the flat metric on the plane. This time, we use these transformations to map an arbitrary metric on the torus to the flat metric on the torus. But there's a new subtlety that arises: not all flat metrics on the torus are equivalent.

6.4.1 The Moduli Space of the Torus

Let's spell out what we mean by this. We can construct a torus by identifying a region in the complex z-plane as shown in the figure. In general, this identification depends on a single complex parameter, $\tau \in \mathbf{C}$.

$$z \equiv z + 2\pi \quad \text{and} \quad z \equiv z + 2\pi\tau$$

Do not confuse τ with the Minkowski worldsheet time: we left that behind way back in Section 3. Everything here is Euclidean worldsheet and τ is just a parameter telling us how skewed the torus is. The flat metric on the torus is now simply

$$ds^2 = dz d\bar{z}$$

subject to the identifications above.

A general metric on a torus can always be transformed to a flat metric for some value of τ. But the question that interests us is whether two tori, parameterized by different τ, are conformally equivalent. In general, the answer is no. The space of conformally inequivalent tori, parameterized by τ, is called the *moduli space* $\mathcal{M}$.

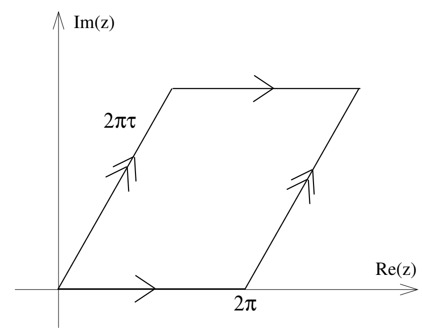


Figure 40:

However, there are some values of τ that do correspond to the same torus. In particular, there are a couple of obvious ways in which we can change τ without changing the torus. They go by the names of the S and T transformations:

- $T : \tau \to \tau + 1$: This clearly gives rise to the same torus, because the identification is now

$$z \equiv z + 2\pi \quad \text{and} \quad z \equiv z + 2\pi(\tau + 1) \equiv z + 2\pi\tau$$

- $S : \tau \to -1/\tau$: This simply flips the sides of the torus. For example, if $\tau = ia$ is purely imaginary, then this transformation maps $\tau \to i/a$, which can then be undone by a scaling.

It turns out that these two changes S and T are the only ones that keep the torus intact. They are sometimes called *modular transformations*. A general modular transformations is constructed from combinations of S and T and takes the form,

$$\tau \to \frac{a\tau + b}{c\tau + d} \qquad \text{with } ad - bc = 1 \tag{6.19}$$

where a, b, c and $d \in \mathbf{Z}$. This is the group $SL(2, \mathbf{Z})$. (In fact, we have our usual $\mathbf{Z}_2$ identification and the group is actually $PSL(2, \mathbf{Z}) = SL(2; \mathbf{Z})/\mathbf{Z}_2$). The moduli space $\mathcal{M}$ of the torus is given by

$$\mathcal{M} \cong \mathbf{C}/SL(2; \mathbf{Z})$$

What does this space look like? Using $T : \tau \to \tau + 1$, we can always shift τ until it lies within the interval

$$\mathrm{Re}\,\tau \in [-\tfrac{1}{2}, +\tfrac{1}{2}]$$

where the edges of the interval are identified. Meanwhile, $S : \tau \to -1/\tau$ inverts the

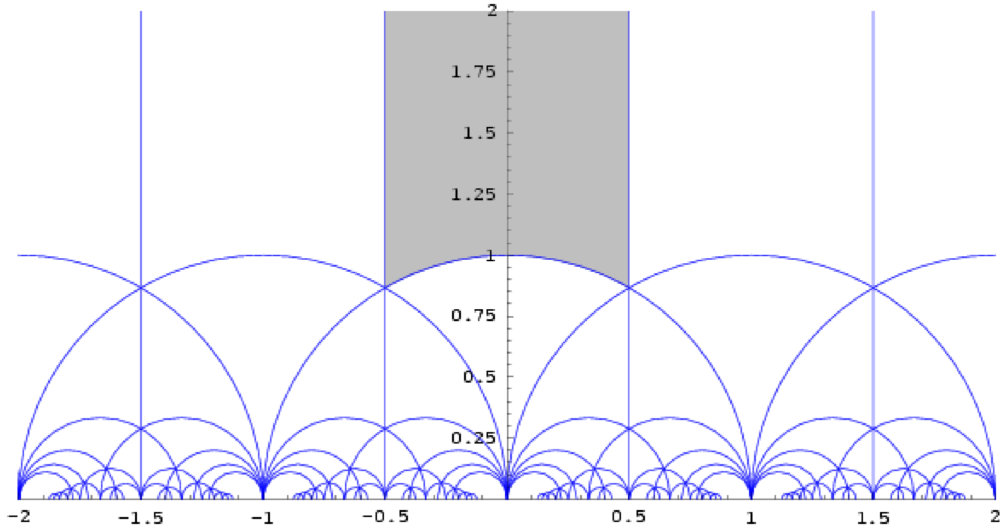


Figure 41: The fundamental domain.

modulus $|\tau|$, so we can use this to map a point inside the circle $|\tau| < 1$ to a point outside $|\tau| > 1$. One can show that by successive combinations of S and T, it is possible to map any point to lie within the shaded region shown in the figure, defined by

$$|\tau| \geq 1 \quad \text{ and } \quad \mathrm{Re}\,\tau \in [-\tfrac{1}{2}, +\tfrac{1}{2}\,]$$

This is referred to as the *fundamental domain* of $SL(2;\mathbf{Z})$.

We could have just as easily chosen one of the other fundamental domains shown in the figure. But the shaded region is the standard one.

Integrating over the Moduli Space

In string theory we're invited to sum over all metrics. After gauge fixing diffeomorphisms and Weyl invariance, we still need to integrate over all inequivalent tori. In other words, we integrate over the fundamental domain. The $SL(2;\mathbf{Z})$ invariant measure over the fundamental domain is

$$\int \frac{d^2\tau}{(\mathrm{Im}\,\tau)^2}$$

To see that this is $SL(2;\mathbf{Z})$ invariant, note that under a general transformation of the form (6.19) we have

$$d^2\tau \;\rightarrow\; \frac{d^2\tau}{|c\tau+d|^4} \quad \text{ and } \quad \mathrm{Im}\,\tau \;\rightarrow\; \frac{\mathrm{Im}\,\tau}{|c\tau+d|^2}$$

There's some physics lurking within these rather mathematical statements. The integration over the fundamental domain in string theory is analogous to the loop integral over momentum in quantum field theory. Consider the square tori defined by $\mathrm{Re}\,\tau = 0$. The tori with $\mathrm{Im}\,\tau \to \infty$ are squashed and chubby. They correspond to the infra-red region of loop momenta in a Feynman diagram. Those with $\mathrm{Im}\,\tau \to 0$ are long and thin. Those correspond to the ultra-violet limit of loop momenta in a Feynman diagram. Yet, as we have seen, we should not integrate over these UV regions of the loop since the fundamental domain does not stretch down that far. Or, more precisely, the thin tori are mapped to chubby tori. This corresponds to the fact that any putative UV divergence of string theory can always be reinterpreted as an IR divergence. This is the second manifestation of the well-behaved UV nature of string theory. We will see this more explicitly in the example of Section 6.4.2.

Finally, when computing a loop amplitude in string theory, we still need to worry about the residual gauge symmetry that is left unfixed after the map to the flat torus. In the case of tree-level amplitudes on the sphere, this residual gauge symmetry was due to the conformal Killing group $SL(2;\mathbf{C})$. For the torus, the conformal Killing group is generated by the obvious generators ∂_z and $\bar{\partial}_{\bar{z}}$. It is $U(1) \times U(1)$.

Higher Genus Surfaces

The moduli space $\mathcal{M}_g$ of the Riemann surface of genus $g > 1$ can be shown to have dimension,

$$\dim \mathcal{M}_g = 3g - 3$$

There are no conformal Killing vectors when $g > 1$. These facts can be demonstrated as an application of the Riemann-Roch theorem. For more details, see section 5.2 of Polchinski, or sections 3.3 and 8.2 of Green, Schwarz and Witten.

6.4.2 The One-Loop Partition Function

We won't compute any one-loop scattering amplitudes in string theory. Instead, we will look at something a little simpler: the one-loop vacuum to vacuum amplitude. A Euclidean worldsheet with periodic time has the interpretation of a finite temperature partition function for the theory defined on a cylinder. In $D = 26$ dimensional spacetime, it is related to the cosmological constant in bosonic string theory.

Consider firstly the partition function of a theory on a square torus, with $\mathrm{Re}\,\tau = 0$. Compactifying Euclidean time, with period $(\mathrm{Im}\,\tau)$ is equivalent to putting the theory at temperature $T = 1/(\mathrm{Im}\,\tau)$,

$$Z[\tau] = \mathrm{Tr}\; e^{-2\pi(\mathrm{Im}\,\tau)H}$$

where the Tr is over all states in the theory. For any CFT defined on a cylinder, the Hamiltonian given by

$$H = L_0 + \tilde{L}_0 - \frac{c + \tilde{c}}{24}$$

where the final term is the Casimir energy computed in Section 4.4.1.

What then is the interpretation of the vacuum amplitude computed on a torus with $\mathrm{Re}\,\tau \neq 0$? From the diagram, we see that the effect of such a skewed torus is to translate a given point around the cylinder by $\mathrm{Re}\,\tau$. But we know which operator implements such a translation: it is $\exp(2\pi i(\mathrm{Re}\,\tau)P)$, where P is the momentum operator on the cylinder. After the map to the plane, this becomes the rotation operator

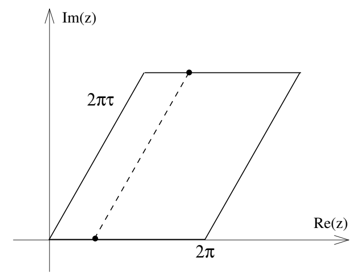


Figure 42:

$$P = L_0 - \tilde{L}_0$$

So the vacuum amplitude on the torus has the interpretation of the sum over all states in the theory, weighted by

$$Z[\tau] = \mathrm{Tr}\; e^{-2\pi(\mathrm{Im}\,\tau)(L_0+\tilde{L}_0)}\, e^{-2\pi i(\mathrm{Re}\,\tau)(L_0-\tilde{L}_0)}\, e^{2\pi(\mathrm{Im}\,\tau)(c+\tilde{c})/24}$$

We define

$$q = e^{2\pi i\tau} \quad , \quad \bar{q} = e^{-2\pi i\bar{\tau}}$$

The partition function can then be written in slick notation as

$$Z[\tau] = \mathrm{Tr}\; q^{L_0 - c/24}\; \bar{q}^{\tilde{L}_0 - \tilde{c}/24}$$

Let's compute this for the free string. We know that each scalar field X decomposes into a zero mode and an infinite number harmonic oscillator modes α_{-n} which create states of energy n. We'll deal with the zero mode shortly but, for now, we focus on the oscillators. Acting d times with the operator α_{-n} creates states with energy dn. This gives a contribution to $\mathrm{Tr}q^{L_0}$ of the form

$$\sum_{d=0}^{\infty} q^{nd} = \frac{1}{1 - q^n}$$

But the Fock space of a single scalar field is built by acting with oscillator modes $n \in \mathbf{Z}^+$. Including the central charge, $c = 1$, the contribution from the oscillator modes of a single scalar field is therefore

$$\mathrm{Tr}\ q^{L_0 - c/24} = \frac{1}{q^{1/24}} \prod_{n=1}^{\infty} \frac{1}{1-q^n}$$

There is a similar expression from the $\bar{q}^{\tilde{L}_0 - \tilde{c}/24}$ sector. We're still left with the contribution from the zero mode p of the scalar field. The contribution to the energy H of the state on the worldsheet is

$$\frac{1}{4\pi\alpha'} \int d\sigma\ (\alpha' p)^2 = \frac{1}{2}\alpha' p^2$$

The trace in the partition function requires us to sum over all states, which gives

$$\int \frac{dp}{2\pi}\ e^{-\pi\alpha' (\mathrm{Im}\,\tau) p^2} \sim \frac{1}{\sqrt{\alpha' \mathrm{Im}\,\tau}}$$

So, including both the zero mode and oscillators, we get the partition function for a single free scalar field,

$$Z_{\text{scalar}}[\tau] \sim \frac{1}{\sqrt{\alpha' \mathrm{Im}\,\tau}} \frac{1}{(q\bar{q})^{1/24}} \prod_{n=1}^{\infty} \frac{1}{1-q^n} \prod_{n=1}^{\infty} \frac{1}{1-\bar{q}^n} \tag{6.20}$$

where I haven't been careful to keep track of constant factors.

To build the string partition function, we should really work in covariant quantization and include the ghost fields. Here we'll cheat and work in lightcone gauge. This is dodgy because, if we do it honestly, much of the physics gets pushed to the $p^+ = 0$ limit of the lightcone momentum where the gauge choice breaks down. So instead we'll do it dishonestly.

In lightcone gauge, we have 24 oscillator modes. But we have 26 zero modes. (You may worry that we still have to impose level matching...this is the dishonest part of the calculation. We'll see partly where it comes from shortly). Finally, there's a couple of extra steps. We need to divide by the volume of the conformal Killing group. This is just $U(1) \times U(1)$, acting by translations along the cycles of the torus. The volume is just $\mathrm{Vol} = 4\pi^2\, \mathrm{Im}\,\tau$. Finally, we also need to integrate over the moduli space of the torus. Our final result, neglecting all constant factors, is

$$Z_{\text{string}} = \int d^2\tau\ \frac{1}{(\mathrm{Im}\,\tau)} \frac{1}{(\alpha' \mathrm{Im}\,\tau)^{13}}\ \frac{1}{q\bar{q}} \left(\prod_{n=1}^{\infty} \frac{1}{1-q^n} \right)^{24} \left(\prod_{n=1}^{\infty} \frac{1}{1-\bar{q}^n} \right)^{24} \tag{6.21}$$

Modular Invariance

The function appearing in the partition function for the scalar field has a name: it is the inverse of the Dedekind eta function

$$\eta(q) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n)$$

It was studied in the 1800s by mathematicians interested in the properties of functions under modular transformations $T : \tau \to \tau + 1$ and $S : \tau \to -1/\tau$. The eta-function satisfies the identities

$$\eta(\tau + 1) = e^{2\pi i/24} \eta(\tau) \quad \text{and} \quad \eta(-1/\tau) = \sqrt{-i\tau}\, \eta(\tau)$$

These two statements ensure that the scalar partition function (6.20) is a modular invariant function. Of course, that kinda had to be true: it follows from the underlying physics.

Written in terms of η, the string partition function (6.21) takes the form

$$Z_{\text{string}} = \int \frac{d^2\tau}{(\operatorname{Im}\tau)^2} \left(\frac{1}{\sqrt{\operatorname{Im}\tau}} \frac{1}{\eta(q)} \frac{1}{\bar{\eta}(\bar{q})} \right)^{24}$$

Both the measure and the integrand, are individually modular invariant.

6.4.3 Interpreting the String Partition Function

It's probably not immediately obvious what the string partition function (6.21) is telling us. Let's spend some time trying to understand it in terms of some simpler concepts.

We know that the free string describes an infinite number of particles with mass $m_n^2 = 4(n-1)/\alpha'$, $n = 0, 1, \ldots$. The string partition function should just be a sum over vacuum loops of each of these particles. We'll now show that it almost has this interpretation.

Firstly, let's figure out what the contribution from a single particle would be? We'll consider a free massive scalar field ϕ in D dimensions. The partition function is given by,

$$\begin{aligned} Z &= \int \mathcal{D}\phi \exp\left(-\frac{1}{2} \int d^D x \; \phi(-\partial^2 + m^2)\phi \right) \\ &\sim \det{}^{-1/2}(-\partial^2 + m^2) \\ &= \exp\left(\frac{1}{2} \int \frac{d^D p}{(2\pi)^D} \ln(p^2 + m^2) \right) \end{aligned}$$

This is the partition function of a field theory. It contains vacuum loops for all numbers of particles. To compare to the string partition function, we want the vacuum amplitude for just a single particle. But that's easy to extract. We write the field theory partition function as,

$$Z = \exp\left(Z_1\right) = \sum_{n=0}^{\infty} \frac{Z_1^n}{n!}$$

Each term in the sum corresponds to n particles propagating in a vacuum loop, with the $n!$ factor taking care of Bosonic statistics. So the vacuum amplitude for a single, free massive particle is simply

$$Z_1 = \frac{1}{2}\int \frac{d^Dp}{(2\pi)^D}\ \ln(p^2+m^2)$$

Clearly this diverges in the UV range of the integral, $p \to \infty$. There's a nice way to rewrite this integral using something known as Schwinger parameterization. We make use of the identity

$$\int_0^\infty dl\ e^{-xl} = \frac{1}{x} \quad \Rightarrow \quad \int_0^\infty dl\ \frac{e^{-xl}}{l} = -\ln x$$

We then write the single particle partition function as

$$Z_1 = \int \frac{d^Dp}{(2\pi)^D} \int_0^\infty \frac{dl}{2l}\ e^{-(p^2+m^2)l} \tag{6.22}$$

It's worth mentioning that there's another way to see that this is the single particle partition function that is a little closer in spirit to the method we used in string theory. We could start with the einbein form of the relativistic particle action (1.8). After fixing the gauge to $e = 1$, the exponent in (6.22) is the energy of the particle traversing a loop of length l. The integration measure dl/l sums over all possible sizes of loops.

We can happily perform the $\int d^Dp$ integral in (6.22). Ignoring numerical factors, we have

$$Z_1 = \int_0^\infty dl\ \frac{1}{l^{1+D/2}}\ e^{-m^2 l} \tag{6.23}$$

Note that the UV divergence as $p \to \infty$ has metamorphosised into a divergence associated to small loops as $l \to 0$.

Equation (6.23) gives the answer for a single particle of mass m. In string theory, we expect contributions from an infinite number of species of particles of mass m_n. Specializing to $D = 26$, we expect the partition function to be

$$Z = \int_0^\infty dl\, \frac{1}{l^{14}} \sum_{n=0}^{\infty} e^{-m_n^2 l}$$

But we know that the mass spectrum of the free string: it is given in terms of the L_0 and $\tilde{L}_0$ operators by

$$m^2 = \frac{4}{\alpha'}(L_0 - 1) = \frac{4}{\alpha'}(\tilde{L}_0 - 1) = \frac{2}{\alpha'}(L_0 + \tilde{L}_0 - 2)$$

subject to the constraint of level matching, $L_0 = \tilde{L}_0$. It's easy to impose level matching: we simply throw in a Kronecker delta in its integral representation,

$$\frac{1}{2\pi} \int_{-1/2}^{+1/2} ds\; e^{2\pi i s(L_0 - \tilde{L}_0)} = \delta_{L_0, \tilde{L}_0} \tag{6.24}$$

Replacing the sum over species, with the trace over the spectrum of states subject to level matching, the partition function becomes,

$$Z = \int_0^\infty dl\, \frac{1}{l^{14}} \int_{-1/2}^{+1/2} ds\; \mathrm{Tr}\, e^{2\pi i s(L_0 - \tilde{L}_0)}\, e^{-2(L_0 + \tilde{L}_0 - 2)l/\alpha'} \tag{6.25}$$

We again use the definition $q = \exp(2\pi i \tau)$, but this time the complex parameter τ is a combination of the length of the loop l and the auxiliary variable that we introduced to impose level matching,

$$\tau = s + \frac{2li}{\alpha'}$$

The trace over the spectrum of the string once gives the eta-functions, just as it did before. We're left with the result for the partition function,

$$Z_{\text{string}} = \int \frac{d^2\tau}{(\mathrm{Im}\,\tau)^2} \left(\frac{1}{\sqrt{\mathrm{Im}\,\tau}} \frac{1}{\eta(q)} \frac{1}{\bar{\eta}(\bar{q})} \right)^{24}$$

But this is exactly the same expression that we saw before. With a difference! In fact, the difference is hidden in the notation: it is the range of integration for $d^2\tau$ which can be found in the original expressions (6.23) and (6.24). $\mathrm{Re}\,\tau$ runs over the same interval $[-\frac{1}{2}, +\frac{1}{2}]$ that we saw in string theory. As is clear from this discussion, it is this integral which implements level matching. The difference comes in the range of $\mathrm{Im}\,\tau$ which, in this naive analysis, runs over $[0, \infty)$. This is in stark contrast to string theory where we only integrate over the fundamental domain.

This highlights our previous statement: the potential UV divergences in field theory are encountered in the region $\mathrm{Im}\,\tau \sim l \to 0$. In the above analysis, this corresponds to particles traversing small loops. But this region is simply absent in the correct string theory computation. It is mapped, by modular invariance, to the infra-red region of large loops.

It is often said that in the $g_s \to 0$ limit string theory becomes a theory of an infinite number of free particles. This is true of the spectrum. But this calculation shows that it's not really true when we compute loops because the modular invariance means that we integrate over a different range of momenta in string theory than in a naive field theory approach.

So what happens in the infra-red region of our partition function? The easiest place to see it is in the $l \to \infty$ limit of the integral (6.25). We see that the integral is dominated by the lightest state which, for the bosonic string is the tachyon. This has $m^2 = -4/\alpha'$, or $(L_0 + \tilde{L}_0 - 2) = -2$. This gives a contribution to the partition function of,

$$\int^{\infty} \frac{dl}{l^{14}}\; e^{+4l/\alpha'}$$

which clearly diverges. This IR divergence of the one-loop partition function is another manifestation of tachyonic trouble. In the superstring, there is no tachyon and the IR region is well-behaved.

6.4.4 So is String Theory Finite?

The honest answer is that we don't know. The UV finiteness that we saw above holds for all one-loop amplitudes. This means, in particular, that we have a one-loop finite theory of gravity interacting with matter in higher dimensions. This is already remarkable.

There is more good news: One can show that UV finiteness continues to hold at the two-loops. And, for the superstring, state-of-the-art techniques using the "pure-spinor" formalism show that certain objects remain finite up to five-loops. Moreover, the exponential suppression (6.14) that we saw when all momentum exchanges are large continues to hold for all amplitudes.

However, no general statement of finiteness has been proven. The danger lurks in the singular points in the integration over Riemann surfaces of genus 3 and higher.

6.4.5 Beyond Perturbation Theory?

From the discussion in this section, it should be clear that string perturbation theory is entirely analogous to the Feynman diagram expansion in field theory. Just as in field theory, one can show that the expansion in g_s is asymptotic. This means that the series does not converge, but we can nonetheless make sense of it.

However, we know that there are many phenomena in quantum field theory that aren't captured by Feynman diagrams. These include confinement in the strongly coupled regime and instantons and solitons in the weakly coupled regime. Does this mean that we are missing similarly interesting phenomena in string theory? The answer is almost certainly yes! In this section, I'll very briefly allude to a couple of more advanced topics which allow us to go beyond the perturbative expansion in string theory. The goal is not really to teach you these things, but merely to familiarize you with some words.

One way to proceed is to keep quantum field theory as our guide and try to build a non-perturbative definition of string theory in terms of a path integral. We've already seen that the Polyakov path integral over worldsheets is equivalent to Feynman diagrams. So we need to go one step further. What does this mean? Recall that in QFT, a field creates a particle. In string theory, we are now looking for a field which creates a loop of string. We should have a different field for each configuration of the string. In other words, our field should itself be a function of a function: $\Phi(X^\mu(\sigma))$. Needless to say, this is quite a complicated object. If we were brave, we could then consider the path integral for this field,

$$Z = \int \mathcal{D}\Phi \; e^{iS[\Phi(X(\sigma))]}$$

for some suitable action $S[\Phi]$. The idea is that this path integral should reproduce the perturbative string expansion and, furthermore, defines a non-perturbative completion of the theory. This line of ideas is known as *string field theory*. It should be clear that this is one step further in the development: particles $\rightarrow$ fields $\rightarrow$ string fields. Or, in more historical language, if field theory is "second quantization", then string field theory is "third quantization".

String field theory has been fairly successful for the open string and some interesting non-perturbative results have been obtained in this manner. However, for the closed string this approach has been much less useful. It is usually thought that there are deep reasons behind the failure of closed string field theory, related to issues that we mentioned at the beginning of this section: there are no off-shell quantities in a theory of gravity. Moreover, we mentioned in Section 4 that a theory of interacting open strings necessarily includes closed strings, so somehow the open string field theory should already contain gravity and closed strings. Quite how this comes about is still poorly understood.

There are other ways to get a handle on non-perturbative aspects of string theory using the low-energy effective action (we will describe what the "low-energy effective action" is in the next section). Typically these techniques rely on supersymmetry to provide a window into the strongly coupled regime and so work only for the superstring. These methods have been extremely successful and any course on superstring theory would be devoted to explaining various aspects of such as dualities and M-theory.

Finally, in asymptotically AdS spacetimes, the AdS/CFT correspondence gives a non-perturbative definition of string theory and quantum gravity in the bulk in terms of Yang-Mills theory, or something similar, on the boundary. In some sense, the boundary field theory is a "string field theory".

6.5 Appendix: Games with Integrals and Gamma Functions

The gamma function is defined by the integral representation

$$\Gamma(z) = \int_0^\infty dt\; t^{z-1} e^{-t} \tag{6.26}$$

which converges if Re$z > 0$. It has a unique analytic expression to the whole z-plane. The absolute value of the gamma function over the z-plane is shown in the figure.

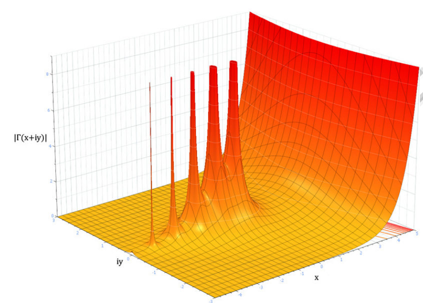


Figure 43:

The gamma function has a couple of important properties. Firstly, it can be thought of as the analytic continuation of the factorial function for positive integers, meaning

$$\Gamma(n) = (n-1)! \qquad n \in \mathbf{Z}^+$$

Secondly, $\Gamma(z)$ has poles at non-positive integers. More precisely when $z \approx -n$, with $n = 0, 1, \ldots$, there is the expansion

$$\Gamma(z) \approx \frac{1}{z+n}\,\frac{(-1)^n}{n!}$$

The Euler Beta Function

The Euler beta function is defined for $x,\, y \in \mathbf{C}$ by

$$B(x,y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

It has the integral representation

$$B(x,y)=\int_0^1 dt\ t^{x-1}(1-t)^{y-1} \tag{6.27}$$

Let's prove this statement. We start by looking at

$$\Gamma(x)\Gamma(y)=\int_0^\infty du\int_0^\infty dv\ e^{-u}u^{x-1}e^{-v}v^{y-1}$$

We write $u=a^2$ and $v=b^2$ so the integral becomes

$$\begin{aligned}\Gamma(x)\Gamma(y)&=4\int_0^\infty da\int_0^\infty db\ e^{-(a^2+b^2)}a^{2x-1}b^{2y-1}\\&=\int_{-\infty}^\infty da\int_{-\infty}^\infty db\ e^{-(a^2+b^2)}|a|^{2x-1}|b|^{2y-1}\end{aligned}$$

We now change coordinates once more, this time to polar $a=r\cos\theta$ and $b=r\sin\theta$. We get

$$\begin{aligned}\Gamma(x)\Gamma(y)&=\int_0^\infty rdr\ e^{-r^2}r^{2x+2y-2}\int_0^{2\pi}d\theta\ |\cos\theta|^{2x-1}|\sin\theta|^{2y-1}\\&=\frac{1}{2}\Gamma(x+y)\times 4\int_0^{\pi/2}d\theta\ (\cos\theta)^{2x-1}(\sin\theta)^{2y-1}\\&=\Gamma(x+y)\int_0^1 dt\ (1-t)^{y-1}t^{x-1}\end{aligned}$$

where, in the final line, we made the substitution $t=\cos^2\theta$. This completes the proof.

The Virasoro-Shapiro Amplitude

In the closed string computation, we came across the integral

$$C(a,b)=\int d^2z\ |z|^{2a-2}|1-z|^{2b-2}$$

We will now evaluate this and show that it is given by (6.11). We start by using a trick. We can write

$$|z|^{2a-2}=\frac{1}{\Gamma(1-a)}\int_0^\infty dt\ t^{-a}e^{-|z|^2t}$$

which follows from the definition (6.26) of the gamma function. Similarly, we can write

$$|1-z|^{2b-2} = \frac{1}{\Gamma(1-b)} \int_0^\infty du\ u^{-b} e^{-|1-z|^2 u}$$

We decompose the complex coordinate $z = x + iy$, so that the measure of the integral is $d^2z = 2dxdy$. We can then write the integral $C(a,b)$ as

$$\begin{aligned} C(a,b) &= \int \frac{d^2z\, du\, dt}{\Gamma(1-a)\Gamma(1-b)}\ t^{-a} u^{-b} e^{-|z|^2 t} e^{-|1-z|^2 u} \\ &= 2\int \frac{dx\, dy\, du\, dt}{\Gamma(1-a)\Gamma(1-b)}\ t^{-a} u^{-b}\, e^{-(t+u)(x^2+y^2)+2xu-u} \\ &= 2\int \frac{dx\, dy\, du\, dt}{\Gamma(1-a)\Gamma(1-b)}\ t^{-a} u^{-b} \exp\left(-(t+u)\left[\left(x-\frac{u}{t+u}\right)^2 + y^2\right] - u + \frac{u^2}{t+u}\right) \end{aligned}$$

Now we do the $dxdy$ integral which is simply Gaussian. We find

$$C(a,b) = \frac{2\pi}{\Gamma(1-a)\Gamma(1-b)} \int_0^\infty du\, dt\ \frac{t^{-a}u^{-b}}{t+u}\ e^{-tu/(t+u)}$$

Finally, we make a change of variables. We write $t = \alpha\beta$ and $u = (1-\beta)\alpha$. In order for t and u to take values in the range $[0,\infty)$, we require $\alpha \in [0,\infty)$ and $\beta \in [0,1]$. Taking into account the Jacobian arising from this transformation, which is simply α, the integral becomes

$$C(a,b) = \frac{2\pi}{\Gamma(1-a)\Gamma(1-b)} \int d\alpha\, d\beta\ \frac{\alpha^{1-a-b}}{\alpha}\ \beta^{-a}(1-\beta)^{-b} e^{-\alpha\beta(1-\beta)}$$

But we recognize the integral over $d\alpha$: it is simply

$$\int_0^\infty d\alpha\ \alpha^{-a-b} e^{-\beta\alpha(1-\beta)} = [\beta(1-\beta)]^{a+b-1}\Gamma(1-a-b)$$

We write $c = 1-a-b$. Finally, we're left with

$$C(a,b) = \frac{2\pi\Gamma(c)}{\Gamma(1-a)\Gamma(1-b)} \int_0^1 d\beta\ (1-\beta)^{a-1}\beta^{b-1}$$

But the final integral is the Euler beta function (6.27). This gives us our promised result,

$$C(a,b) = \frac{2\pi\Gamma(a)\Gamma(b)\Gamma(c)}{\Gamma(1-a)\Gamma(1-b)\Gamma(1-c)}$$

7. Low Energy Effective Actions

So far, we've only discussed strings propagating in flat spacetime. In this section we will consider strings propagating in different backgrounds. This is equivalent to having different CFTs on the worldsheet of the string.

There is an obvious generalization of the Polyakov action to describe a string moving in curved spacetime,

$$S = \frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{g}\, g^{\alpha\beta}\,\partial_\alpha X^\mu\,\partial_\beta X^\nu\, G_{\mu\nu}(X) \tag{7.1}$$

Here $g_{\alpha\beta}$ is again the worldsheet metric. This action describes a map from the worldsheet of the string into a spacetime with metric $G_{\mu\nu}(X)$. (Despite its name, this metric is not to be confused with the Einstein tensor which we won't have need for in this lecture notes).

Actions of the form (7.1) are known as *non-linear sigma models*. (This strange name has its roots in the history of pions). In this context, the D-dimensional spacetime is sometimes called the *target space*. Theories of this type are important in many aspects of physics, from QCD to condensed matter.

Although it's obvious that (7.1) describes strings moving in curved spacetime, there's something a little fishy about just writing it down. The problem is that the quantization of the closed string already gave us a graviton. If we want to build up some background metric $G_{\mu\nu}(X)$, it should be constructed from these gravitons, in much the same manner that a laser beam is made from the underlying photons. How do we see that the metric in (7.1) has anything to do with the gravitons that arise from the quantization of the string?

The answer lies in the use of vertex operators. Let's expand the metric as a small fluctuation around flat space

$$G_{\mu\nu}(X) = \delta_{\mu\nu} + h_{\mu\nu}(X)$$

Then the partition function that we build from the action (7.1) is related to the partition function for a string in flat space by

$$Z = \int \mathcal{D}X\mathcal{D}g\; e^{-S_{\rm Poly}-V} = \int \mathcal{D}X\mathcal{D}g\; e^{-S_{\rm Poly}}(1 - V + \frac{1}{2}V^2 + \ldots)$$

where S_{Poly} is the action for the string in flat space given in (1.22) and V is the expression

$$V = \frac{1}{4\pi\alpha'}\int d^2\sigma \sqrt{g}\, g^{\alpha\beta}\partial_\alpha X^\mu\, \partial_\beta X^\nu\, h_{\mu\nu}(X) \tag{7.2}$$

But we've seen this before: it's the vertex operator associated to the graviton state of the string! For a plane wave, corresponding to a graviton with polarization given by the symmetric, traceless tensor $\zeta_{\mu\nu}$ and momentum p^μ, the fluctuation is given by

$$h_{\mu\nu}(X) = \zeta_{\mu\nu}\, e^{ip\cdot X}$$

With this choice, the expression (7.2) agrees with the vertex operator (5.9). But in general, we could take any linear superposition of plane waves to build up a general fluctuation $h_{\mu\nu}(X)$.

We know that inserting a single copy of V in the path integral corresponds to the introduction of a single graviton state. Inserting e^V in the path integral corresponds to a coherent state of gravitons, changing the metric from $\delta_{\mu\nu}$ to $\delta_{\mu\nu} + h_{\mu\nu}$. In this way we see that the background curved metric of (7.1) is indeed built of the quantized gravitons that we first met back in Section 2.

7.1 Einstein's Equations

In conformal gauge, the Polyakov action in flat space reduces to a free theory. This fact was extremely useful, allowing us to compute the spectrum of the theory. But on a curved background, it is no longer the case. In conformal gauge, the worldsheet theory is described by an interacting two-dimensional field theory,

$$S = \frac{1}{4\pi\alpha'}\int d^2\sigma\; G_{\mu\nu}(X)\,\partial_\alpha X^\mu\,\partial^\alpha X^\nu \tag{7.3}$$

To understand these interactions in more detail, let's expand around a classical solution which we take to simply be a string sitting at a point $\bar{x}^\mu$.

$$X^\mu(\sigma) = \bar{x}^\mu + \sqrt{\alpha'}\,Y^\mu(\sigma)$$

Here Y^μ are the dynamical fluctuations about the point which we assume to be small. The factor of $\sqrt{\alpha'}$ is there for dimensional reasons: since $[X] = -1$, we have $[Y] = 0$ and statements like $Y \ll 1$ make sense. Expanding the Lagrangian gives

$$G_{\mu\nu}(X)\,\partial X^\mu \partial X^\nu = \alpha'\left[G_{\mu\nu}(\bar{x}) + \sqrt{\alpha'}G_{\mu\nu,\omega}(\bar{x})\,Y^\omega + \frac{\alpha'}{2}\,G_{\mu\nu,\omega\rho}(\bar{x})\,Y^\omega Y^\rho + \ldots\right]\partial Y^\mu\,\partial Y^\nu$$

Each of the coefficients $G_{\mu\nu,\ldots}$ in the Taylor expansion are coupling constants for the interactions of the fluctuations Y^μ. The theory has an infinite number of coupling constants and they are nicely packaged into the function $G_{\mu\nu}(X)$.

We want to know when this field theory is weakly coupled. Obviously this requires the whole infinite set of coupling constants to be small. Let's try to characterize this in a crude manner. Suppose that the target space has characteristic radius of curvature r_c, meaning schematically that

$$\frac{\partial G}{\partial X} \sim \frac{1}{r_c}$$

The radius of curvature is a length scale, so $[r_c] = -1$. From the expansion of the metric, we see that the effective dimensionless coupling is given by

$$\frac{\sqrt{\alpha'}}{r_c} \tag{7.4}$$

This means that we can use perturbation theory to study the CFT (7.3) if the spacetime metric only varies on scales much greater than $\sqrt{\alpha'}$. The perturbation series in $\sqrt{\alpha'}/r_c$ is usually called the α'-expansion to distinguish it from the g_s expansion that we saw in the previous section. Typically a quantity computed in string theory is given by a double perturbation expansion: one in α' and one in g_s.

If there are regions of spacetime where the radius of curvature becomes comparable to the string length scale, $r_c \sim \sqrt{\alpha'}$, then the worldsheet CFT is strongly coupled and we will need to develop new methods to solve it. Notice that strong coupling in α' is hard, but the problem is at least well-defined in terms of the worldsheet path integral. This is qualitatively different to the question of strong coupling in g_s for which, as discussed in Section 6.4.5, we're really lacking a good definition of what the problem even means.

7.1.1 The Beta Function

Classically, the theory defined by (7.3) is conformally invariant. But this is not necessarily true in the quantum theory. To regulate divergences we will have to introduce a UV cut-off and, typically, after renormalization, physical quantities depend on the scale of a given process μ. If this is the case, the theory is no longer conformally invariant. There are plenty of theories which classically possess scale invariance which is broken quantum mechanically. The most famous of these is Yang-Mills.

As we've discussed several times, in string theory conformal invariance is a gauge symmetry and we can't afford to lose it. Our goal in this section is to understand the circumstances under which (7.3) retains conformal invariance at the quantum level.

The object which describes how couplings depend on a scale μ is called the β-function. Since we have a functions worth of couplings, we should really be talking about a β-functional, schematically of the form

$$\beta_{\mu\nu}(G) \sim \mu \frac{\partial G_{\mu\nu}(X;\mu)}{\partial \mu}$$

The quantum theory will be conformally invariant only if

$$\beta_{\mu\nu}(G) = 0$$

We now compute this for the non-linear sigma model at one-loop. Our strategy will be to isolate the UV divergence of the theory and figure out what kind of counterterm we should add. The beta-function will vanish if this counterterm vanishes.

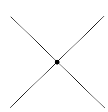
The analysis is greatly simplified by a cunning choice of coordinates. Around any point $\bar{x}$, we can always pick Riemann normal coordinates such that the expansion in $X^\mu = \bar{x}^\mu + \sqrt{\alpha'}\, Y^\mu$ gives

$$G_{\mu\nu}(X) = \delta_{\mu\nu} - \frac{\alpha'}{3}\mathcal{R}_{\mu\lambda\nu\kappa}(\bar{x})Y^\lambda Y^\kappa + \mathcal{O}(Y^3)$$

To quartic order in the fluctuations, the action becomes

$$S = \frac{1}{4\pi}\int d^2\sigma\ \partial Y^\mu \partial Y^\nu \delta_{\mu\nu} - \frac{\alpha'}{3}\mathcal{R}_{\mu\lambda\nu\kappa}\, Y^\lambda Y^\kappa \partial Y^\mu \partial Y^\nu$$

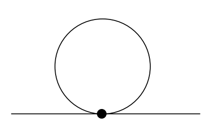
We can now treat this as an interacting quantum field theory in two dimensions. The quartic interaction gives a vertex with the Feynman rule,



$$\sim \mathcal{R}_{\mu\lambda\nu\kappa}\,(k^\mu \cdot k^\nu)$$

where k^μ_α is the 2d momentum ($\alpha = 1,2$ is a worldsheet index) for the scalar field Y^μ. It sits in the Feynman rules because we are talking about derivative interactions.

Now we've reduced the problem to a simple interacting quantum field theory, we can compute the β-function using whatever method we like. The divergence in the theory comes from the one-loop diagram



It's actually simplest to think about this diagram in position space. The propagator for a scalar particle is

$$\langle Y^\lambda(\sigma) Y^\kappa(\sigma')\rangle = -\frac{1}{2}\,\delta^{\lambda\kappa}\,\ln|\sigma-\sigma'|^2$$

For the scalar field running in the loop, the beginning and end point coincide. The propagator diverges as $\sigma\to\sigma'$, which is simply reflecting the UV divergence that we would see in the momentum integral around the loop.

To isolate this divergence, we choose to work with dimensional regularization, with $d = 2+\epsilon$. The propagator then becomes,

$$\begin{aligned}\langle Y^\lambda(\sigma) Y^\kappa(\sigma')\rangle &= 2\pi\delta^{\lambda\kappa}\int \frac{d^{2+\epsilon}k}{(2\pi)^{2+\epsilon}}\,\frac{e^{ik\cdot(\sigma-\sigma')}}{k^2} \\ &\longrightarrow\ \frac{\delta^{\lambda\kappa}}{\epsilon} \qquad \text{as } \sigma\to\sigma'\end{aligned}$$

The necessary counterterm for this divergence can be determined simply by replacing $Y^\lambda Y^\kappa$ in the action with $\langle Y^\lambda Y^\kappa\rangle$. To subtract the $1/\epsilon$ term, we add the counterterm

$$\mathcal{R}_{\mu\lambda\nu\kappa}\,Y^\lambda Y^\kappa\partial Y^\mu\partial Y^\nu\ \to\ \mathcal{R}_{\mu\lambda\nu\kappa}\,Y^\lambda Y^\kappa\partial Y^\mu\partial Y^\nu - \frac{1}{\epsilon}\mathcal{R}_{\mu\nu}\,\partial Y^\mu\partial Y^\nu$$

One can check that this can be absorbed by a wavefunction renormalization $Y^\mu \to Y^\mu + (\alpha'/6\epsilon)\,\mathcal{R}^\mu{}_\nu Y^\nu$, together with the renormalization of the coupling constant which, in our theory, is the metric $G_{\mu\nu}$. We require,

$$G_{\mu\nu}\ \to\ G_{\mu\nu} + \frac{\alpha'}{\epsilon}\,\mathcal{R}_{\mu\nu} \tag{7.5}$$

From this we learn the beta function of the theory and the condition for conformal invariance. It is

$$\beta_{\mu\nu}(G) = \alpha'\mathcal{R}_{\mu\nu} = 0 \tag{7.6}$$

This is a magical result! The requirement for the sigma-model to be conformally invariant is that the target space must be Ricci flat: $\mathcal{R}_{\mu\nu}=0$. Or, in other words, the background spacetime in which the string moves must obey the vacuum Einstein equations! We see that the equations of general relativity also describe the renormalization group flow of 2d sigma models.

There are several more magical things just around the corner, but it's worth pausing to make a few diverse comments.

Beta Functions and Weyl Invariance

The above calculation effectively studies the breakdown of conformal invariance in the CFT (7.3) on a flat worldsheet. We know that this should be the same thing as the breakdown of Weyl invariance on a curved worldsheet. Since this is such an important result, let's see how it works from this other perspective. We can consider the worldsheet metric

$$g_{\alpha\beta} = e^{2\phi}\delta_{\alpha\beta}$$

Then, in dimensional regularization, the theory is not Weyl invariant in $d = 2 + \epsilon$ dimensions because the contribution from $\sqrt{g}$ does not quite cancel that from the inverse metric $g^{\alpha\beta}$. The action is

$$\begin{aligned} S &= \frac{1}{4\pi\alpha'}\int d^{2+\epsilon}\sigma\ e^{\phi\epsilon}\partial_\alpha X^\mu\,\partial^\alpha X^\nu\,G_{\mu\nu}(X) \\ &\approx \frac{1}{4\pi\alpha'}\int d^{2+\epsilon}\sigma\ (1+\phi\epsilon)\,\partial_\alpha X^\mu\,\partial^\alpha X^\nu\,G_{\mu\nu}(X) \end{aligned}$$

where, in this expression, the $\alpha = 1, 2$ index is now raised and lowered with $\delta_{\alpha\beta}$. If we replace $G_{\mu\nu}$ in this expression with the renormalized metric (7.5), we see that there's a term involving ϕ which remains even as $\epsilon \to 0$,

$$S = \frac{1}{4\pi\alpha'}\int d^2\sigma\ \partial_\alpha X^\mu\partial^\alpha X^\nu\ \left[G_{\mu\nu}(X) + \alpha'\phi\,\mathcal{R}_{\mu\nu}(X)\right]$$

This indicates a breakdown of Weyl invariance. Indeed, we can look at our usual diagnostic for Weyl invariance, namely the vanishing of $T^\alpha_{\ \alpha}$. In conformal gauge, this is given by

$$T_{\alpha\beta} = +\frac{4\pi}{\sqrt{g}}\frac{\partial S}{\partial g^{\alpha\beta}} = -2\pi\frac{\partial S}{\partial\phi}\delta_{\alpha\beta} \quad\Rightarrow\quad T^\alpha_{\ \alpha} = -\frac{1}{2}\mathcal{R}_{\mu\nu}\,\partial X^\mu\,\partial X^\nu$$

In this way of looking at things, we define the β-function to be the coefficient in front of $\partial X\partial X$, namely

$$T^\alpha_{\ \alpha} = -\frac{1}{2\alpha'}\,\beta_{\mu\nu}\,\partial X^\mu\partial X^\nu$$

Again, we have the result

$$\beta_{\mu\nu} = \alpha'\mathcal{R}_{\mu\nu}$$

7.1.2 Ricci Flow

In string theory we only care about conformal theories with Ricci flat metrics. (And generalizations of this result that we will discuss shortly). However, in other areas of physics and mathematics, the RG flow itself is important. It is usually called Ricci flow,

$$\mu \frac{\partial G_{\mu\nu}}{\partial \mu} = \alpha' \mathcal{R}_{\mu\nu} \tag{7.7}$$

which dictates how the metric changes with scale μ.

As an illustrative and simple example, consider the target space $\mathbf{S}^2$ with radius r. This is an important model in condensed matter physics where it describes the low-energy limit of a one-dimensional Heisenberg spin chain. It is sometimes called the $O(3)$ sigma-model. Because the sphere is a symmetric space, the only effect of the RG flow is to make the radius scale dependent: $r = r(\mu)$. The beta function is given by

$$\mu \frac{\partial r^2}{\partial \mu} = \frac{\alpha'}{2\pi}$$

Hence r gets large as we go towards the UV and small towards the IR. Since the coupling is $1/r$, this means that the non-linear sigma model with $\mathbf{S}^2$ target space is asymptotically free. At low energies, the theory is strongly coupled and perturbative calculations — such as this one-loop beta function — are no longer trusted. In particular, one can show that the $\mathbf{S}^2$ sigma-model develops a mass gap in the IR.

The idea of Ricci flow (7.7) was recently used by Perelman to prove the Poincaré conjecture. In fact, Perelman used a slightly generalized version of Ricci flow which we will see shortly. In the language of string theory, he introduced the dilaton field.

7.2 Other Couplings

We've understood how strings couple to a background spacetime metric. But what about the other modes of the string? In Section 2, we saw that a closed string has further massless states which are associated to the anti-symmetric tensor $B_{\mu\nu}$ and the dilaton Φ. We will now see how the string reacts if these fields are turned on in spacetime.

7.2.1 Charged Strings and the B field

Let's start by looking at how strings couple to the anti-symmetric field $B_{\mu\nu}$. We discussed the vertex operator associated to this state in Section 5.4.1. It is given in

(5.9) and takes the same form as the graviton vertex operator, but with $\zeta_{\mu\nu}$ anti-symmetric. It is a simple matter to exponentiate this, to get an expression for how strings propagate in background $B_{\mu\nu}$ field. We'll keep the curved metric $G_{\mu\nu}$ as well to get the general action,

$$S = \frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{g}\, \left(G_{\mu\nu}(X)\, \partial_\alpha X^\mu\, \partial_\beta X^\nu g^{\alpha\beta} + iB_{\mu\nu}(X)\, \partial_\alpha X^\mu\, \partial_\beta X^\nu\, \epsilon^{\alpha\beta} \right) \tag{7.8}$$

Where $\epsilon^{\alpha\beta}$ is the anti-symmetric 2-tensor, normalized such that $\sqrt{g}\epsilon^{12} = +1$. (The factor of i is there in the action because we're in Euclidean space and this new term has a single "time" derivative). The action retains invariance under worldsheet reparameterizations and Weyl rescaling.

So what is the interpretation of this new term? We will now show that we should think of the field $B_{\mu\nu}$ as analogous to the gauge potential A_μ in electromagnetism. The action (7.8) is telling us that the string is "electrically charged" under $B_{\mu\nu}$.

Gauge Potentials

We'll take a short detour to remind ourselves about some pertinent facts in electromagnetism. Let's start by returning to a point particle. We know that a charged point particle couples to a background gauge potential A_μ through the addition of a worldline term to the action,

$$\int d\tau\ A_\mu(X)\, \dot{X}^\mu\ . \tag{7.9}$$

If this relativistic form looks a little unfamiliar, we can deconstruct it by working in static gauge with $X^0 \equiv t = \tau$, where it reads

$$\int dt\ A_0(X) + A_i(X)\, \dot{X}^i\ ,$$

which should now be recognizable as the Lagrangian that gives rise to the Coulomb and Lorentz force laws for a charged particle.

So what is the generalization of this kind of coupling for a string? First note that (7.9) has an interesting geometrical structure. It is the pull-back of the one-form $A = A_\mu dX^\mu$ in spacetime onto the worldline of the particle. This works because A is a one-form and the worldline is one-dimensional. Since the worldsheet of the string is two-dimensional, the analogous coupling should be to a two-form in spacetime. This is an anti-symmetric tensor field with two indices, $B_{\mu\nu}$. The pull-back of $B_{\mu\nu}$ onto the worldsheet gives the interaction,

$$\int d^2\sigma\; B_{\mu\nu}(X)\,\partial_\alpha X^\mu\,\partial_\beta X^\nu\,\epsilon^{\alpha\beta}\ . \tag{7.10}$$

This is precisely the form of the interaction we found in (7.8).

The point particle coupling (7.9) is invariant under gauge transformations of the background field $A_\mu \to A_\mu + \partial_\mu\alpha$. This follows because the Lagrangian changes by a total derivative. There is a similar statement for the two-form $B_{\mu\nu}$. The spacetime gauge symmetry is,

$$B_{\mu\nu} \to B_{\mu\nu} + \partial_\mu C_\nu - \partial_\nu C_\mu \tag{7.11}$$

under which the Lagrangian (7.10) changes by a total derivative.

In electromagnetism, one can construct the gauge invariant electric and magnetic fields which are packaged in the two-form field strength $F = dA$. Similarly, for $B_{\mu\nu}$, the gauge invariant field strength $H = dB$ is a three-form,

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

This 3-form H is sometimes known as the *torsion*. It plays the same role as torsion in general relativity, providing an anti-symmetric component to the affine connection.

7.2.2 The Dilaton

Let's now figure out how the string couples to a background dilaton field $\Phi(X)$. This is more subtle. A naive construction of the vertex operator is not primary and one must work a little harder. The correct derivation of the vertex operators can be found in Polchinski. Here I will simply give the coupling and explain some important features.

The action of a string moving in a background involving profiles for the massless fields $G_{\mu\nu}$, $B_{\mu\nu}$ and $\Phi(X)$ is given by

$$S = \frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{g}\ \Big(G_{\mu\nu}(X)\,\partial_\alpha X^\mu\,\partial_\beta X^\nu g^{\alpha\beta} + iB_{\mu\nu}(X)\,\partial_\alpha X^\mu\,\partial_\beta X^\nu\,\epsilon^{\alpha\beta} + \alpha'\,\Phi(X)\,R^{(2)}\Big) \tag{7.12}$$

where $R^{(2)}$ is the two-dimensional Ricci scalar of the worldsheet. (Up until now, we've always denoted this simply as R but we'll introduce the superscript from hereon to distinguish the worldsheet Ricci scalar from the spacetime Ricci scalar).

The coupling to the dilaton is surprising for several reasons. Firstly, we see that the term in the action vanishes on a flat worldsheet, $R^{(2)} = 0$. This is one of the reasons that it's a little trickier to determine this coupling using vertex operators.

However, the most surprising thing about the coupling to the dilaton is that it *does not* respect Weyl invariance! Since a large part of this course has been about understanding the implications of Weyl invariance, why on earth are we willing to throw it away now?! The answer, of course, is that we're not. Although the dilaton coupling does violate Weyl invariance, there is a way to restore it. We will explain this shortly. But firstly, let's discuss one crucially important implication of the dilaton coupling (7.12).

The Dilaton and the String Coupling

There is an exception to the statement that the classical coupling to the dilaton violates Weyl invariance. This arises when the dilaton is constant. For example, suppose

$$\Phi(X) = \lambda \quad , \text{ a constant}$$

Then the dilaton coupling reduces to something that we've seen before: it is

$$S_{\text{dilaton}} = \lambda\chi$$

where χ is the Euler character of the worldsheet that we introduced in (6.4). This tells us something important: the constant mode of the dilaton, $\langle\Phi\rangle$ determines the string coupling constant. This constant mode is usually taken to be the asymptotic value of the dilaton,

$$\Phi_0 = \underset{X\to\infty}{\text{limit}}\ \Phi(X) \tag{7.13}$$

The string coupling is then given by

$$g_s = e^{\Phi_0} \tag{7.14}$$

So the string coupling is not an independent parameter of string theory: it is the expectation value of a field. This means that, just like the spacetime metric $G_{\mu\nu}$ (or, indeed, like the Higgs vev) it can be determined dynamically.

We've already seen that our perturbative expansion around flat space is valid as long as $g_s \ll 1$. But now we have a stronger requirement: we can only trust perturbation theory if the string is localized in regions of space where $e^{\Phi(X)} \ll 1$ for all X. If the string ventures into regions where $e^{\Phi(X)}$ is of order 1, then we will need to use techniques that don't rely on string perturbation theory as described in Section 6.4.5.

7.2.3 Beta Functions

We now return to understanding how we can get away with the violation of Weyl invariance in the dilaton coupling (7.12). The key to this is to notice the presence of α' in front of the dilaton coupling. It's there simply on dimensional grounds. (The other two terms in the action both come with derivatives $[\partial X] = -1$, so don't need any powers of α').

However, recall that α' also plays the role of the loop-expansion parameter (7.4) in the non-linear sigma model. This means that the classical lack of Weyl invariance in the dilaton coupling can be compensated by a one-loop contribution arising from the couplings to $G_{\mu\nu}$ and $B_{\mu\nu}$.

To see this explicitly, one can compute the beta-functions for the two-dimensional field theory (7.12). In the presence of the dilaton coupling, it's best to look at the breakdown of Weyl invariance as seen by $\langle T^\alpha_{\ \alpha}\rangle$. There are three different kinds of contribution that the stress-tensor can receive, related to the three different spacetime fields. Correspondingly, we define three different beta functions,

$$\langle T^\alpha_{\ \alpha}\rangle = -\frac{1}{2\alpha'}\beta_{\mu\nu}(G)\, g^{\alpha\beta}\partial_\alpha X^\mu\, \partial_\beta X^\nu - \frac{i}{2\alpha'}\beta_{\mu\nu}(B)\,\epsilon^{\alpha\beta}\partial_\alpha X^\mu\, \partial_\beta X^\nu - \frac{1}{2}\beta(\Phi)R^{(2)} \tag{7.15}$$

We will not provide the details of the one-loop beta function computations. We merely state the results[8],

$$\begin{aligned}
\beta_{\mu\nu}(G) &= \alpha'\mathcal{R}_{\mu\nu} + 2\alpha'\nabla_\mu\nabla_\nu\Phi - \frac{\alpha'}{4}H_{\mu\lambda\kappa}H_\nu^{\ \lambda\kappa} \\
\beta_{\mu\nu}(B) &= -\frac{\alpha'}{2}\nabla^\lambda H_{\lambda\mu\nu} + \alpha'\nabla^\lambda\Phi\, H_{\lambda\mu\nu} \\
\beta(\Phi) &= -\frac{\alpha'}{2}\nabla^2\Phi + \alpha'\nabla_\mu\Phi\,\nabla^\mu\Phi - \frac{\alpha'}{24}H_{\mu\nu\lambda}H^{\mu\nu\lambda}
\end{aligned}$$

A consistent background of string theory must preserve Weyl invariance, which now requires $\beta_{\mu\nu}(G) = \beta_{\mu\nu}(B) = \beta(\Phi) = 0$.

7.3 The Low-Energy Effective Action

The equations $\beta_{\mu\nu}(G) = \beta_{\mu\nu}(B) = \beta(\Phi) = 0$ can be viewed as the equations of motion for the background in which the string propagates. We now change our perspective: we

[8] The relationship between the beta function and Einstein's equations was first shown by Friedan in his 1980 PhD thesis. A readable account of the full beta functions can be found in the paper by Callan, Friedan, Martinec and Perry "*Strings in Background Fields*", Nucl. Phys. B262 (1985) 593. The full calculational details can be found in TASI lecture notes by Callan and Thorlacius which can be downloaded from the course webpage.

look for a $D = 26$ dimensional spacetime action which reproduces these beta-function equations as the equations of motion. This is the *low-energy effective action* of the bosonic string,

$$S = \frac{1}{2\kappa_0^2}\int d^{26}X\sqrt{-G}\,e^{-2\Phi}\left(\mathcal{R} - \frac{1}{12}H_{\mu\nu\lambda}H^{\mu\nu\lambda} + 4\partial_\mu\Phi\,\partial^\mu\Phi\right) \tag{7.16}$$

where we have taken the liberty of Wick rotating back to Minkowski space for this expression. Here the overall constant involving κ_0 is not fixed by the field equations but can be determined by coupling these equations to a suitable source as described, for example, in 7.4.2. On dimensional grounds alone, it scales as $\kappa_0^2 \sim l_s^{24}$ where $\alpha' = l_s^2$.

Varying the action with respect to the three fields can be shown to yield the beta functions thus,

$$\delta S = \frac{1}{2\kappa_0^2\alpha'}\int d^{26}X\sqrt{-G}\,e^{-2\Phi}\ \Big(\delta G_{\mu\nu}\,\beta^{\mu\nu}(G) - \delta B_{\mu\nu}\,\beta^{\mu\nu}(B)$$
$$-(2\delta\Phi + \frac{1}{2}G^{\mu\nu}\,\delta G_{\mu\nu})(\beta^\lambda{}_\lambda(G) - 4\beta(\Phi))\Big)$$

Equation (7.16) governs the low-energy dynamics of the spacetime fields. The caveat "low-energy" refers to the fact that we only worked with the one-loop beta functions which requires large spacetime curvature.

Something rather remarkable has happened here. We started, long ago, by looking at how a single string moves in flat space. Yet, on grounds of consistency alone, we're led to the action (7.16) governing how spacetime and other fields fluctuate in $D = 26$ dimensions. It feels like the tail just wagged the dog. That tiny string is seriously high-maintenance: its requirements are so stringent that they govern the way the whole universe moves.

You may also have noticed that we now have two different methods to compute the scattering of gravitons in string theory. The first is in terms of scattering amplitudes that we discussed in Section 6. The second is by looking at the dynamics encoded in the low-energy effective action (7.16). Consistency requires that these two approaches agree. They do.

7.3.1 String Frame and Einstein Frame

The action (7.16) isn't quite of the familiar Einstein-Hilbert form because of that strange factor of $e^{-2\Phi}$ that's sitting out front. This factor simply reflects the fact that the action has been computed at tree level in string perturbation theory and, as we saw in Section 6, such terms typically scale as $1/g_s^2$.

It's also worth pointing out that the kinetic terms for Φ in (7.16) seem to have the wrong sign. However, it's not clear that we should be worried about this because, again, the factor of $e^{-2\Phi}$ sits out front meaning that the kinetic terms are not canonically normalized anyway.

To put the action in more familiar form, we can make a field redefinition. Firstly, it's useful to distinguish between the constant part of the dilaton, Φ_0, and the part that varies which we call $\tilde{\Phi}$. We defined the constant part in (7.13); it is related to the string coupling constant. The varying part is simply given by

$$\tilde{\Phi} = \Phi - \Phi_0 \tag{7.17}$$

In D dimensions, we define a new metric $\tilde{G}_{\mu\nu}$ as a combination of the old metric and the dilaton,

$$\tilde{G}_{\mu\nu}(X) = e^{-4\tilde{\Phi}/(D-2)}\, G_{\mu\nu}(X) \tag{7.18}$$

Note that this isn't to be thought of as a coordinate transformation or symmetry of the action. It's merely a relabeling, a mixing-up, of the fields in the theory. We could make such redefinitions in any field theory. Typically, we choose not to because the fields already have canonical kinetic terms. The point of the transformation (7.18) is to get the fields in (7.16) to have canonical kinetic terms as well.

The new metric (7.18) is related to the old by a conformal rescaling. One can check that two metrics related by a general conformal transformation $\tilde{G}_{\mu\nu} = e^{2\omega}G_{\mu\nu}$, have Ricci scalars related by

$$\tilde{\mathcal{R}} = e^{-2\omega}\left(\mathcal{R} - 2(D-1)\nabla^2\omega - (D-2)(D-1)\partial_\mu\omega\,\partial^\mu\omega\right)$$

(We used a particular version of this earlier in the course when considering $D = 2$ conformal transformations). With the choice $\omega = -2\tilde{\Phi}/(D-2)$ in (7.18), and restricting back to $D = 26$, the action (7.16) becomes

$$S = \frac{1}{2\kappa^2}\int d^{26}X\,\sqrt{-\tilde{G}}\left(\tilde{\mathcal{R}} - \frac{1}{12}e^{-\tilde{\Phi}/3}H_{\mu\nu\lambda}H^{\mu\nu\lambda} - \frac{1}{6}\partial_\mu\tilde{\Phi}\partial^\mu\tilde{\Phi}\right) \tag{7.19}$$

The kinetic terms for $\tilde{\Phi}$ are now canonical and come with the right sign. Notice that there is no potential term for the dilaton and therefore nothing that dynamically sets its expectation value in the bosonic string. However, there do exist backgrounds of the superstring in which a potential for the dilaton develops, fixing the string coupling constant.

The gravitational part of the action takes the standard Einstein-Hilbert form. The gravitational coupling is given by

$$\kappa^2 = \kappa_0^2\, e^{2\Phi_0} \sim l_s^{24} g_s^2 \tag{7.20}$$

The coefficient in front of Einstein-Hilbert term is usually identified with Newton's constant

$$8\pi G_N = \kappa^2$$

Note, however, that this is Newton's constant in $D = 26$ dimensions: it will differ from Newton's constant measured in a four-dimensional world. From Newton's constant, we define the $D = 26$ Planck length $8\pi G_N = l_p^{24}$ and Planck mass $M_p = l_p^{-1}$. (With the factor of 8π sitting there, this is usually called the reduced Planck mass). Comparing to (7.20), we see that weak string coupling, $g_s \ll 1$, provides a parameteric separation between the Planck scale and the string scale,

$$g_s \ll 1 \quad \Rightarrow \quad l_p \ll l_s$$

Often the mysteries of gravitational physics are associated with the length scale l_p. We understand string theory best when $g_s \ll 1$ where much of stringy physics occurs at $l_s \gg l_p$ and can be disentangled from strong coupling effects in gravity.

The original metric $G_{\mu\nu}$ is usually called the *string metric* or *sigma-model metric*. It is the metric that strings see, as reflected in the action (7.1). In contrast, $\tilde{G}_{\mu\nu}$ is called the *Einstein metric*. Of course, the two actions (7.16) and (7.19) describe the same physics: we have simply chosen to package the fields in a different way in each. The choice of metric — $G_{\mu\nu}$ or $\tilde{G}_{\mu\nu}$ — is usually referred to as a choice of *frame*: string frame, or Einstein frame.

The possibility of defining two metrics really arises because we have a massless scalar field Φ in the game. Whenever such a field exists, there's nothing to stop us measuring distances in different ways by including Φ in our ruler. Said another way, massless scalar fields give rise to long range attractive forces which can mix with gravitational forces and violate the principle of equivalence. Ultimately, if we want to connect to Nature, we need to find a way to make Φ massive. Such mechanisms exist in the context of the superstring.

7.3.2 Corrections to Einstein's Equations

Now that we know how Einstein's equations arise from string theory, we can start to try to understand new physics. For example, what are the quantum corrections to Einstein's equations?

On general grounds, we expect these corrections to kick in when the curvature r_c of spacetime becomes comparable to the string length scale $\sqrt{\alpha'}$. But that dovetails very nicely with the discussion above where we saw that the perturbative expansion parameter for the non-linear sigma model is α'/r_c^2. Computing the next loop correction to the beta function will result in corrections to Einstein's equations!

If we ignore H and Φ , the 2-loop sigma-model beta function can be easily computed and results in the α' correction to Einstein's equations:

$$\beta_{\mu\nu} = \alpha' \mathcal{R}_{\mu\nu} + \frac{1}{2}\alpha'^2 \mathcal{R}_{\mu\lambda\rho\sigma}\mathcal{R}_{\nu}{}^{\lambda\rho\sigma} + \ldots = 0$$

Such two loop corrections also appear in the heterotic superstring. However, they are absent for the type II string theories, with the first corrections appearing at 4-loops from the perspective of the sigma-model.

String Loop Corrections

Perturbative string theory has an α' expansion and g_s expansion. We still have to discuss the latter. Here an interesting subtlety arises. The sigma-model beta functions arise from regulating the UV divergences of the worldsheet. Yet the g_s expansion cares only about the topology of the string. How can the UV divergences care about the global nature of the worldsheet. Or, equivalently, how can the higher-loop corrections to the beta-functions give anything interesting?

The resolution to this puzzle is to remember that, when computing higher g_s corrections, we have to integrate over the moduli space of Riemann surfaces. But this moduli space will include some tricky points where the Riemann surface degenerates. (For example, one cycle of the torus may pinch off). At these points, the UV divergences suddenly do care about global topology and this results in the g_s corrections to the low-energy effective action.

7.3.3 Nodding Once More to the Superstring

In section 2.5, we described the massless bosonic content for the four superstring theories: Heterotic $SO(32)$, Heterotic $E_8 \times E_8$, Type IIA and Type IIB. Each of them contains the fields $G_{\mu\nu}$, $B_{\mu\nu}$ and Φ that appear in the bosonic string, together with a collection of further massless fields. For each, the low-energy effective action describes the dynamics of these fields in $D = 10$ dimensional spacetime. It naturally splits up into three pieces,

$$S_{\text{superstring}} = S_1 + S_2 + S_{\text{fermi}}$$

Here $S_{\rm fermi}$ describes the interactions of the spacetime fermions. We won't describe these here. But we will briefly describe the low-energy bosonic action $S_1 + S_2$ for each of these four superstring theories.

S_1 is essentially the same for all theories and is given by the action we found for the bosonic string in string frame (7.16). We'll start to use form notation and denote $H_{\mu\nu\lambda}$ simply as H_3, where the subscript tells us the degree of the form. Then the action reads

$$S_1 = \frac{1}{2\kappa_0^2}\int d^{10}X\sqrt{-G}\,e^{-2\Phi}\left(\mathcal{R} - \frac{1}{2}|\tilde{H}_3|^2 + 4\partial_\mu\Phi\,\partial^\mu\Phi\right) \tag{7.21}$$

There is one small difference, which is that the field $\tilde{H}_3$ that appears here for the heterotic string is not quite the same as the original H_3; we'll explain this further shortly.

The second part of the action, S_2, describes the dynamics of the extra fields which are specific to each different theory. We'll now go through the four theories in turn, explaining S_2 in each case.

- **Type IIA**: For this theory, $\tilde{H}_3$ appearing in (7.21) is $H_3 = dB_2$, just as we saw in the bosonic string. In Section 2.5, we described the extra bosonic fields of the Type IIA theory: they consist of a 1-form C_1 and a 3-form C_3. The dynamics of these fields is governed by the so-called Ramond-Ramond part of the action and is written in form notation as,

$$S_2 = -\frac{1}{4\kappa_0^2}\int d^{10}X\,\left[\sqrt{-G}\,\left(|F_2|^2 + |\tilde{F}_4|^2\right) + B_2\wedge F_4\wedge F_4\right]$$

  Here the field strengths are given by $F_2 = dC_1$ and $F_4 = dC_3$, while the object that appears in the kinetic terms is $\tilde{F}_4 = F_4 - C_1\wedge H_3$. Notice that the final term in the action does not depend on the metric: it is referred to as a *Chern-Simons* term.

- **Type IIB**: Again, $\tilde{H}_3 \equiv H_3$. The extra bosonic fields are now a scalar C_0, a 2-form C_2 and a 4-form C_4. Their action is given by

$$S_2 = -\frac{1}{4\kappa_0^2}\int d^{10}X\,\left[\sqrt{-G}\,\left(|F_1|^2 + |\tilde{F}_3|^2 + \frac{1}{2}|\tilde{F}_5|^2\right) + C_4\wedge H_3\wedge F_3\right]$$

  where $F_1 = dC_0$, $F_3 = dC_2$ and $F_5 = dC_4$. Once again, the kinetic terms involve more complicated combinations of the forms: they are $\tilde{F}_3 = F_3 - C_0\wedge H_3$ and

$\tilde{F}_5 = F_5 - \frac{1}{2}C_2 \wedge H_3 + \frac{1}{2}B_2 \wedge F_3$. However, for type IIB string theory, there is one extra requirement on these fields that cannot be implemented in any simple way in terms of a Lagrangian: $\tilde{F}_5$ must be self-dual

$$\tilde{F}_5 = {}^\star\tilde{F}_5$$

Strictly speaking, one should say that the low-energy dynamics of type IIB theory is governed by the equations of motion that we get from the action, supplemented with this self-duality requirement.

- **Heterotic**: Both heterotic theories have just one further massless bosonic ingredient: a non-Abelian gauge field strength F_2, with gauge group $SO(32)$ or $E_8 \times E_8$. The dynamics of this field is simply the Yang-Mills action in ten dimensions,

$$S_2 = \frac{\alpha'}{8\kappa_0^2} \int d^{10}X \, \sqrt{-G} \, \mathrm{Tr}\, |F_2|^2$$

The one remaining subtlety is to explain what $\tilde{H}_3$ means in (7.21): it is defined as $\tilde{H}_3 = dB_2 - \alpha'\omega_3/4$ where ω_3 is the Chern-Simons three form constructed from the non-Abelian gauge field A_1

$$\omega_3 = \mathrm{Tr}\left(A_1 \wedge dA_1 + \frac{2}{3}A_1 \wedge A_1 \wedge A_1\right)$$

The presence of this strange looking combination of forms sitting in the kinetic terms is tied up with one of the most intricate and interesting aspects of the heterotic string, known as anomaly cancelation.

The actions that we have written down here probably look a little arbitrary. But they have very important properties. In particular, the full action $S_{\text{superstring}}$ of each of the Type II theories is invariant under $\mathcal{N} = 2$ spacetime supersymmetry. (That means 32 supercharges). They are the unique actions with this property. Similarly, the heterotic superstring actions are invariant under $\mathcal{N} = 1$ supersymmetry and, crucially, do not suffer from anomalies. The second book by Polchinski is a good place to start learning more about these ideas.

7.4 Some Simple Solutions

The spacetime equations of motion,

$$\beta_{\mu\nu}(G) = \beta_{\mu\nu}(B) = \beta(\Phi) = 0$$

have many solutions. This is part of the story of vacuum selection in string theory. What solution, if any, describes the world we see around us? Do we expect this putative solution to have other special properties, or is it just a random choice from the many possibilities? The answer is that we don't really know, but there is currently no known principle which uniquely selects a solution which looks like our world — with the gauge groups, matter content and values of fundamental constants that we observe — from the many other possibilities. Of course, these questions should really be asked in the context of the superstring where a greater understanding of various non-perturbative effects such as D-branes and fluxes leads to an even greater array of possible solutions.

Here we won't discuss these problems. Instead, we'll just discuss a few simple solutions that are well known. The first plays a role when trying to make contact with the real world, while the value of the others lies mostly in trying to better understand the structure of string theory.

7.4.1 Compactifications

We've seen that the bosonic string likes to live in $D = 26$ dimensions. But we don't. Or, more precisely, we only observe three macroscopically large spatial dimensions. How do we reconcile these statements?

Since string theory is a theory of gravity, there's nothing to stop extra dimensions of the universe from curling up. Indeed, under certain circumstances, this may be required dynamically. Here we exhibit some simple solutions of the low-energy effective action which have this property. We set $H_{\mu\nu\rho} = 0$ and Φ to a constant. Then we are simply searching for Ricci flat backgrounds obeying $\mathcal{R}_{\mu\nu} = 0$. There are solutions where the metric is a direct product of metrics on the space

$$\mathbf{R}^{1,3} \times \mathbf{X} \tag{7.22}$$

where $\mathbf{X}$ is a compact 22-dimensional Ricci-flat manifold.

The simplest such manifold is just $\mathbf{X} = \mathbf{T}^{22}$, the torus endowed with a flat metric. But there are a whole host of other possibilities. Compact, complex manifolds that admit such Ricci-flat metrics are called *Calabi-Yau* manifolds. (Strictly speaking, Calabi-Yau manifolds are complex manifolds with vanishing first Chern class. Yau's theorem guarantees the existence of a unique Ricci flat metric on these spaces).

The idea that there may be extra, compact directions in the universe was considered long before string theory and goes by the name of *Kaluza-Klein compactification*. If the characteristic length scale L of the space $\mathbf{X}$ is small enough then the presence of these extra dimensions would not have been observed in experiment. The standard model of particle physics has been accurately tested to energies of a TeV or so, meaning that if the standard model particles can roam around $\mathbf{X}$, then the length scale must be $L \lesssim (\mathrm{TeV})^{-1} \sim 10^{-16}$ cm.

However, one can cook up scenarios in which the standard model is stuck somewhere in these extra dimensions (for example, it may be localized on a D-brane). Under these circumstances, the constraints become much weaker because we would rely on gravitational experiments to detect extra dimensions. Present bounds require only $L \lesssim 10^{-5}$ cm.

Consider the Einstein-Hilbert term in the low-energy effective action. If we are interested only in the dynamics of the 4d metric on $\mathbf{R}^{1,3}$, this is given by

$$S_{EH} = \frac{1}{2\kappa^2}\int d^{26}X\sqrt{-\tilde{G}}\,\tilde{\mathcal{R}} = \frac{\mathrm{Vol}(\mathbf{X})}{2\kappa^2}\int d^4X\sqrt{-G_{4d}}\,\mathcal{R}_{4d}$$

(There are various moduli of the internal manifold $\mathbf{X}$ that are being neglected here). From this equation, we learn that effective 4d Newton constant is given in terms of 26d Newton constant by,

$$8\pi G_N^{4d} = \frac{\kappa^2}{\mathrm{Vol}(\mathbf{X})}$$

Rewriting this in terms of the 4d Planck scale, we have $l_p^{(4d)} \sim g_s l_s^{12}/\sqrt{\mathrm{Vol}(\mathbf{X})}$. To trust this whole analysis, we require $g_s \ll 1$ and all length scales of the internal space to be bigger than l_s. This ensures that $l_p^{(4d)} < l_s$. Although the 4d Planck length is ludicrously small, $l_p^{(4d)} \sim 10^{-33}$ cm, it may be that we don't have to probe to this distance to uncover UV gravitational physics. The back-of-the-envelope calculation above shows that the string scale l_s could be much larger, enhanced by the volume of extra dimensions.

7.4.2 The String Itself

We've seen that quantizing small loops of string gives rise to the graviton and $B_{\mu\nu}$ field. Yet, from the sigma model action (7.12), we also know that the string is charged under the $B_{\mu\nu}$. Moreover, the string has tension, which ensures that it also acts as a source for the metric $G_{\mu\nu}$. So what does the back-reaction of the string look like? Or, said another way: what is the sigma-model describing a string moving in the background of another string?

Consider an infinite, static, straight string stretched in the X^1 direction. We can solve for the background fields by coupling the equations of motion to a delta-function string source. This is the same kind of calculation that we're used to in electromagnetism. The resulting spacetime fields are given by

$$ds^2 = f(r)^{-1}\,(-dt^2 + dX_1^2) + \textstyle\sum_{i=2}^{25} dX_i^2$$
$$B = (f(r)^{-1} - 1)\,dt \wedge dX_1 \quad , \quad e^{2\Phi} = f(r)^{-1} \tag{7.23}$$

The function $f(r)$ depends only on the transverse direction $r^2 = \sum_{i=2}^{25} X_i^2$ and is given by

$$f(r) = 1 + \frac{g_s^2 N l_s^{22}}{r^{22}}$$

Here N is some constant which we will shortly demonstrate counts the number of strings which source the background. The string length scale in the solutions is $l_s = \sqrt{\alpha'}$. The function $f(r)$ has the property that it is harmonic in the space transverse to the string, meaning that it satisfies $\nabla^2_{\mathbf{R}^{24}} f(r) = 0$ except at $r = 0$.

Let's compute the B-field charge of this solution. We do exactly what we do in electromagnetism: we integrate the total flux through a sphere which surrounds the object. The string lies along the X^1 direction so the transverse space is $\mathbf{R}^{24}$. We can consider a sphere $\mathbf{S}^{23}$ at the boundary of this transverse space. We should be integrating the flux over this sphere. But what is the expression for the flux?

To see what we should do, let's look at the action for $H_{\mu\nu\rho}$ in the presence of a string source. We will use form notation since this is much cleaner and refer to $H_{\mu\nu\rho}$ simply as H_3. Schematically, the action takes the form

$$\frac{1}{g_s^2}\int_{\mathbf{R}^{26}} H_3 \wedge {}^\star H_3 + \int_{\mathbf{R}^2} B_2 = \frac{1}{g_s^2}\int_{\mathbf{R}^{26}} H_3 \wedge {}^\star H_3 + g_s^2 B_2 \wedge \delta(\omega)$$

Here $\delta(\omega)$ is a delta-function source with support on the 2d worldsheet of the string. The equation of motion is

$$d^\star H_3 \sim g_s^2 \delta(\omega)$$

From this we learn that to compute the charge of a single string we need to integrate

$$\frac{1}{g_s^2}\int_{\mathbf{S}^{23}} {}^\star H_3 = 1$$

After these general comments, we now return to our solution (7.23). The above discussion was schematic and no attention was paid to factors of 2 and π. Keeping in this spirit, the flux of the solution (7.23) can be checked to be

$$\frac{1}{g_s^2}\int_{\mathbf{S}^{23}} {}^\star H_3 = N$$

This is telling us that the solution (7.23) describes the background sourced by N coincident, parallel fundamental strings. Another way to check this is to compute the ADM mass per unit length of the solution: it is $NT \sim N/\alpha'$ as expected.

Note as far as the low-energy effective action is concerned, there is nothing that insists $N \in \mathbf{Z}$. This is analogous to the statement that nothing in classical Maxwell theory requires e to be quantized. However, in string theory, as in QED, we know the underlying sources of the microscopic theory and N must indeed take integer values.

Finally, notice that as $r \to 0$, the solution becomes singular. It is not to be trusted in this regime where higher order α' corrections become important.

7.4.3 Magnetic Branes

We've already seen that string theory is not just a theory of strings; there are also D-branes, defined as surfaces on which strings can end. We'll have much more to say about D-branes in Section 7.5. Here, we will consider a third kind of object that exists in string theory. It is again a brane – meaning that it is extended in some number of spacetime directions — but it is not a D-brane because the open string cannot end there. In these lectures we will call it the *magnetic brane*.

Electric and Magnetic Charges

You're probably not used to talking about magnetically charged objects in electromagnetism. Indeed, in undergraduate courses we usually don't get much further than pointing out that $\nabla \cdot B = 0$ does not allow point-like magnetic charges. However, in the context of quantum field theory, much of the interesting behaviour often boils down to understanding how magnetic charges behave. And the same is true of string theory. Because this may be unfamiliar, let's take a minute to discuss the basics.

In electromagnetism in $d = 3 + 1$ dimensions, we measure electric charge q by integrating the electric field $\vec{E}$ over a sphere $\mathbf{S}^2$ that surrounds the particle,

$$q = \int_{\mathbf{S}^2} \vec{E} \cdot d\vec{S} = \int_{\mathbf{S}^2} {}^\star F_2 \tag{7.24}$$

In the second equality we have introduced the notation of differential forms that we also used in the previous example to discuss the string solutions.

Suppose now that a particle carries magnetic charge g. This can be measured by integrating the magnetic field $\vec{B}$ over the same sphere. This means

$$g = \int_{\mathbf{S}^2} \vec{B} \cdot d\vec{S} = \int_{\mathbf{S}^2} F_2 \tag{7.25}$$

In $d = 3{+}1$ dimensions, both electrically and magnetically charged objects are particles. But this is not always true in any dimension! The reason that it holds in $4d$ is because both the field strength F_2 and the dual field strength ${}^\star F_2$ are 2-forms. Clearly, this is rather special to four dimensions.

In general, suppose that we have a p-brane that is electrically charged under a suitable gauge field. As we discussed in Section 7.2.1, a $(p+1)$-dimensional object naturally couples to a $(p+1)$-form gauge potential C_{p+1} through,

$$\mu \int_W C_{p+1}$$

where μ is the charge of the object, while W is the worldvolume of the brane. The $(p+1)$-form gauge potential has a $(p+2)$-form field strength

$$G_{p+2} = dC_{p+1}$$

To measure the electric charge of the p-brane, we need to integrate the field strength over a sphere that completely surrounds the object. A p-brane in D-dimensions has a transverse space $\mathbf{R}^{D-p-1}$. We can integrate the flux over the sphere at infinity, which is $\mathbf{S}^{D-p-2}$. And, indeed, the counting works out nicely because, in D dimensions, the dual field strength is a $(D-p-2)$-form, ${}^\star G_{p+2} = \tilde{G}_{D-p-2}$, which we can happily integrate over the sphere to find the charge sitting inside,

$$q = \int_{\mathbf{S}^{D-p-2}} {}^\star G_{p+2}$$

This equation is the generalized version of (7.24)

Now let's think about magnetic charges. The generalized version of (7.25) suggest that we should compute the magnetic charge by integrating G_{p+2} over a sphere $\mathbf{S}^{p+2}$. What kind of object sits inside this sphere to emit the magnetic charge? Doing the sums backwards, we see that it should be a $(D-p-4)$-brane.

We can write down the coupling between the $(D-p-4)$-brane and the field strength. To do so, we first need to introduce the magnetic gauge potential defined by

$$ {}^\star G_{p+2} = \tilde{G}_{D-p-2} = d\tilde{C}_{D-p-3} \tag{7.26}$$

We can then add the magnetic coupling to the worldvolume $\tilde{W}$ of a $(D-p-4)$-brane simply by writing

$$\tilde{\mu} \int_{\tilde{W}} \tilde{C}_{D-p-3}$$

where $\tilde{\mu}$ is the magnetic charge. Note that it's typically not possible to write down a Lagrangian that includes both magnetically charged object and electrically charged objects at the same time. This would need us to include both C_{p+1} and $\tilde{C}_{D-p-3}$ in the Lagrangian, but these are not independent fields: they're related by the rather complicated differential equations (7.26).

The Magnetic Brane in Bosonic String Theory

After these generalities, let's see what it means for the bosonic string. The fundamental string is a 1-brane and, as we saw in Section 7.2.1, carries electric charge under the 2-form B. The appropriate object carrying magnetic charge under B is therefore a $(D-p-4)=(26-1-4)=21$-brane.

To stress a point: neither the fundamental string, nor the magnetic 21-brane are D-branes. They are not surfaces where strings can end. We are calling them *branes* only because they are extended objects.

The magnetic 21-brane of the bosonic string can be found as a solution to the low-energy equations of motion. The solution can be written in terms of the dual potential $\tilde{B}_{22}$ such that $d\tilde{B}_{22} = {}^\star dB_2$. It is

$$\begin{aligned} ds^2 &= \left(-dt^2 + \sum_{i=1}^{21} dX_i^2\right) + h(r)\left(dX_{22}^2 + \ldots + dX_{25}^2\right) \\ \tilde{B}_{22} &= (1-h(r)^{-2})\, dt \wedge dX_1 \wedge \ldots \wedge dX_{21} \\ e^{2\Phi} &= h(r) \end{aligned} \tag{7.27}$$

The function $h(r)$ depends only on the radial direction in $\mathbf{R}^4$ transverse to the brane: $r^2 = \sum_{i=22}^{25} X_i^2$. It is a harmonic function in $\mathbf{R}^4$, given by

$$h(r) = 1 + \frac{N l_s^2}{r^2}$$

The role of this function in the metric (7.27) is to warp the transverse $\mathbf{R}^4$ directions. Distances get larger as you approach the brane and the origin, $r=0$, is at infinite distance.

It can be checked that the solution carried N units of magnetic charge and has tension

$$T \sim \frac{N}{l_s^{22}} \frac{1}{g_s^2}$$

Let's summarize how the tension of different objects scale in string theory. The powers of $\alpha' = l_s^2$ are entirely fixed on dimensional grounds. (Recall that the tension is mass per spatial volume, so the tension of a p-brane has $[T_p] = p+1$). More interesting is the dependence on the string coupling g_s. The tension of the fundamental string does not depend on g_s, while the magnetic brane scales as $1/g_s^2$. This kind of $1/g^2$ behaviour is typical of solitons in field theories. The D-branes sit between the two: their tension scales as $1/g_s$. Objects with this behaviour are somewhat rarer (although not unheard of) in field theory.

In the perturbative limit, $g_s \to 0$, both D-branes and magnetic branes are heavy. The coupling of an object with tension T to gravity is governed by $T\kappa^2$ where the gravitational coupling scales as $\kappa \sim g_s^2$ (7.20). This means that in the weak coupling limit, the gravitational backreaction of the string and D-branes can be neglected. However, the coupling of the magnetic brane to gravity is always of order one.

The Magnetic Brane in Superstring Theory

Superstring theories also have a brane magnetically charged under B. It is a $(D-p-4) = (10-1-4) = 5$-brane and is usually referred to as the NS5-brane. The solution in the transverse $\mathbf{R}^4$ again takes the form (7.27).

The NS5-brane exists in both type II and heterotic string. In many ways it is more mysterious than D-branes and its low-energy effective dynamics is still poorly understood. It is closely related to the 5-brane of M-theory.

7.4.4 Moving Away from the Critical Dimension

The beta function equations provide a new view on the critical dimension $D = 26$ of the bosonic string. To see this, let's look more closely at the dilaton beta function $\beta(\Phi)$ defined in (7.15): it takes the same form as the Weyl anomaly that we discussed back in Section 4.4.2. This means that if we consider a string propagating in $D \neq 26$ then the Weyl anomaly simply arises as the leading order term in the dilaton beta function. So let's relax the requirement of the critical dimension. The equations of motion arising from $\beta_{\mu\nu}(G)$ and $\beta_{\mu\nu}(B)$ are unchanged, while the dilaton beta function equation becomes

$$\beta(\Phi) = \frac{D-26}{6} - \frac{\alpha'}{2}\nabla^2\Phi + \alpha'\nabla_\mu\Phi\,\nabla^\mu\Phi - \frac{\alpha'}{24}H_{\mu\nu\lambda}H^{\mu\nu\lambda} = 0 \tag{7.28}$$

The low-energy effective action in string frame picks up an extra term which looks like a run-away potential for Φ,

$$S = \frac{1}{2\kappa_0^2}\int d^DX\sqrt{-G}\,e^{-2\Phi}\left(\mathcal{R} - \frac{1}{12}H_{\mu\nu\lambda}H^{\mu\nu\lambda} + 4\partial_\mu\Phi\,\partial^\mu\Phi - \frac{2(D-26)}{3\alpha'}\right)$$

This sounds quite exciting. Can we really get string theory living in $D = 4$ dimensions so easily? Well, yes and no. Firstly, with this extra potential term, flat D-dimensional Minkowski space no longer solves the equations of motion. This is in agreement with the analysis in Section 2 where we showed that full Lorentz invariance was preserved only in $D = 26$.

Another, technical, problem with solving the string equations of motion this way is that we're playing tree-level term off against a one-loop term. But if tree-level and one-loop terms are comparable, then typically all higher loop contributions will be as well and it is likely that we can't trust our analysis.

The Linear Dilaton CFT

In fact, there is one simple solution to (7.28) which we can trust. It is the solution to

$$\partial_\mu \Phi \, \partial^\mu \Phi = \frac{26 - D}{6\alpha'}$$

Recall that we're working in signature $(-, +, +, \ldots)$, meaning that Φ takes a spacelike profile if $D < 26$ and a timelike profile if $D > 26$,

$$\Phi = \sqrt{\frac{26 - D}{6\alpha'}}\, X^1 \qquad D < 26$$
$$\Phi = \sqrt{\frac{D - 26}{6\alpha'}}\, X^0 \qquad D > 26$$

This gives a dilaton which is linear in one direction. This can be compared to the study of the path integral for non-critical strings that we saw in 5.3.2. There are two ways of seeing the same physics.

The reason that we can trust this solution is that there is an exact CFT underlying it which we can analyze to all orders in α'. It's called, for obvious reasons, the *linear dilaton CFT*. Let's now look at this in more detail.

Firstly, consider the worldsheet action associated to the dilaton coupling. For now we'll consider an arbitrary dilaton profile $\Phi(X)$,

$$S_{\text{dilaton}} = \frac{1}{4\pi} \int d^2\sigma \sqrt{g}\ \Phi(X) R^{(2)} \tag{7.29}$$

Although this term vanishes on a flat worldsheet, it nonetheless changes the stress-energy tensor $T_{\alpha\beta}$ because this is defined as

$$T_{\alpha\beta} = -4\pi \left. \frac{\partial S}{\partial g^{\alpha\beta}} \right|_{g_{\alpha\beta} = \delta_{\alpha\beta}}$$

The variation of (7.29) is straightforward. Indeed, the term is akin to the Einstein-Hilbert term in general relativity but things are simpler in 2d because, for example $R_{\alpha\beta} = \frac{1}{2}\, g_{\alpha\beta} R$. We have

$$\delta(\sqrt{g} g^{\alpha\beta} R_{\alpha\beta}) = \sqrt{g} g^{\alpha\beta}\, \delta R_{\alpha\beta} = \sqrt{g}\, \nabla^\alpha v_\alpha$$

where

$$v_\alpha = \nabla^\beta \delta g_{\alpha\beta} - g^{\gamma\delta} \nabla_\alpha \delta g_{\gamma\delta}$$

Using this, the variation of the dilaton term in the action is given by

$$\delta S_{\text{dilaton}} = \frac{1}{4\pi} \int d^2\sigma \sqrt{g}\, \left(\nabla^\alpha \nabla^\beta \Phi - \nabla^2 \Phi\, g^{\alpha\beta}\right)\, \delta g_{\alpha\beta}$$

which, restricting to flat space $g_{\alpha\beta} = \delta_{\alpha\beta}$, finally gives us the stress-energy tensor of a theory with dilaton coupling

$$T^{\text{dilaton}}_{\alpha\beta} = -\partial_\alpha \partial_\beta \Phi + \partial^2 \Phi\, \delta_{\alpha\beta}$$

Note that this stress tensor is not traceless. This is to be expected because, as we described above, the dilaton coupling is not Weyl invariant at tree-level. In complex coordinates, the stress tensor is

$$T^{\text{dilaton}} = -\partial^2 \Phi \quad , \quad \bar{T}^{\text{dilaton}} = -\bar{\partial}^2 \Phi$$

Linear Dilaton OPE

The stress tensor above holds for any dilaton profile $\Phi(X)$. Let's now restrict to a linear dilaton profile for a single scalar field X,

$$\Phi = QX$$

where Q is some constant. We also include the standard kinetic terms for D scalar fields, of which X is a chosen one, giving the stress tensor

$$T = -\frac{1}{\alpha'} : \partial X\, \partial X : - Q\, \partial^2 X$$

It is a simple matter to compute the TT OPE using the techniques described in Section 4. We find,

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

where the central charge of the theory is given by

$$c = D + 6\alpha' Q^2$$

Note that Q^2 can be positive or negative depending on the whether we have a timelike or spacelike linear dilaton. In this way, we see explicitly how a linear dilaton gradient can absorb central charge.

7.4.5 The Elephant in the Room: The Tachyon

We've been waxing lyrical about the details of solutions to the low-energy effective action, all the while ignoring the most important, relevant field of them all: the tachyon. Since our vacuum is unstable, this is a little like describing all the beautiful pictures we could paint if only that damn paintbrush would balance, unaided, on its tip.

Of course, the main reason for discussing these solutions is that they all carry directly over to the superstring where the tachyon is absent. Nonetheless, it's interesting to ask what happens if the tachyon is turned on. Its vertex operator is simply

$$V_{\text{tachyon}} \sim \int d^2\sigma \sqrt{g}\, e^{ip\cdot X}$$

where $p^2 = 4/\alpha'$. Piecing together a general tachyon profile $V(X)$ from these Fourier modes and exponentiating, results in a potential on the worldsheet of the string

$$S_{\text{potential}} = \int d^2\sigma \sqrt{g}\; \alpha' \, V(X)$$

This is a relevant operator for the worldsheet CFT. Whenever such a relevant operator turns on, we should follow the RG flow to the infra-red until we land on another CFT. The c-theorem tells us that $c_{IR} < c_{UV}$, but in string theory we always require $c = 26$. The deficit, at least initially, is soaked up by the dilaton in the manner described above. The end point of the tachyon RG flow for the bosonic string is not understood. It may be that there is no end point and the bosonic string simply doesn't make sense once the tachyon is turned on. Or perhaps we haven't yet understood the true ground state of the bosonic string.

7.5 D-Branes Revisited: Background Gauge Fields

Understanding the constraints of conformal invariance on the closed string backgrounds led us to Einstein's equations and the low-energy effective action in spacetime. Now we would like to do the same for the open string. We want to understand the restrictions that consistency places on the dynamics of D-branes.

We saw in Section 3 that there are two types of massless modes that arise from the quantization of an open string: scalars, corresponding to the fluctuation of the D-brane, and a $U(1)$ gauge field. We will ignore the scalar fluctuations for now, but will return to them later. We focus initially on the dynamics of a gauge field A_a, $a = 0, \ldots, p$ living on a Dp-brane

The first question that we ask is: how does the end of the string react to a background gauge field? To answer this, we need to look at the vertex operator associated to the photon. It was given in (5.10)

$$V_{\text{photon}} \sim \int_{\partial\mathcal{M}} d\tau\ \zeta_a\, \partial^\tau X^a\, e^{ip\cdot X}$$

which is Weyl invariant and primary only if $p^2 = 0$ and $p^a\zeta_a = 0$. Exponentiating this vertex operator, as described at the beginning of Section 7, gives the coupling of the open string to a general background gauge field $A_a(X)$,

$$S_{\text{end-point}} = \int_{\partial\mathcal{M}} d\tau\ A_a(X)\,\frac{dX^a}{d\tau}$$

But this is a very familiar coupling — we've already mentioned it in (7.9). It is telling us that the end of the string is charged under the background gauge field A_a on the brane.

7.5.1 The Beta Function

We can now perform the same type of beta function calculation that we saw for the closed string[9]. To do this, it's useful to first use conformal invariance to map the open string worldsheet to the Euclidean upper-half plane as we described in Section 4.7. The action describing an open string propagating in flat space, with its ends subject to a background gauge field on the D-brane splits up into two pieces

$$S = S_{\text{Neumann}} + S_{\text{Dirichlet}}$$

where S_{Neumann} describes the fluctuations parallel to the Dp-brane and is given by

$$S_{\text{Neumann}} = \frac{1}{4\pi\alpha'}\int_{\mathcal{M}} d^2\sigma\ \partial^\alpha X^a\, \partial_\alpha X^b\, \delta_{ab} + i\int_{\partial\mathcal{M}} d\tau\ A_a(X)\dot{X}^a \tag{7.30}$$

Here $a, b = 0, \ldots, p$. The extra factor of i arises because we are in Euclidean space. Meanwhile, the fields transverse to the brane have Dirichlet boundary conditions and take range $I = p+1, \ldots, D-1$. Their dynamics is given by

$$S_{\text{Dirichlet}} = \frac{1}{4\pi\alpha'}\int_{\mathcal{M}} d^2\sigma\ \partial^\alpha X^I\, \partial_\alpha X^J\, \delta_{IJ}$$

[9] We'll be fairly explicit here, but if you want to see more details then the best place to look is the original paper by Abouelsaood, Callan, Nappi and Yost, "*Open Strings in Background Gauge Fields*", Nucl. Phys. B280 (1987) 599.

The action $S_{\rm Dirichlet}$ describes free fields and doesn't play any role in the computation of the beta-function. The interesting part is $S_{\rm Neumann}$ which, for non-zero $A_a(X)$, is an interacting quantum field theory with boundary. Our task is to compute the beta function associated to the coupling $A_a(X)$. We use the same kind of technique that we earlier applied to the closed string. We expand the fields $X^a(\sigma)$ as

$$X^a(\sigma) = \bar{x}^a(\sigma) + \sqrt{\alpha'}\, Y^a(\sigma)$$

where $\bar{x}^a(\sigma)$ is taken to be some fixed background which obeys the classical equations of motion,

$$\partial^2 \bar{x}^a = 0$$

(In the analogous calculation for the closed string we chose the special case of $\bar{x}^a$ constant. Here we are more general). However, we also need to impose boundary conditions for this classical solution. In the absence of the gauge field A_a, we require Neumann boundary conditions $\partial_\sigma X^a = 0$ at $\sigma = 0$. However, the presence of the gauge field changes this. Varying the full action (7.30) shows that the relevant boundary condition is supplemented by an extra term,

$$\partial_\sigma \bar{x}^a + 2\pi\alpha' i\, F^{ab}\, \partial_\tau \bar{x}_b = 0 \qquad \text{at } \sigma = 0 \tag{7.31}$$

where the F_{ab} is the field strength

$$F_{ab}(X) = \frac{\partial A_b}{\partial X^a} - \frac{\partial A_a}{\partial X^b} \equiv \partial_a A_b - \partial_b A_a$$

The fields $Y^a(\sigma)$ are the fluctuations which are taken to be small. Again, the presence of $\sqrt{\alpha'}$ in the expansion ensures that Y^a are dimensionless. Expanding the action $S_{\rm Neumann}$ (which we'll just call S from now on) to second order in fluctuations gives,

$$\begin{aligned} S[\bar{x} + \sqrt{\alpha'} Y] = S[\bar{x}] &+ \frac{1}{4\pi}\int_{\mathcal{M}} d^2\sigma\ \partial Y^a\, \partial Y^b \delta_{ab} \\ &+ i\alpha' \int_{\partial\mathcal{M}} d\tau\ \left(\partial_a A_b\, Y^a\, \dot{Y}^b + \frac{1}{2}\partial_a\partial_b A_c\, Y^a\, Y^b\, \dot{\bar{x}}^c\right) + \ldots \end{aligned}$$

where all expressions involving the background gauge fields are now evaluated on the classical solution $\bar{x}$. We can rearrange the boundary terms by splitting the first term up into two halves and integrating one of these pieces by parts,

$$\int d\tau\ (\partial_a A_b) Y^a \dot{Y}^b = \frac{1}{2}\int d\tau\ \partial_a A_b\, Y^a\, \dot{Y}^b - \partial_a A_b\, \dot{Y}^a Y^b - \partial_c\partial_a A_b\, Y^a Y^b \dot{\bar{x}}^c$$

Combining this with the second term means that we can write all interactions in terms of the gauge invariant field strength F_{ab},

$$S[\bar{x}+\sqrt{\alpha'}Y] = S[\bar{x}] + \frac{1}{4\pi}\int_{\mathcal{M}} d^2\sigma\; \partial Y^a\, \partial Y^b \delta_{ab} + \frac{i\alpha'}{2}\int_{\partial\mathcal{M}} d\tau\; \left(F_{ab}\, Y^a \dot{Y}^b + \partial_b F_{ac}\, Y^a Y^b \dot{\bar{x}}^c\right) + \ldots \tag{7.32}$$

where the $+\ldots$ refer to the higher terms in the expansion which come with higher derivatives of F_{ab}, accompanied by powers of α'. We can neglect them for the purposes of computing the one-loop beta function.

The Propagator

This Lagrangian describes our interacting boundary theory to leading order. We can now use this to compute the beta function. Firstly, we should determine where possible divergences arise. The offending term is the last one in (7.32). This will lead to a divergence when the fluctuation fields Y^a are contracted with their propagator

$$\langle Y^a(z,\bar{z}) Y^b(w,\bar{w})\rangle = G^{ab}(z,\bar{z};w,\bar{w})$$

We should be used to these free field Green's functions by now. The propagator satisfies

$$\partial\bar{\partial}\, G^{ab}(z,\bar{z}) = -2\pi\delta^{ab}\delta(z,\bar{z}) \tag{7.33}$$

in the upper half plane. But now there's a subtlety. The Y^a fields need to satisfy a boundary condition at $\mathrm{Im}\, z = 0$ and this should be reflected in the boundary condition for the propagator. We discussed this briefly for Neumann boundary conditions in Section 4.7. But we've also seen that the background field strength shifts the Neumann boundary conditions to (7.31). Correspondingly, the propagator $G(z,\bar{z};w,\bar{w})$ must now satisfy

$$\partial_\sigma G^{ab}(z,\bar{z};w,\bar{w}) + 2\pi\alpha' i\, F^a_{\ c}\, \partial_\tau G^{cb}(z,\bar{z};w,\bar{w}) = 0 \qquad \text{at } \sigma = 0 \tag{7.34}$$

In Section 4.7, we showed how Neumann boundary conditions could be imposed by considering an image charge in the lower half plane. A similar method works here. We extend $G^{ab} \equiv G^{ab}(z,\bar{z};w,\bar{w})$ to the entire complex plane. The solution to (7.33) subject to (7.34) is given by

$$G^{ab} = -\delta^{ab}\ln|z-w| - \frac{1}{2}\left(\frac{1-2\pi\alpha' F}{1+2\pi\alpha' F}\right)^{ab}\ln(z-\bar{w}) - \frac{1}{2}\left(\frac{1+2\pi\alpha' F}{1-2\pi\alpha' F}\right)^{ab}\ln(\bar{z}-w)$$

The Counterterm and Beta Function

Let's now return to the interacting theory (7.32) and see what counterterm is needed to remove the divergence. Since all interactions take place on the boundary, we should evaluate our propagator on the boundary, which means $z = \bar{z}$ and $w = \bar{w}$. In this case, all the logarithms become the same and, in the limit that $z \to w$, gives the leading divergence $\ln |z - w| \to \epsilon^{-1}$. We learn that the UV divergence takes the form,

$$-\frac{1}{\epsilon}\left[\delta^{ab} + \frac{1}{2}\left(\frac{1-2\pi\alpha' F}{1+2\pi\alpha' F}\right)^{ab} + \frac{1}{2}\left(\frac{1+2\pi\alpha' F}{1-2\pi\alpha' F}\right)^{ab}\right] = -\frac{2}{\epsilon}\left(\frac{1}{1-4\pi^2\alpha'^2 F^2}\right)^{ab}$$

It's now easy to determine the necessary counterterm. We simply replace $Y^a Y^b$ in the final term with $\langle Y^a Y^b \rangle$. This yields

$$-\frac{i2\pi\alpha'^2}{\epsilon}\int_{\partial\mathcal{M}} d\tau\ \partial_b F_{ac}\left[\frac{1}{1-4\pi^2\alpha'^2\, F^2}\right]^{ab}\ \dot{x}^c$$

For the open string theory to retain conformal invariance, we need the associated beta function to vanish. This gives us the condition on the field strength F_{ab}: it must satisfy the equation

$$\partial_b F_{ac}\left[\frac{1}{1-4\pi^2\alpha'^2 F^2}\right]^{ab} = 0 \tag{7.35}$$

This is our final equation governing the equations of motion that F_{ab} must satisfy to provide a consistent background for open string propagation.

7.5.2 The Born-Infeld Action

Equation (7.35) probably doesn't look too familiar! Following the path we took for the closed string, we wish to write down an action whose equations of motion coincide with (7.35). The relevant action was actually constructed many decades ago as a non-linear alternative to Maxwell theory: it goes by the name of the *Born-Infeld action*:

$$S = -T_p \int d^{p+1}\xi\ \sqrt{-\det\left(\eta_{ab} + 2\pi\alpha'\, F_{ab}\right)} \tag{7.36}$$

Here ξ are the worldvolume coordinates on the brane and T_p is the tension of the Dp-brane (which, since it multiplies the action, doesn't affect the equations of motion). The gauge potential is to be thought of as a function of the worldvolume coordinates: $A_a = A_a(\xi)$. It actually takes a little work to show that the equations of motion that we derive from this action coincide with the vanishing of the beta function (7.35). Some hints on how to proceed are provided on Example Sheet 4.

For small field strengths, $F_{ab} \ll 1/\alpha'$, the action (7.36) coincides with Maxwell's action. To see this, we need simply expand to get

$$S = -T_p \int d^{p+1}\xi \, \left(1 + \frac{(2\pi\alpha')^2}{4}\, F_{ab}F^{ab} + \ldots\right)$$

The leading order term, quadratic in field strengths, is the Maxwell action. Terms with higher powers of F_{ab} are suppressed by powers of α'.

So, for small field strengths, the dynamics of the gauge field on a D-brane is governed by Maxwell's equations. However, as the electric and magnetic field strengths increase and become of order $1/\alpha'$, non-linear corrections to the dynamics kick in and are captured by the Born-Infeld action.

The Born-Infeld action arises from the one-loop beta function. It is the exact result for constant field strengths. If we want to understand the dynamics of gauge fields with large gradients, ∂F, then we will have determine the higher loop contributions to the beta function.

7.6 The DBI Action

We've understood that the dynamics of gauge fields on the brane is governed by the Born-Infeld action. But what about the fluctuations of the brane itself. We looked at this briefly in Section 3.2 and suggested, on general grounds, that the action should take the Dirac form (3.6). It would be nice to show this directly by considering the beta function equations for the scalar fields ϕ^I on the brane. Turning these on corresponds to considering boundary conditions where the brane is bent. It is indeed possible to compute something along the lines of beta-function equations and to show directly that the fluctuations of the brane are governed by the Dirac action[10].

More generally, one could consider both the dynamics of the gauge field and the fluctuation of the brane. This is governed by a mixture of the Dirac action and the Born-Infeld action which is usually referred to as the *DBI action*,

$$S_{DBI} = -T_p \int d^{p+1}\xi \, \sqrt{-\det(\gamma_{ab} + 2\pi\alpha'\, F_{ab})}$$

As in Section (3.2), γ_{ab} is the pull-back of the the spacetime metric onto the worldvolume,

$$\gamma_{ab} = \frac{\partial X^\mu}{\partial \xi^a}\frac{\partial X^\nu}{\partial \xi^b}\,\eta_{\mu\nu}$$

[10] A readable discussion of this calculation can be found in the original paper by Leigh, *Dirac-Born-Infeld Action from Dirichlet Sigma Model*, Mod. Phys. Lett. A4: 2767 (1989).

The new dynamical fields in this action are the embedding coordinates $X^\mu(\xi)$, with $\mu = 0, \ldots, D-1$. This appears to be D new degrees of freedom while we expect only $D-p-1$ transverse physical degrees of freedom. The resolution to this should be familiar by now: the DBI action enjoys a reparameterization invariance which removes the longitudinal fluctuations of the brane.

We can use this reparameterization invariance to work in static gauge. For an infinite, flat Dp-brane, it is useful to set

$$X^a = \xi^a \qquad a = 0, \ldots, p$$

so that the pull-back metric depends only on the transverse fluctuations X^I,

$$\gamma_{ab} = \eta_{ab} + \frac{\partial X^I}{\partial \xi^a}\frac{\partial X^J}{\partial \xi^b}\,\delta_{IJ}$$

If we are interested in situations with small field strengths F_{ab} and small derivatives $\partial_a X$, then we can expand the DBI action to leading order. We have

$$S = -(2\pi\alpha')^2 T_p \int d^{p+1}\xi\ \left(\frac{1}{4}F_{ab}F^{ab} + \frac{1}{2}\partial_a\phi^I\partial^a\phi^I + \ldots\right)$$

where we have rescaled the positions to define the scalar fields $\phi^I = X^I/2\pi\alpha'$. We have also dropped an overall constant term in the action. This is simply free Maxwell theory coupled to free massless scalar fields ϕ^I. The higher order terms that we have dropped are all suppressed by powers of α'.

7.6.1 Coupling to Closed String Fields

The DBI action describes the low-energy dynamics of a Dp-brane in flat space. We could now ask how the motion of the D-brane is affected if it moves in a background created by closed string modes $G_{\mu\nu}$, $B_{\mu\nu}$ and Φ. Rather than derive this, we'll simply write down the answer and then justify each term in turn. The answer is:

$$S_{DBI} = -T_p \int d^{p+1}\xi\ e^{-\tilde{\Phi}}\sqrt{-\det(\gamma_{ab} + 2\pi\alpha' F_{ab} + B_{ab})}$$

Let's start with the coupling to the background metric $G_{\mu\nu}$. It's actually hidden in the notation in this expression: it appears in the pull-back metric γ_{ab} which is now given by

$$\gamma_{ab} = \frac{\partial X^\mu}{\partial \xi^a}\frac{\partial X^\nu}{\partial \xi^b}\,G_{\mu\nu}$$

It should be clear that this is indeed the natural place for it to sit.

Next up is the dilaton. As in (7.17), we have decomposed the dilaton into a constant piece and a varying piece: $\Phi = \Phi_0 + \tilde{\Phi}$. The constant piece governs the asymptotic string coupling, $g_s = e^{\Phi_0}$, and is implicitly sitting in front of the action because the tension of the D-brane scales as

$$T_p \sim 1/g_s$$

This, then, explains the factor of $e^{-\tilde{\Phi}}$ in front of the action: it simply reunites the varying part of the dilaton with the constant piece. Physically, it's telling us that the tension of the D-brane depends on the local value of the dilaton field, rather than its asymptotic value. If the dilaton varies, the effective string coupling at a point X in spacetime is given by $g_s^{eff} = e^{\Phi(X)} = g_s\, e^{\tilde{\Phi}(X)}$. This, in turn, changes the tension of the D-brane. It can lower its tension by moving to regions with larger g_s^{eff}.

Finally, let's turn to the $B_{\mu\nu}$ field. This is a 2-form in spacetime. The function B_{ab} appearing in the DBI action is the pull-back to the worldvolume

$$B_{ab} = \frac{\partial X^\mu}{\partial \xi^a}\frac{\partial X^\nu}{\partial \xi^b}\, B_{\mu\nu}$$

Its appearance in the DBI action is actually required on grounds of gauge invariance alone. This can be seen by considering an open string, moving in the presence of both a background $B_{\mu\nu}(X)$ in spacetime and a background $A_a(X)$ on the worldvolume of a brane. The relevant terms on the string worldsheet are

$$\frac{1}{4\pi\alpha'}\int_{\mathcal{M}} d^2\sigma\ \epsilon^{\alpha\beta}\partial_\alpha X^\mu\,\partial_\beta X^\nu B_{\mu\nu} + \int_{\partial\mathcal{M}} d\tau\ A_a \dot{X}^a$$

Under a spacetime gauge transformation

$$B_{\mu\nu} \to B_{\mu\nu} + \partial_\mu C_\nu - \partial_\nu C_\mu \tag{7.37}$$

the first term changes by a total derivative. This is fine for a closed string, but it doesn't leave the action invariant for an open string because we pick up the boundary term. Let's quickly look at what we get in more detail. Under the gauge transformation (7.37), we have

$$\begin{aligned} S_B &= \frac{1}{4\pi\alpha'}\int_{\mathcal{M}} d^2\sigma\ \epsilon^{\alpha\beta}\partial_\alpha X^\mu\,\partial_\beta X^\nu B_{\mu\nu} \\ &\longrightarrow\ S_B + \frac{1}{2\pi\alpha'}\int_{\mathcal{M}} d\sigma d\tau\ \epsilon^{\alpha\beta}\partial_\alpha X^\mu\,\partial_\beta X^\nu\,\partial_\mu C_\nu \\ &= S_B + \frac{1}{2\pi\alpha'}\int_{\mathcal{M}} d\sigma d\tau\ \epsilon^{\alpha\beta}\partial_\alpha\ (\partial_\beta X^\nu C_\nu) \\ &= S_B + \frac{1}{2\pi\alpha'}\int_{\partial\mathcal{M}} d\tau \dot{X}^\nu C_\nu = S_B + \frac{1}{2\pi\alpha'}\int_{\partial\mathcal{M}} d\tau \dot{X}^a C_a \end{aligned}$$

where, in the last line, we have replaced the sum over all directions X^ν with the sum over those directions obeying Neumann boundary conditions X^a, since $\dot{X}^I = 0$ at the end-points for any directions with Dirichlet boundary conditions.

The result of this short calculation is to see that the string action is not invariant under (7.37). To restore this spacetime gauge invariance, this boundary contribution must be canceled by an appropriate shift of A_a in the second term,

$$A_a \to A_a - \frac{1}{2\pi\alpha'} C_a \tag{7.38}$$

Note that this is not the usual kind of gauge transformation that we consider in electrodynamics. In particular, the field strength F_{ab} is not invariant. Rather, the gauge invariant combination under (7.37) and (7.38) is

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

This is the reason that this combination must appear in the DBI action. This is also related to an important physical effect. We have already seen that the string in spacetime is charged under $B_{\mu\nu}$. But we've also seen that the end of the string is charged under the gauge field A_a on the D-brane. This means that the open string deposits B charge on the brane, where it is converted into A charge. The fact that the gauge invariant field strength involves a combination of both F_{ab} and B_{ab} is related to this interplay of charges.

7.7 The Yang-Mills Action

Finally, let's consider the case of N coincident D-branes. We discussed this in Section 3.3 where we showed that the massless fields on the brane could be naturally packaged as $N \times N$ Hermitian matrices, with the element of the matrix telling us which brane the end points terminate on. The gauge field then takes the form

$$(A_a)^m{}_n$$

with $a = 0, \ldots, p$ and $m, n = 1, \ldots, N$. Written this way, it looks rather like a $U(N)$ gauge connection. Indeed, this is the correct interpretation. But how do we see this? Why is the gauge field describing a $U(N)$ gauge symmetry rather than, say, $U(1)^{N^2}$?

The quickest way to see that coincident branes give rise to a $U(N)$ gauge symmetry is to recall that the end point of the string is charged under the $U(1)$ gauge field that inhabits the brane it's ending on. Let's illustrate this with the simplest example. Suppose that we have two branes. The diagonal components $(A_a)^1{}_1$ and $(A_a)^2{}_2$ arise from strings which begin and end on the same brane. Each is a $U(1)$ gauge field. What about the off-diagonal terms $(A_a)^1{}_2$ and $(A_a)^2{}_1$? These come from strings stretched between the two branes. They are again massless gauge bosons, but they are charged under the two original $U(1)$ symmetries; they carry charge $(+1, -1)$ and $(-1, +1)$ respectively. But this is precisely the structure of a $U(2)$ gauge theory, with the off-diagonal terms playing a role similar to W-bosons. In fact, the only way to make sense of massless, charged spin 1 particles is through non-Abelian gauge symmetry.

So the massless excitations of N coincident branes are a $U(N)$ gauge field $(A_a)^m{}_n$, together with scalars $(\phi^I)^m{}_n$ which transform in the adjoint representation of the $U(N)$ gauge group. We saw in Section 3 that the diagonal components $(\phi^I)^m{}_m$ have the interpretation of the transverse fluctuations of the m^{th} brane. Can we now write down an action describing the interactions of these fields?

In fact, there are several subtleties in writing down a non-Abelian generalization of the DBI action and such an action is not known (if, indeed, it makes sense at all). However, we can make progress by considering the low-energy limit, corresponding to small field strengths. The field strength in question is now the appropriate non-Abelian expression which, neglecting the matrix indices, reads

$$F_{ab} = \partial_a A_b - \partial_b A_a + i[A_a, A_b]$$

The low-energy action describing the dynamics of N coincident Dp-branes can be shown to be (neglecting an overall constant term),

$$S = -(2\pi\alpha')^2 T_p \int d^{p+1}\xi \;\text{Tr}\left(\frac{1}{4}F_{ab}F^{ab} + \frac{1}{2}\mathcal{D}_a\phi^I\mathcal{D}^a\phi^I - \frac{1}{4}\sum_{I\neq J}[\phi^I, \phi^J]^2\right) \qquad (7.39)$$

We recognize the first term as the $U(N)$ Yang-Mills action. The coefficient in front of the Yang-Mills action is the coupling constant $1/g_{YM}^2$. For a Dp-brane, this is given by $\alpha'^2 T_p$, or

$$g_{YM}^2 \sim l_s^{p-3} g_s$$

The kinetic term for ϕ^I simply reflects the fact that these fields transform in the adjoint representation of the gauge group,

$$\mathcal{D}_a\phi^I = \partial_a\phi^I + i[A_a, \phi^I]$$

We won't derive this action in these lectures: the first two terms basically follow from gauge invariance alone. The potential term is harder to see directly: the quick ways to derive it use T-duality or, in the case of the superstring, supersymmetry.

A flat, infinite Dp-brane breaks the Lorentz group of spacetime to

$$S(1, D-1) \ \to SO(1,p) \times SO(D-p-1) \tag{7.40}$$

This unbroken group descends to the worldvolume of the D-brane where it classifies all low-energy excitations of the D-brane. The $SO(1,p)$ is simply the Lorentz group of the D-brane worldvolume. The $SO(D-p-1)$ is a global symmetry of the D-brane theory, rotating the scalar fields ϕ^I.

The potential term in (7.39) is particularly interesting,

$$V = -\frac{1}{4}\sum_{I\neq J} \text{Tr}\ [\phi^I, \phi^J]^2$$

The potential is positive semi-definite. We can look at the fields that can be turned on at no cost of energy, $V = 0$. This requires that all ϕ^I commute which means that, after a suitable gauge transformation, they take the diagonal form,

$$\phi^I = \begin{pmatrix} \phi_1^I & & \\ & \ddots & \\ & & \phi_N^I \end{pmatrix} \tag{7.41}$$

The diagonal component ϕ_n^I describes the position of the n^{th} brane in transverse space $\mathbf{R}^{D-p-1}$. We still need to get the dimensions right. The scalar fields have dimension $[\phi] = 1$. The relationship to the position in space (which we mentioned before in 3.2) is

$$\vec{X}_n = 2\pi\alpha'\vec{\phi}_n \tag{7.42}$$

where we've swapped to vector notation to replace the I index.

The eigenvalues ϕ_n^I are not quite gauge invariant: there is a residual gauge symmetry — the Weyl group of $U(N)$ — which leaves ϕ^I in the form (7.41) but permutes the entries by S_N, the permutation group of N elements. But this has a very natural interpretation: it is simply telling us that the D-branes are indistinguishable objects.

When all branes are separated, the vacuum expectation value (7.41) breaks the gauge group from $U(N) \to U(1)^N$. The W-bosons gain a mass M_W through the Higgs mechanism. Let's compute this mass. We'll consider a $U(2)$ theory and we'll separate the two D-branes in the direction $X^D \equiv X$. This means that we turn on a vacuum expectation value for $\phi^D = \phi$, which we write as

$$\phi = \begin{pmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{pmatrix} \tag{7.43}$$

The values of ϕ_1 and ϕ_2 are the positions of the first and second brane. Or, more precisely, we need to multiply by the conversion factor $2\pi\alpha'$ as in (7.42) to get the position X_m of the $m = 1^{\rm st}, 2^{\rm nd}$ brane,

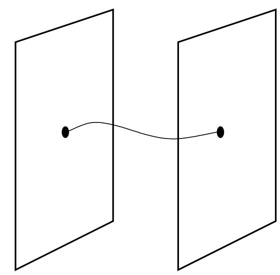


Figure 44:

Let's compute the mass of the W-boson from the Yang-Mills action (7.39). It comes from the covariant derivative terms $\mathcal{D}\phi$. We expand out the gauge field as

$$A_a = \begin{pmatrix} A_a^{11} & W_a \\ W_a^\dagger & A_a^{22} \end{pmatrix}$$

with A^{11} and A^{22} describing the two $U(1)$ gauge fields and W the W-boson. The mass of the W-boson comes from the $[A_a, \phi]$ term inside the covariant derivative which, using the expectation value (7.43), is given by

$$\frac{1}{2}\mathrm{Tr}\,[A_a, \phi]^2 = -(\phi_2 - \phi_1)^2 |W_a|^2$$

This gives us the mass of the W-boson: it is

$$M_W^2 = (\phi_2 - \phi_1)^2 = T^2 |X_2 - X_1|^2$$

where $T = 1/2\pi\alpha'$ is the tension of the string. But this has a very natural interpretation. It is precisely the mass of a string stretched between the two D-branes as shown in the figure above. We see that D-branes provide a natural geometric interpretation of the Higgs mechanism using adjoint scalars.

Notice that when branes are well separated, and the strings that stretch between them are heavy, their positions are described by the diagonal elements of the matrix given in (7.41). However, as the branes come closer together, these stretched strings become light and are important for the dynamics of the branes. Now the positions of the branes should be described by the full $N \times N$ matrices, including the off-diagonal elements. In this manner, D-branes begin to see space as something non-commutative at short distances.

In general, we can consider N D-branes located at positions $\vec{X}_m$, $m = 1, \ldots, N$ in transverse space. The string stretched between the m^{th} and n^{th} brane has mass

$$M_W = |\vec{\phi}_n - \vec{\phi}_m| = T|\vec{X}_n - \vec{X}_m|$$

which again coincides with the mass of the appropriate W-boson computed using (7.39).

7.7.1 D-Branes in Type II Superstring Theories

As we mentioned previously, D-branes are ingredients of the Type II superstring theories. Type IIA has Dp-branes with p even, while Type IIB is home to Dp-branes with p odd. The D-branes have a very important property in these theories: they preserve half the supersymmetries.

Let's take a moment to explain what this means. We'll start by returning to the Lorentz group $SO(1, D-1)$ now, of course, with $D = 10$. We've already seen that an infinite, flat Dp-brane is not invariant under the full Lorentz group, but only the subgroup (7.40). If we act with either $SO(1, p)$ or $SO(D-p-1)$ then the D-brane solution remains invariant. We say that these symmetries are preserved by the solution.

However, the role of the preserved symmetries doesn't stop there. The next step is to consider small excitations of the D-brane. These must fit into representations of the preserved symmetry group (7.40). This ensures that the low-energy dynamics of the D-brane must be governed by a theory which is invariant under (7.40) and we have indeed seen that the Lagrangian (7.39) has $SO(1, p)$ as a Lorentz group and $SO(D-p-1)$ as a global symmetry group which rotates the scalar fields.

Now let's return to supersymmetry. The Type II string theories enjoy a lot of supersymmetry: 32 supercharges in total. The infinite, flat D-branes are invariant under half of these; if we act with one half of the supersymmetry generators, the D-brane solutions don't change. Objects that have this property are often referred to as *BPS* states. Just as with the Lorentz group, these unbroken symmetries descend to the worldvolume of the D-brane. This means that the low-energy dynamics of the D-branes is described by a theory which is itself invariant under 16 supersymmetries.

There is a unique class of theories with 16 supersymmetries and a non-Abelian gauge field and matter in the adjoint representation. This class is known as maximally supersymmetric Yang-Mills theory and the bosonic part of the action is given by (7.39). Supersymmetry is realized only after the addition of fermionic fields which also live on the brane. These theories describe the low-energy dynamics of multiple D-branes.

As an illustrative example, consider D3-branes in the Type IIB theory. The theory describing N D-branes is $U(N)$ Yang-Mills with 16 supercharges, usually referred to as $U(N)$ $\mathcal{N} = 4$ super-Yang-Mills. The bosonic part of the action is given by (7.39), where there are $D - p - 1 = 6$ scalar fields ϕ^I in the adjoint representation of the gauge group. These are augmented with four Weyl fermions, also in the adjoint representation.

8. Compactification and T-Duality

In this section, we will consider the simplest compactification of the bosonic string: a background spacetime of the form

$$\mathbf{R}^{1,24} \times \mathbf{S}^1 \tag{8.1}$$

The circle is taken to have radius R, so that the coordinate on $\mathbf{S}^1$ has periodicity

$$X^{25} \equiv X^{25} + 2\pi R$$

We will initially be interested in the physics at length scales $\gg R$ where motion on the $\mathbf{S}^1$ can be ignored. Our goal is to understand what physics looks like to an observer living in the non-compact $\mathbf{R}^{1,24}$ Minkowski space. This general idea goes by the name of *Kaluza-Klein compactification*. We will view this compactification in two ways: firstly from the perspective of the spacetime low-energy effective action and secondly from the perspective of the string worldsheet.

8.1 The View from Spacetime

Let's start with the low-energy effective action. Looking at length scales $\gg R$ means that we will take all fields to be independent of X^{25}: they are instead functions only on the non-compact $\mathbf{R}^{1,24}$.

Consider the metric in Einstein frame. This decomposes into three different fields on $\mathbf{R}^{1,24}$: a metric $\tilde{G}_{\mu\nu}$, a vector A_μ and a scalar σ which we package into the $D = 26$ dimensional metric as

$$ds^2 = \tilde{G}_{\mu\nu}\, dX^\mu\, dX^\nu + e^{2\sigma}\left(dX^{25} + A_\mu\, dX^\mu\right)^2 \tag{8.2}$$

Here all the indices run over the non-compact directions $\mu, \nu = 0, \ldots 24$ only.

The vector field A_μ is an honest gauge field, with the gauge symmetry descending from diffeomorphisms in $D = 26$ dimensions. To see this recall that under the transformation $\delta X^\mu = V^\mu(X)$, the metric transforms as

$$\delta G_{\mu\nu} = \nabla_\mu \Lambda_\nu + \nabla_\nu \Lambda_\mu$$

This means that diffeomorphisms of the compact direction, $\delta X^{25} = \Lambda(X^\mu)$, turn into gauge transformations of A_μ,

$$\delta A_\mu = \partial_\mu \Lambda$$

We'd like to know how the fields $G_{\mu\nu}$, A_μ and σ interact. To determine this, we simply insert the ansatz (8.2) into the $D = 26$ Einstein-Hilbert action. The $D = 26$ Ricci scalar $\mathcal{R}^{(26)}$ is given by

$$\mathcal{R}^{(26)} = \mathcal{R} - 2e^{-\sigma}\nabla^2 e^{\sigma} - \frac{1}{4}e^{2\sigma}F_{\mu\nu}F^{\mu\nu}$$

where $\mathcal{R}$ in this formula now refers to the $D = 25$ Ricci scalar. The action governing the dynamics becomes

$$S = \frac{1}{2\kappa^2}\int d^{26}X\sqrt{-\tilde{G}^{(26)}}\;\mathcal{R}^{(26)} = \frac{2\pi R}{2\kappa^2}\int d^{25}X\sqrt{-\tilde{G}}\;e^{\sigma}\left(\mathcal{R} - \frac{1}{4}e^{2\sigma}F_{\mu\nu}F^{\mu\nu} + \partial_\mu\sigma\partial^\mu\sigma\right)$$

The dimensional reduction of Einstein gravity in D dimensions gives Einstein gravity in $D-1$ dimensions, coupled to a $U(1)$ gauge theory and a single massless scalar. This illustrates the original idea of Kaluza and Klein, with Maxwell theory arising naturally from higher-dimensional gravity.

The gravitational action above is not quite of the Einstein-Hilbert form. We need to again change frames, absorbing the scalar σ in the same manner as we absorbed the dilaton in Section 7.3.1. Moreover, just as for the dilaton, there is no potential dictating the vacuum expectation value of σ. Changing the vev of σ corresponds to changing R, so this is telling us that nothing in the gravitational action fixes the radius R of the compact circle. This is a problem common to all Kaluza-Klein compactifications[11]: there are always massless scalar fields, corresponding to the volume of the internal space as well as other deformations. Massless scalar fields, such as the dilaton Φ or the volume σ, are usually referred to as *moduli*.

If we want this type of Kaluza-Klein compactification to describe our universe — where we don't see massless scalar fields — we need to find a way to "fix the moduli". This means that we need a mechanism which gives rise to a potential for the scalar fields, making them heavy and dynamically fixing their vacuum expectation value. Such mechanisms exist in the context of the superstring.

Let's now also look at the Kaluza-Klein reduction of the other fields in the low-energy effective action. The dilaton is easy: a scalar in D dimensions reduces to a scalar in $D-1$ dimensions. The anti-symmetric 2-form has more structure: it reduces to a 2-form $B_{\mu\nu}$, together with a vector field $\tilde{A}_\mu = B_{\mu\,25}$.

[11] The description of compactification on more general manifolds is a beautiful story involving aspects differential geometry and topology. This story is told in the second volume of Green, Schwarz and Witten.

In summary, the low-energy physics of the bosonic string in $D-1$ dimensions consists of a metric $G_{\mu\nu}$, two $U(1)$ gauge fields A_μ and $\tilde{A}_\mu$ and two massless scalars Φ and σ.

8.1.1 Moving around the Circle

In the above discussion, we assumed that all fields are independent of the periodic direction X^{25}. Let's now look at what happens if we relax this constraint. It's simplest to see the resulting physics if we look at the scalar field Φ where we don't have to worry about cluttering equations with indices. In general, we can expand this field in Fourier modes around the circle

$$\Phi(X^\mu; X^{25}) = \sum_{n=-\infty}^{\infty} \Phi_n(X^\mu) e^{inX^{25}/R}$$

where reality requires $\Phi_n^\star = \Phi_{-n}$. Ignoring the coupling to gravity for now, the kinetic terms for this scalar are

$$\int d^{26}X \; \partial_\mu \Phi \, \partial^\mu \Phi + (\partial_{25}\Phi)^2 = 2\pi R \int d^{25}X \sum_{n=-\infty}^{\infty} \left(\partial_\mu \Phi_n \, \partial^\mu \Phi_{-n} + \frac{n^2}{R^2} \,|\, \Phi_n|^2 \right)$$

This simple Fourier decomposition is telling us something very important: a single scalar field on $\mathbf{R}^{1,D-1} \times \mathbf{S}^1$ splits into an infinite number of scalar fields on $\mathbf{R}^{1,D-2}$, indexed by the integer n. These have mass

$$M_n^2 = \frac{n^2}{R^2} \tag{8.3}$$

For R small, all particles are heavy except for the massless zero mode $n = 0$. The heavy particles are typically called Kaluza-Klein (KK) modes and can be ignored if we're probing energies $\ll 1/R$ or, equivalently, distance scales $\gg R$.

There is one further interesting property of the KK modes Φ_n with $n \neq 0$: they are charged under the gauge field A_μ arising from the metric. The simplest way to see this is to look at the appropriate gauge transformation which, from the spacetime perspective, is the diffeomorphism $X^{25} \to X^{25} + \Lambda(X^\mu)$. Clearly, this shifts the KK modes

$$\Phi_n \to \exp\left(\frac{in\Lambda}{R}\right) \Phi_n$$

This tells us that the n^{th} KK mode has charge n/R. In fact, one usually rescales the gauge field to $A'_\mu = A_\mu/R$, under which the charge of the KK mode Φ_n is simply $n \in \mathbf{Z}$.

8.2 The View from the Worldsheet

We now consider the Kaluza-Klein reduction from the perspective of the string. We want to study a string moving in the background $\mathbf{R}^{1,24} \times \mathbf{S}^1$. There are two ways in which the compact circle changes the string dynamics.

The first effect of the circle is that the spatial momentum, p, of the string in the circle direction can no longer take any value, but is quantized in integer units

$$p^{25} = \frac{n}{R} \qquad n \in \mathbf{Z}$$

The simplest way to see this is simply to require that the string wavefunction, which includes the factor $e^{ip\cdot X}$, is single valued.

The second effect is that we can allow more general boundary conditions for the mode expansion of X. As we move around the string, we no longer need $X(\sigma + 2\pi) = X(\sigma)$, but can relax this to

$$X^{25}(\sigma + 2\pi) = X^{25}(\sigma) + 2\pi m R \qquad m \in \mathbf{Z}$$

The integer m tells us how many times the string winds around $\mathbf{S}^1$. It is usually simply called the *winding number.*

Let's now follow the familiar path that we described in Section 2 to study the spectrum of the string on the spacetime (8.1). We start by considering only the periodic field X^{25}, highlighting the differences with our previous treatment. The mode expansion of X^{25} is now given by

$$X^{25}(\sigma, \tau) = x^{25} + \frac{\alpha' n}{R}\tau + mR\sigma + \text{oscillator modes}$$

which incorporates both the quantized momentum and the possibility of a winding number. Before splitting $X^{25}(\sigma, \tau)$ into right-moving and left-moving parts, it will be useful to introduce the quantities

$$p_L = \frac{n}{R} + \frac{mR}{\alpha'} \quad , \quad p_R = \frac{n}{R} - \frac{mR}{\alpha'} \tag{8.4}$$

Then we have $X^{25}(\sigma, \tau) = X_L^{25}(\sigma^+) + X_R^{25}(\sigma^-)$, where

$$X_L^{25}(\sigma^+) = \tfrac{1}{2}x^{25} + \tfrac{1}{2}\alpha' p_L\,\sigma^+ + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\,\tilde{\alpha}_n^{25}\,e^{-in\sigma^+} \;,$$

$$X_R^{25}(\sigma^-) = \tfrac{1}{2}x^{25} + \tfrac{1}{2}\alpha' p_R\,\sigma^- + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\,\alpha_n^{25}\,e^{-in\sigma^-}$$

This differs from the mode expansion (1.36) only in the terms p_L and p_R. The mode expansion for all the other scalar fields on flat space $\mathbf{R}^{1,24}$ remains unchanged and we don't write them explicitly.

Let's think about what the spectrum of this theory looks like to an observer living in $D = 25$ non-compact directions. Each particle state will be described by a momentum p^μ with $\mu = 0, \ldots, 24$. The mass of the particle is

$$M^2 = -\sum_{\mu=0}^{24} p_\mu p^\mu$$

As before, the mass of these particles is fixed in terms of the oscillator modes of the string by the L_0 and $\tilde{L}_0$ equations. These now read

$$M^2 = p_L^2 + \frac{4}{\alpha'}(\tilde{N} - 1) = p_R^2 + \frac{4}{\alpha'}(N - 1)$$

where N and $\tilde{N}$ are the levels, defined in lightcone quantization by (2.24). (One should take the lightcone coordinate inside $\mathbf{R}^{1,24}$ rather than along the $\mathbf{S}^1$). The factors of -1 are the necessary normal ordering coefficients that we've seen in several guises in this course.

These equations differ from (2.25) by the presence of the momentum and winding terms around $\mathbf{S}^1$ on the right-hand side. In particular, level matching no longer tells us that $N = \tilde{N}$, but instead

$$N - \tilde{N} = nm \tag{8.5}$$

Expanding out the mass formula, we have

$$M^2 = \frac{n^2}{R^2} + \frac{m^2 R^2}{\alpha'^2} + \frac{2}{\alpha'}(N + \tilde{N} - 2) \tag{8.6}$$

The new terms in this formula have a simple interpretation. The first term tells us that a string with $n > 0$ units of momentum around the circle gains a contribution to its mass of n/R. This agrees with the result (8.3) that we found from studying the KK reduction of the spacetime theory. The second term is even easier to understand: a string which winds $m > 0$ times around the circle picks up a contribution $2\pi m R T$ to its mass, where $T = 1/2\pi\alpha'$ is the tension of the string.

8.2.1 Massless States

We now restrict attention to the massless states in $\mathbf{R}^{1,24}$. This can be achieved in the mass formula (8.6) by looking at states with zero momentum $n = 0$ and zero winding $m = 0$, obeying the level matching condition $N = \tilde{N} = 1$. The possibilities are

- $\alpha^{\mu}_{-1}\tilde{\alpha}^{\nu}_{-1}\,|0;p\rangle$: Under the $SO(1,24)$ Lorentz group, these states decompose into a metric $G_{\mu\nu}$, an anti-symmetric tensor $B_{\mu\nu}$ and a scalar Φ.
- $\alpha^{\mu}_{-1}\tilde{\alpha}^{25}_{-1}\,|0;p\rangle$ and $\alpha^{25}_{-1}\tilde{\alpha}^{\mu}_{-1}\,|0;p\rangle$: These are two vector fields. We can identify the sum of these $(\alpha^{\mu}_{-1}\tilde{\alpha}^{25}_{-1} + \alpha^{25}_{-1}\tilde{\alpha}^{\mu}_{-1})\,|0;p\rangle$ with the vector field A_μ coming from the metric and the difference $(\alpha^{\mu}_{-1}\tilde{\alpha}^{25}_{-1} - \alpha^{25}_{-1}\tilde{\alpha}^{\mu}_{-1})\,|0;p\rangle$ with the vector field $\tilde{A}_\mu$ coming from the anti-symmetric field.
- $\alpha^{25}_{-1}\tilde{\alpha}^{25}_{-1}\,|0;p\rangle$: This is another scalar. It is identified with the scalar σ associated to the radius of $\mathbf{S}^1$.

We see that the massless spectrum of the string coincides with the massless spectrum associated with the Kaluza-Klein reduction of the previous section.

8.2.2 Charged Fields

One can also check that the KK modes with $n \neq 0$ have charge n under the gauge field A_μ. We can determine the charge of a state under a given $U(1)$ by computing the 3-point function in which two legs correspond to the state of interest, while the third is the appropriate photon. We have two photons, with vertex operators given by,

$$V_\pm(p) \sim \int d^2z\; \zeta_\mu(\partial X^\mu \bar{\partial}\bar{X}^{25} \pm \partial X^{25}\bar{\partial}\bar{X}^\mu)e^{ip\cdot X}$$

where $+$ corresponds to A_μ and $-$ to $\tilde{A}_\mu$ and we haven't been careful about the overall normalization. Meanwhile, any state can be assigned momentum n and winding m by dressing the operator with the factor $e^{ip_L X^{25}(z)+ip_R\bar{X}^{25}(\bar{z})}$. As always, it's simplest to work with the momentum and winding modes of the tachyon, whose vertex operators are of the form

$$V_{m,n}(p) \sim \int d^2z\; e^{ip\cdot X}\, e^{ip_L X^{25}+ip_R\bar{X}^{25}}$$

The charge of a state is the coefficient in front of the 3-point coupling of the field and the photon,

$$\langle V_\pm(p_1)V_{m,n}(p_2)V_{-m,-n}(p_3)\rangle \sim \delta^{25}(\sum_i p_i)\,\zeta_\mu(p_2^\mu - p_3^\mu)\,(p_L \pm p_R)$$

The first few factors are merely kinematical. The interesting information is in the last factor. It is telling us that under A_μ, fields have charge $p_L + p_R \sim n/R$. This is in agreement with the Kaluza-Klein analysis that we saw before. However, it's also telling us something new: under $\tilde{A}_\mu$, fields have charge $p_L - p_R \sim mR/\alpha'$. In other words, winding modes are charged under the gauge field that arises from the reduction of $B_{\mu\nu}$. This is not surprising: winding modes correspond to strings wrapping the circle and we saw in Section 7 that strings are electrically charged under $B_{\mu\nu}$.

8.2.3 Enhanced Gauge Symmetry

With a circle in the game, there are other ways to build massless states that don't require us to work at level $N = \tilde{N} = 1$. For example, we can set $N = \tilde{N} = 0$ and look at winding modes $m \neq 0$. The level matching condition (8.5) requires $n = 0$ and the mass of the states is

$$M^2 = \left(\frac{mR}{\alpha'}\right)^2 - \frac{4}{\alpha'}$$

and states can be massless whenever the radius takes special values $R^2 = 4\alpha'/m^2$ with $m \in \mathbf{Z}$. Similarly, we can set the winding to zero $m = 0$ and consider the KK modes of the tachyon which have mass

$$M^2 = \frac{n^2}{R^2} - \frac{4}{\alpha'}$$

which become massless when $R^2 = n^2\alpha'/4$.

However, the richest spectrum of massless states occurs when the radius takes a very special value, namely

$$R = \sqrt{\alpha'}$$

Solutions to the level matching condition (8.5) with $M^2 = 0$ are now given by

- $N = \tilde{N} = 1$ with $m = n = 0$. These give the states described above: a metric, two $U(1)$ gauge fields and two neutral scalars.
- $N = \tilde{N} = 0$ with $n = \pm 2$ and $m = 0$. These are KK modes of the tachyon field. They are scalars in spacetime with charges $(\pm 2, 0)$ under the $U(1) \times U(1)$ gauge symmetry.
- $N = \tilde{N} = 0$ with $n = 0$ and $m = \pm 2$. This is a winding mode of the tachyon field. They are scalars in spacetime with charges $(0, \pm 2)$ under $U(1) \times U(1)$.

- $N = 1$ and $\tilde{N} = 0$ with $n = m = \pm 1$. These are two new spin 1 fields, $\alpha^\mu_{-1}\,|0;p\rangle$. They carry charge $(\pm 1, \pm 1)$ under the two $U(1) \times U(1)$.
- $N = 1$ and $\tilde{N} = 0$ with $n = -m = \pm 1$. These are a further two spin 1 fields, $\tilde{\alpha}^\mu_{-1}\,|0;p\rangle$, with charge $(\pm 1, \mp 1)$ under $U(1) \times U(1)$.

How do we interpret these new massless states? Let's firstly look at the spin 1 fields. These are charged under $U(1) \times U(1)$. As we mentioned in Section 7.7, the only way to make sense of charged massless spin 1 fields is in terms of a non-Abelian gauge symmetry. Looking at the charges, we see that at the critical radius $R = \sqrt{\alpha'}$, the theory develops an enhanced gauge symmetry

$$U(1) \times U(1) \rightarrow SU(2) \times SU(2)$$

The massless scalars from the $N = \tilde{N} = 0$ now join with the previous scalars to form adjoint representations of this new symmetry. We move away from the critical radius by changing the vacuum expectation value for σ. This breaks the gauge group back to the Cartan subalgebra by the Higgs mechanism.

From the discussion above, it's clear that this mechanism for generating non-Abelian gauge symmetries relies on the existence of the tachyon. For this reason, this mechanism doesn't work in Type II superstring theories. However, it turns out that it does work in the heterotic string, even though it has no tachyon in its spectrum.

8.3 Why Big Circles are the Same as Small Circles

The formula (8.6) has a rather remarkable property: it is invariant under the exchange

$$R \;\leftrightarrow\; \frac{\alpha'}{R} \tag{8.7}$$

if, at the same time, we swap the quantum numbers

$$m \;\leftrightarrow\; n \tag{8.8}$$

This means that a string moving on a circle of radius R has the same spectrum as a string moving on a circle of radius α'/R. It achieves this feat by exchanging what it means to wind with that it means to move.

As the radius of the circle becomes large, $R \rightarrow \infty$, the winding modes become very heavy with mass $\sim R/\alpha'$ and are irrelevant for the low-energy dynamics. But the momentum modes become very light, $M \sim 1/R$, and, in the strict limit form a continuum. From the perspective of the energy spectrum, this continuum of energy states is exactly what we mean by the existence of a non-compact direction in space.

In the other limit, $R \to 0$, the momentum modes become heavy and can be ignored: it takes way too much energy to get anything to move on the $\mathbf{S}^1$. In contrast, the winding modes become light and start to form a continuum. The resulting energy spectrum looks as if another dimension of space is opening up!

The equivalence of the string spectrum on circles of radii R and α'/R extends to the full conformal field theory and hence to string interactions. Strings are unable to tell the difference between circles that are very large and circles that are very small. This striking statement has a rubbish name: it is called *T-duality*.

This provides another mechanism in which string theory exhibits a minimum length scale: as you shrink a circle to smaller and smaller sizes, at $R = \sqrt{\alpha'}$, the theory acts as if the circle is growing again, with winding modes playing the role of momentum modes.

The New Direction in Spacetime

So how do we describe this strange new spatial direction that opens up as $R \to 0$? Under the exchange (8.7) and (8.8), we see that p_L and p_R transform as

$$p_L \to p_L \quad , \quad p_R \to -p_R$$

Motivated by this, we define a new scalar field,

$$Y^{25} = X_L^{25}(\sigma^+) - X_R^{25}(\sigma^-)$$

It is simple to check that in the CFT for a free, compact scalar field all OPEs of Y^{25} coincide with the OPEs of X^{25}. This is sufficient to ensure that all interactions defined in the CFT are the same.

We can write the new spatial direction Y directly in terms of the old field X, without first doing the split into left and right-moving pieces. From the definition of Y, one can check that $\partial_\tau X = \partial_\sigma Y$ and $\partial_\sigma X = \partial_\tau Y$. We can write this in a unified way as

$$\partial_\alpha X = \epsilon_{\alpha\beta}\, \partial^\beta Y \tag{8.9}$$

where $\epsilon_{\alpha\beta}$ is the antisymmetric matrix with $\epsilon_{\tau\sigma} = -\epsilon_{\sigma\tau} = +1$. (The minus sign from $\epsilon_{\sigma\tau}$ in the above equation is canceled by another from the Minkowski worldsheet metric when we lower the index on ∂^β).

The Shift of the Dilaton

The dilaton, or string coupling, also transforms under T-duality. Here we won't derive this in detail, but just give a plausible explanation for why it's the case. The main idea is that a scientist living in a stringy world shouldn't be able to do any experiments that distinguish between a compact circle of radius R and one of radius α'/R. But the first place you would look is simply the low-energy effective action which, working in Einstein frame, contains terms like

$$\frac{2\pi R}{2l_s^{24}g_s^2}\int d^{25}X\sqrt{-\tilde{G}}\;e^{\sigma}\,\mathcal{R}+\ldots$$

A scientist cannot tell the difference between R and $\tilde{R}=\alpha'/R$ only if the value of the dilaton is also ambiguous so that the term in front of the action remains invariant: i.e. $R/g_s^2=\tilde{R}/\tilde{g}_s^2$. This means that, under T-duality, the dilaton must shift so that the coupling constant becomes

$$g_s\to\tilde{g}_s=\frac{\sqrt{\alpha'}g_s}{R} \tag{8.10}$$

8.3.1 A Path Integral Derivation of T-Duality

There's a simple way to see T-duality of the quantum theory using the path integral. We'll consider just a single periodic scalar field $X\equiv X+2\pi R$ on the worldsheet. It's useful to change normalization and write $X=R\varphi$, so that the field φ has periodicity 2π. The radius R of the circle now sits in front of the action,

$$S[\varphi]=\frac{R^2}{4\pi\alpha'}\int d^2\sigma\;\partial_\alpha\varphi\,\partial^\alpha\varphi \tag{8.11}$$

The Euclidean partition function for this theory is $Z=\int\mathcal{D}\varphi\,e^{-S[\varphi]}$. We will now play around with this partition function and show that we can rewrite it in terms of new variables that describe the T-dual circle.

The theory (8.11) has a simple shift symmetry $\varphi\to\varphi+\lambda$. The first step is to make this symmetry local by introducing a gauge field A_α on the worldsheet which transforms as $A_\alpha\to A_\alpha-\partial_\alpha\lambda$. We then replace the ordinary derivatives with covariant derivatives

$$\partial_\alpha\varphi\to\mathcal{D}_\alpha\varphi=\partial_\alpha\varphi+A_\alpha$$

This changes our theory. However, we can return to the original theory by adding a new field, θ which couples as

$$S[\varphi,\theta,A]=\frac{R^2}{4\pi\alpha'}\int d^2\sigma\;\mathcal{D}_\alpha\varphi\,\mathcal{D}^\alpha\varphi+\frac{i}{2\pi}\int d^2\sigma\;\theta\,\epsilon^{\alpha\beta}\partial_\alpha A_\beta \tag{8.12}$$

The new field θ acts as a Lagrange multiplier. Integrating out θ sets $\epsilon^{\alpha\beta}\partial_\alpha A_\beta = 0$. If the worldsheet is topologically $\mathbf{R}^2$, then this condition ensures that A_α is pure gauge which, in turn, means that we can pick a gauge such that $A_\alpha = 0$. The quantum theory described by (8.12) is then equivalent to that given by (8.11).

Of course, if the worldsheet is topologically $\mathbf{R}^2$ then we're missing the interesting physics associated to strings winding around φ. On a non-trivial worldsheet, the condition $\epsilon^{\alpha\beta}\partial_\alpha A_\beta = 0$ does not mean that A_α is pure gauge. Instead, the gauge field can have non-trivial holonomy around the cycles of the worldsheet. One can show that these holonomies are gauge trivial if θ has periodicity 2π. In this case, the partition function defined by (8.12),

$$Z = \frac{1}{\text{Vol}} \int \mathcal{D}\varphi \mathcal{D}\theta \mathcal{D}A \; e^{-S[\varphi,\theta,A]}$$

is equivalent to the partition function constructed from (8.11) for worldsheets of any topology.

At this stage, we make use of a clever and ubiquitous trick: we reverse the order of integration. We start by integrating out φ which we can do by simply fixing the gauge symmetry so that $\varphi = 0$. The path integral then becomes

$$Z = \int \mathcal{D}\theta \mathcal{D}A \; \exp\left(-\frac{R^2}{4\pi\alpha'} \int d^2\sigma \; A_\alpha A^\alpha + \frac{i}{2\pi} \int d^2\sigma \; \epsilon^{\alpha\beta} \left(\partial_\alpha \theta\right) A_\beta \right)$$

where we have also taken the opportunity to integrate the last term by parts. We can now complete the procedure and integrate out A_α. We get

$$Z = \int \mathcal{D}\theta \; \exp\left(-\frac{\tilde{R}^2}{4\pi\alpha'} \int d^2\sigma \; \partial_\alpha \theta \, \partial^\alpha \theta \right)$$

with $\tilde{R} = \alpha'/R$ the radius of the T-dual circle. In the final integration, we threw away the overall factor in the path integral, which is proportional to $\sqrt{\alpha'/R}$. A more careful treatment shows that this gives rise to the appropriate shift in the dilaton (8.10).

8.3.2 T-Duality for Open Strings

What happens to open strings and D-branes under T-duality? Suppose firstly that we compactify a circle in direction X transverse to the brane. This means that X has Dirichlet boundary conditions

$$X = \text{const} \quad \Rightarrow \quad \partial_\tau X^{25} = 0 \qquad \text{at } \sigma = 0, \pi$$

But what happens in the T-dual direction Y? From the definition (8.9) we learn that the new direction has Neumann boundary conditions,

$$\partial_\sigma Y = 0 \quad \text{at } \sigma = 0, \pi$$

We see that T-duality exchanges Neumann and Dirichlet boundary conditions. If we dualize a circle transverse to a Dp-brane, then it turns into a D$(p+1)$-brane.

The same argument also works in reverse. We can start with a Dp-brane wrapped around the circle direction X, so that the string has Neumann boundary conditions. After T-duality, (8.9) changes these to Dirichlet boundary conditions and the Dp-brane turns into a D$(p-1)$-brane, localized at some point on the circle Y.

In fact, this was how D-branes were originally discovered: by following the fate of open strings under T-duality.

8.3.3 T-Duality for Superstrings

To finish, let's nod one final time towards the superstring. It turns out that the ten-dimensional superstring theories are not invariant under T-duality. Instead, they map into each other. More precisely, Type IIA and IIB transform into each other under T-duality. This means that Type IIA string theory on a circle of radius R is equivalent to Type IIB string theory on a circle of radius α'/R. This dovetails with the transformation of D-branes, since type IIA has Dp-branes with p even, while IIB has p odd. Similarly, the two heterotic strings transform into each other under T-duality.

8.3.4 Mirror Symmetry

The essence of T-duality is that strings get confused. Their extended nature means that they're unable to tell the difference between big circles and small circles. We can ask whether this confusion extends to more complicated manifolds. The answer is yes. The fact that strings can see different manifolds as the same is known as *mirror symmetry*.

Mirror symmetry is cleanest to state in the context of the Type II superstring, although similar behaviour also holds for the heterotic strings. The simplest example is when the worldsheet of the string is governed by a superconformal non-linear sigma-model with target space given by some Calabi-Yau manifold $\mathbf{X}$. The claim of mirror symmetry is that this CFT is identical to the CFT describing the string moving on a different Calabi-Yau manifold $\mathbf{Y}$. The topology of $\mathbf{X}$ and $\mathbf{Y}$ is not the same. Their Hodge diamonds are the mirror of each other; hence the name. The subject of mirror symmetry is an active area of research in geometry and provides a good example of the impact of string theory on mathematics.

8.4 Epilogue

We are now at the end of this introductory course on string theory. We began by trying to make sense of the quantum theory of a relativistic string moving in flat space. It is, admittedly, an odd place to start. But from then on we had no choices to make. The relativistic string leads us ineluctably to conformal field theory, to higher dimensions of spacetime, to Einstein's theory of gravity at low-energies, to good UV behaviour at high-energies and to Yang-Mills theories living on branes. There are few stories in theoretical physics where such meagre input gives rise to such a rich structure.

This journey continues. There is one further ingredient that it is necessary to add: supersymmetry. Even this is in some sense not a choice, but is necessary to remove the troublesome tachyon that plagued these lectures. From there we may again blindly follow where the string leads, through anomalies (and the lack thereof) in ten dimensions, to dualities and M-theory in eleven dimensions, to mirror symmetry and moduli stabilization and black hole entropy counting and holography and the miraculous AdS/CFT correspondence.

However, the journey is far from complete. There is much about string theory that remains to be understood. This is true both of the mathematical structure of the theory and of its relationship to the world that we observe. The problems that we alluded to in Section 6.4.5 are real. Non-perturbative completions of string theory are only known in spacetimes which are asymptotically anti-de Sitter, but cosmological observations suggest that our home is not among these. In attempts to make contact with the standard models of particle physics and cosmology, we typically return to the old idea of Kaluza-Klein compactifications. Is this the right approach? Or are we missing some important and subtle conceptual ingredient? Or is the existence of this remarkable mathematical structure called string theory merely a red-herring that has nothing to do with the real world?

In the years immediately after its birth, no one knew that string theory was a theory of strings. It seems very possible that we're currently in a similar situation. When the theory is better understood, it may have little to do with strings. We are certainly still some way from answering the simple question: what is string theory really?

White holes and eternal black holes

Stephen D. H. Hsu*
Institute of Theoretical Science, University of Oregon, Eugene, OR 97403

(Dated: November 2011)

We investigate isolated white holes surrounded by vacuum, which correspond to the time reversal of eternal black holes that do not evaporate. We show that isolated white holes produce quasi-thermal Hawking radiation. The time reversal of this radiation, incident on a black hole precursor, constitutes a special preparation that will cause the black hole to become eternal.

What is a white hole?

White holes have received far less attention from researchers than black holes (for a review, see, e.g., [1]). This is understandable, given that conditions in our universe readily lead to black hole formation, whereas white hole creation has neither been observed nor is expected to have occurred in the history of the universe.

However, white holes are themselves fundamental objects and worthy of further study. White holes are time-reversed black holes, and therefore characterized by the same quantum numbers: mass, angular momentum, charge. While a classical black hole spacetime has a singularity in the future, a white hole has one in the past. If quantum gravitational effects can resolve black hole singularities, then white holes need not result from singular initial conditions. (In any case the initial white hole singularity is not directly visible to observers.)

Standard quantum mechanical reasoning suggests that any initial state which evolves into a black hole also has some nonzero probability to evolve into a white hole. Note we are referring to a quantum gravitational process, and are making the assumption that even in quantum gravity tunneling between two states with the same quantum numbers has non-zero (although possibly very small) probability. For example, it is known that the collision of two sufficiently energetic particles can create a black hole [2]. Because the quantum numbers are the same we would expect that the same energetic particles have a small but non-zero probability of producing a white hole with the same quantum numbers [3]. Since large black holes are long-lived, there are some white hole states (corresponding to time slices late in the black hole's existence) that persist for a long time before exploding. Thus, long-lived white holes are a consequence of quantum mechanics and the properties of black holes.

A class of highly entropic objects whose full spacetime evolution is that of a white hole which explodes outwards, is stopped by gravitational self-attraction, and recollapses to form a black hole, are described in [4].

Hawking's arguments and thermal equilibrium

In his 1976 paper *Black holes and thermodynamics* [5], Hawking analyzed the properties of white holes by considering a box in thermal equilibrium, whose temperature and volume are adjusted so that the most probable configuration is a black hole surrounded by a gas of particles whose temperature is equal to that of a black hole. The black hole emits Hawking radiation but absorbs, on average, as much energy from the gas as it emits. Applying time reversal, the configuration describes a white hole emitting and absorbing radiation. Since there is no arrow of time for a system in thermal equilibrium, Hawking argued that black and white holes must be indistinguishable. More precisely, the properties of white and black holes *in equilibrium with their surroundings* are identical. However, the same cannot be said for black and white holes in isolation (i.e., surrounded by empty vacuum)– we shall see that their properties are radically different.

In elementary particle physics we are accustomed to the idea that time reversal maps particles to their antiparticles. However, in the case of black and white holes the subsequent evolution of the time-reversed object depends on more than just quantum numbers such as M, J, Q. For a hot black hole in a cold environment, there is a statistical arrow of time.

Isolated white holes

Consider a white hole (figure 1) in a spacetime with the property that past null infinity is in the empty vacuum state of ordinary flat space. This implies that space far from the hole is empty and that there is no incoming radiation from the past. This white hole spacetime is the time reversal of a black hole spacetime with no Hawking radiation propagating to future infinity (figure 2). One motivation for considering such objects is that they are localized, as opposed to an entire spacelike slice of a box in thermal equilibrium. Do such white holes exist? What are their properties? (For simplicity, we assume all quantum numbers of the hole, other than its mass, are zero.)

In this discussion we will refer to the diagrams in figures 1 and 2, which depict a black hole spacetime (figure 2) and its time reversal (figure 1). We will refer to past and future null infinity of the black hole spacetime as $\mathcal{J}^{\rm bh}_{\pm}$, where the subscript + indicates future, and − indicates past. In the white hole diagram the role of past and future are reversed: $\mathcal{J}^{\rm wh}_{\mp} = \mathcal{J}^{\rm bh}_{\pm}$.

Assuming that the white hole is isolated implies the

*Electronic address: hsu@uoregon.edu

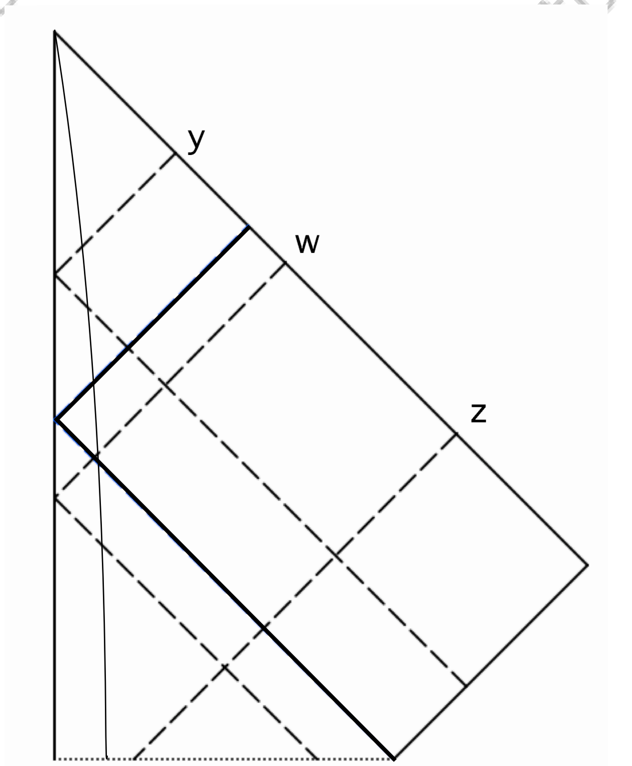


FIG. 1: A white hole spacetime. We impose the condition that past null infinity $\mathcal{J}_-^{\rm wh}$ is in the vacuum state – there is no incoming radiation from the far past. The dotted black line is the initial singularity, and the thick solid line is the path of a null ray on the anti-horizon. The curved line indicates matter which explodes out of the hole. The dashed black lines refer to modes discussed in the text.

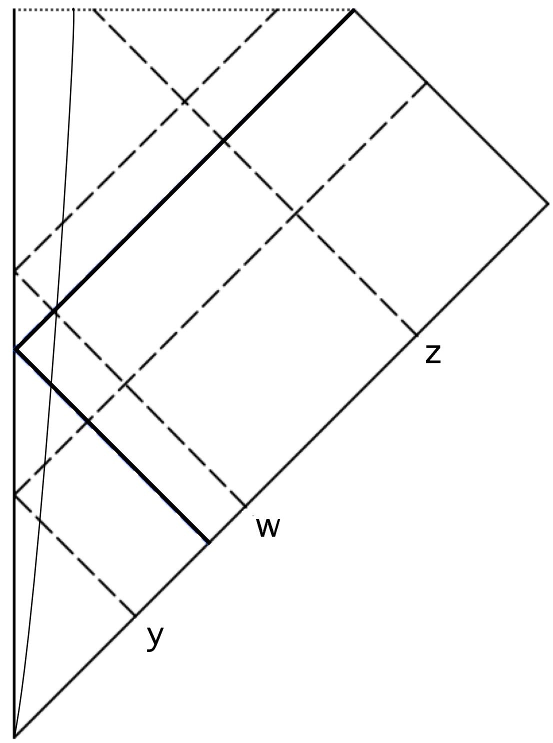


FIG. 2: A black hole spacetime which is the time reversal of figure 1. We impose the condition that future null infinity $\mathcal{J}_+^{\rm bh}$ is in the vacuum state – there is no outgoing Hawking radiation. The dotted black line is the final singularity, and the thick solid line is the path of a null ray which coincides with the horizon at late times. The curved line indicates matter which collapses to form the hole. The dashed black lines refer to modes discussed in the text.

empty vacuum on $\mathcal{J}_-^{\rm wh}$: there is no incoming radiation from the past.

This condition is equivalent, on the black hole spacetime, to no Hawking radiation propagating to future null infinity $\mathcal{J}_+^{\rm bh}$. This sounds strange, but can be accomplished by proper choice of initial state from which the black hole is formed. That is, a special arrangement of incoming modes from $\mathcal{J}_-^{\rm bh}$ is required; see below for details. In the white hole spacetime these modes would be seen exiting the white hole after it explodes from behind its anti-horizon.

In our discussion we *assume* that the black hole spacetime (figure 2) describes a progenitor (e.g., a star) which collapses to form the hole. Because ordinary stars and other progenitors in nearly-flat space obey an entropy bound: $S < A^{3/4}$, where A is their surface area in Planck units, such objects have much lower entropy than a black hole with no constraint on how it was formed [4, 6]. Indeed, a generic black hole, formed in a maximally entropic process (e.g., by allowing an initially small black hole to slowly accrete matter) has entropy of order A, but such objects do not satisfy the isolation condition imposed above. Thus, the objects under study are very exotic (improbable): not only are they white holes, but the condition of isolation further reduces the entropy substantially. Our analysis is mainly of theoretical, rather than practical astrophysical, interest.

In our analysis we only treat the case of a Schwarzschild black hole. Holes with angular momentum or electric charge have a more complex inner structure, including a Cauchy horizon. Interactions between outgoing and backscattered ingoing radiation near this Cauchy horizon lead to a curvature singularity known as mass inflation [7]. The resulting inner structure seems to involve quantum gravitational effects and is still not completely understood.

Following Hawking [8], we define an orthogonal set of modes for a scalar field on the black hole spacetime.

$$\phi = \int d\omega \left(f_\omega a_\omega + \bar{f}_\omega a_\omega^\dagger \right) \quad , \tag{1}$$

where the $\{f_\omega\}$ are a complete, orthonormal family of complex solutions of the wave equation. The notation used here is identical to that in [8], except that (see below) we define the destruction operators of the w, y, z modes to be a^w, a^y, a^z rather than g, h, j as Hawking did. In all other respects, we adhere to his definitions, which we now briefly review.

The modes are defined on an extended Schwarzschild geometry, obtained by analytic continuation, which includes both past and future horizons, $\mathcal{H}^{\pm}$, but does not describe the collapse which forms the black hole. Let u and v be the retarded and advanced time coordinates for the Schwarzschild metric. Let U (Kruskal coordinate) be the affine parameter on the the past horizon $\mathcal{H}^{-}$:

$$u = -4M \ln(-U) \quad , \tag{2}$$

where M is the mass of the black hole. The $\{f^{(1)}_\omega\}$ modes are solutions on the extended spacetime with zero Cauchy data on the past horizon and time dependence of the form $\exp(i\omega v)$ on $\mathcal{J}^{\rm bh}_-$. The $\{f^{(2)}_\omega\}$ are solutions with zero Cauchy data on $\mathcal{J}^{\rm bh}_-$ and dependence $\exp(i\omega U)$ on the past horizon. The analytic continuation of u yields two coordinates

$$u_\pm = -4M(\ln U \mp i\pi) \quad (U > 0) \ , \tag{3}$$

with $u_+ = u_-$ for $U < 0$. These are used to replace the $f^{(2)}$ modes by two orthogonal families of solutions $f^{(3)}$ and $f^{(4)}$. These have zero Cauchy data on $\mathcal{J}^{\rm bh}_-$ and dependence on the past horizon of the form $\exp(i\omega u_+)$ and $\exp(i\omega u_-)$, respectively.

The physical interpretation of these modes, after continuation back to the collapse spacetime, is as follows. The $f^{(1)}$ modes enter the black hole after a horizon has formed. The $f^{(2)}$ modes (equivalently, the $f^{(3),(4)}$ modes) enter the black hole region before the $f^{(1)}$ modes, at earlier advanced time; in the time-reversed spacetime they would exit the white hole before it emerges from behind its anti-horizon (see figure 1). The quantum states associated with these modes are observable to a detector at $\mathcal{J}^{\rm bh}_-$, or equivalently at $\mathcal{J}^{\rm wh}_+$. We define the destruction operators for the f modes to be a^1, a^3, a^4.

It is useful to define additional bases of modes (see figures): $\{w_\omega\}$, $\{y_\omega\}$ and $\{z_\omega\}$, which are linear combinations of the $\{f^{(i)}_\omega\}$, and are observable by a detector at $\mathcal{J}^{\rm bh}_+$. The $\{w_\omega\}$ modes have zero Cauchy data on $\mathcal{J}^{\rm bh}_-$ and on the past horizon for $U < 0$. For $U > 0$ on the past horizon their dependence is of the form $\exp(-i\omega u_+)$. The $\{y_\omega\}$ modes have zero Cauchy data on $\mathcal{J}^{\rm bh}_-$ and on the past horizon for $U > 0$. For $U < 0$ on the past horizon their dependence is of the form $\exp(i\omega u_+)$. The $\{z_\omega\}$ modes are identical to the $f^{(1)}$ modes already defined. The destruction operators for these new modes are a^w, a^y, a^z.

The physical interpretation of these modes, after continuation back to the collapse spacetime, is as follows. The y modes enter the spatial region where the black hole will be formed (i.e., the precursor), but emerge before a horizon appears. The transmitted y modes (which are not reflected by the gravitational potential back into the hole) are observable at future null infinity of the black hole spacetime, $\mathcal{J}^{\rm bh}_+$. The w modes propagate in from $\mathcal{J}^{\rm bh}_-$, enter the black hole region of space (i.e., the precursor) before a horizon is formed, but are trapped and encounter the future singularity. In the white hole spacetime (see figure 1), w modes emerge first from the anti-horizon, followed by the y modes, which appear after matter begins to explode from behind the anti-horizon.

The Hawking radiation modes p_ω, which are observable by a detector at $\mathcal{J}^{\rm bh}_+$, are a complete set of orthonormal solutions which contain only positive frequencies at $\mathcal{J}^{\rm bh}_+$ and are purely outgoing (zero Cauchy data on the horizon of the collapse spacetime). They can be written in terms of the y and z modes [8]:

$$p_\omega = t_\omega y_\omega + r_\omega z_\omega \quad , \tag{4}$$

and the destruction operator for this mode is

$$a^p_\omega = \bar{t}_\omega a^y_\omega + \bar{r}_\omega a^z_\omega \quad . \tag{5}$$

Equation (4) can be understood as follows from figure 2, depicting the black hole spacetime. The modes which reach future infinity are a superposition of transmitted y modes and reflected z modes, where t and r are the transmission and reflection amplitudes for waves incident on the black hole.

The condition that $\mathcal{J}^{\rm bh}_+$ is in the vacuum state (no Hawking radiation; an *eternal* black hole) is

$$a^p_\omega \, |0^{\rm bh}_+\rangle = 0 \ . \tag{6}$$

This condition is not typically imposed on the future state of the black hole spacetime. Instead, one usually requires that the precursor state (i.e., a collapsing star) is surrounded by vacuum, which is a condition on the past rather than on the future. However, the time reversal symmetry of quantum field theory and of general relativity imply that there must exist initial conditions that lead to the future condition (6). In the white hole case it is natural to impose the vacuum condition on the past, and we explore what its consequences are for the future of the hole.

A sufficient, but not necessary, condition for satisfying (6) is to require

$$a^y_\omega \, |0^{\rm bh}_+\rangle = a^z_\omega \, |0^{\rm bh}_+\rangle = 0 \ . \tag{7}$$

In his original discussion of the future vacuum on the black hole spacetime [8], Hawking imposes (7) as well as the additional condition

$$a^w_\omega \, |0^{\rm bh}_+\rangle = 0 \ , \tag{8}$$

requiring that the future vacuum be empty of unnecessary w modes. In our discussion of white holes this condition need not apply since we do not wish to constrain the initial state of the white hole other than to require its isolation.

It is straightforward to calculate the particle number content, mode by mode, for the state defined above. We are specifically interested in the $f^{(i)}$ modes, which are detectable as particles incident on the black hole and its precursor by an observer at $\mathcal{J}^{\rm bh}_-$ (past infinity of the black hole spacetime). Equivalently, these modes are detectable as outgoing particles by an observer far outside the white hole. The condition $a^z_\omega |0^{\rm bh}_+\rangle = 0$, imposed in (7), is identical to the condition $a^1 |0^{\rm bh}_+\rangle = 0$, which implies that there are no $f^{(1)}$ or z modes emitted by the white hole (or absorbed by the eternal black hole). These modes are emitted by the white hole long before it explodes from behind its anti-horizon, or equivalently are absorbed by the black hole long after its horizon forms (see figures). The remaining $f^{(3,4)}$ modes are linear combinations of the w and y modes, which are emitted by the white hole just before and after it explodes. The y modes, in particular, appear to be emitted from the ejecta of the hole.

We obtain

$$\langle 0^{\rm bh}_+|\, a^{3\,\dagger}_\omega a^3_\omega + a^{4\,\dagger}_\omega a^4_\omega\, |0^{\rm bh}_+\rangle = \frac{2x}{1-x} + \frac{1+x}{1-x}\langle 0^{\rm bh}_+|\, a^{w\,\dagger}_\omega a^w_\omega\, |0^{\rm bh}_+\rangle\,, \tag{9}$$

where $x = \exp(-\beta\omega)$ and β the Hawking temperature of the black hole. In the simple case with $a^w_\omega|0^{\rm bh}_+\rangle = 0$, the particle occupation numbers of each of the $f^{(3)}$ and $f^{(4)}$ modes are simply those of the blackbody distribution ($i = 3$ or 4):

$$\langle 0^{\rm bh}_+|\, a^{i\,\dagger}_\omega a^i_\omega |0^{\rm bh}_+\rangle = \frac{x}{1-x} = \frac{1}{\exp(\beta\omega)-1}\quad . \tag{10}$$

Physically, this means that one can construct an *eternal* (i.e., non-radiating) black hole in the minimal state satisfying (7) and (8) by exposing its precursor (and, briefly, the black hole itself) to a special quasi-thermal radiation state. It also implies that an isolated white hole in the state satisfying (8) will radiate quasi-thermally just before and after it explodes from behind its anti-horizon. Note that although the occupation numbers we have calculated are thermal, the state is actually a pure state if the initial white hole state was pure. Unlike in the case of Hawking radiation, we are not forced to trace over any causally disconnected region, and we do not necessarily obtain a mixed state description.

In our analysis so far we have treated the background spacetime as fixed and have neglected backreaction effects. In the original Hawking analysis, one first obtains the thermal spectrum of black hole radiation, and then invokes energy conservation and backreaction to argue that the hole steadily loses mass through radiation, eventually (perhaps) evaporating completely. Since the rate of energy loss is small it is assumed that the semiclassical analysis pertains until the final Planckian stage of evaporation. In our case we can make the same argument regarding the white hole: our calculations initially assume a fixed spacetime, but lead to thermal behavior of the hole just before and after it explodes. Conservation of energy implies that the white hole and the ejecta somehow compensate for the emitted radiation so that the total energy that reaches infinity is the initial ADM mass of the hole. How this happens is not entirely clear, although one can simply regard it as a constraint on possible final states resulting from an isolated white hole. We note that the mode bases used in this analysis only depend on the asymptotic structure of the black or white hole spacetime. The details of how the black hole is formed, or how the white hole explodes, do not affect the results; indeed, the analysis can be formulated on the extended Schwarzschild spacetime which is the analytic continuation of the realistic geometry which contains a collapsing/exploding body.

A necessary and sufficient condition for isolation of the white hole (as opposed to (7), which was sufficient, but a special case) is

$$\bar{t}_\omega a^y_\omega\, |0^{\rm bh}_+\rangle = -\bar{r}_\omega a^z_\omega\, |0^{\rm bh}_+\rangle\ . \tag{11}$$

As mentioned previously, the condition (11) can be understood (see figure 2) as the requirement that reflected z modes interfere perfectly with transmitted y modes so that no Hawking radiation reaches future infinity of the black hole spacetime. This, more general, condition allows for Hawking-like radiation from the white hole in the form of z modes, which leave the white hole long before its explosion and reach future infinity $\mathcal{J}^{\rm wh}_+$.

In the general case, we obtain the following expression for $f^{(3,4)}$ mode occupation numbers:

$$\langle 0^{\rm bh}_+|\sum_{i=3,4} a^{i\,\dagger}_\omega a^i_\omega\, |0^{\rm bh}_+\rangle = \frac{1}{1-x}\Big[2x + \langle 0^{\rm bh}_+|\,(1+x)\times (a^{y\,\dagger}a^y + a^{w\,\dagger}a^w) - 2\sqrt{x}(a^w a^y + a^{w\,\dagger}a^{y\,\dagger})\,|0^{\rm bh}_+\rangle\Big]\,. \tag{12}$$

For $|0^{\rm bh}_+\rangle$ which are particle number eigenstates of y and w, one can use the Cauchy-Schwarz inequality

$$|\langle a^w a^y\rangle|^2 \le \langle\, a^{w\,\dagger}a^w\ a^{y\,\dagger}a^y\,\rangle\ ,$$

and identities $(1+x) \ge 2\sqrt{x}$ and $N_y + N_w \ge 2\sqrt{N_w N_y}$, to see that the expectation value of the mode number is, for every frequency, at least as large as in the simplest case where the conditions (7) and (8) are satisfied.

For a white hole to be indistinguishable from an ordinary black hole it must emit Hawking-like radiation from the beginning, with thermal occupation numbers for $\langle a^{z\,\dagger}a^z\rangle$. Condition (11) then requires non-zero occupation numbers for $\langle a^{y\,\dagger}a^y\rangle$, leading to more energy radiated in $f^{(3,4)}$ modes at late times. A significant amount of energy in this form must emerge *after* the white hole explodes, which limits how much can be radiated before it explodes. It is hard to see how an isolated white hole can behave so as to be indistinguishable from an ordinary black hole of equal mass. This only seems possible if we remove the condition of isolation, allowing the white hole to both emit and absorb energy, as would be the case for the thermal box considered originally by Hawking [5].

Conclusions

We summarize our main results below. These results have not, to our knowledge appeared previously in the literature.

1. Isolated white holes behave very differently from isolated black holes. This is due to the lack of time reversal symmetry in the surrounding environment: the statistical arrow of time implies that isolated black holes evaporate into their cold surroundings, whereas isolated white holes are, by definition, not bathed in incident radiation. Complete time reversal symmetry is only present in thermal equilibrium, the case originally analyzed by Hawking.

2. Isolated white holes with initial state given by the simple conditions (7) and (8) emit quasi-thermal radiation just before and after exploding from behind their anti-horizon. Modifying the initial state, while retaining the condition of isolation, likely implies even more radiation at late stages. There do not seem to be isolated white holes which are indistinguishable from isolated black holes of the same mass.

3. As a byproduct of our investigation, we note the existence of eternal – non-evaporating – black holes, formed from special quantum initial states. We do not know whether such holes are stable against perturbations. That is, if one prepares a black hole in this "eternal" state, but the hole subsequently interacts with a small probe (whose existence was not anticipated in the original preparation), does this cause only a small leakage of Hawking radiation, or does the hole revert to ordinary evaporation? Another interesting question is the relative entropy of eternal and ordinary black holes.

Acknowledgments — The author thanks Roberto Casadio, Ted Jacobson and David Reeb for comments or discussions. This work is supported by the Department of Energy under grant DE-FG02-96ER40969.

---

[1] V. P. Frolov and I. D. Novikov, Black Hole Physics: Basic Concepts and New Developments, Kluwer Academic Publishers (1998).
[2] D. M. Eardley and S. B. Giddings, Phys. Rev. D **66**, 044011 (2002) [gr-qc/0201034]; S. D. H. Hsu, Phys. Lett. B **555**, 92 (2003) [hep-ph/0203154].
[3] To be more specific, take a spacelike slice (e.g., horizontal line) near the top of figure 2 (black hole Penrose diagram), but below the singularity. The time reversal of this state has the same values (and time derivatives) of metric and matter fields as a corresponding spacelike slice near the bottom of figure 1 (white hole diagram). If colliding particles can tunnel to the first (black hole) state, they can also tunnel to the corresponding white hole state (i.e., time reversal does not alter any quantum numbers). Note that because this is a quantum process, there is no singularity in the past of the white hole created. The spacetime with two colliding particles is simply patched onto the initial condition described by the spacelike slice from figure 1.
[4] S. D. H. Hsu and D. Reeb, Phys. Lett. B **658**, 244 (2008) [arXiv:0706.3239 [hep-th]]; Class. Quant. Grav. **25**, 235007 (2008) [arXiv:0803.4212 [hep-th]]; Mod. Phys. Lett. A **24**, 1875 (2009) [arXiv:0908.1265 [gr-qc]].
[5] S. W. Hawking, Phys. Rev. D **13**, 191 (1976).
[6] P. Frampton, S. D. H. Hsu, D. Reeb and T. W. Kephart, Class. Quant. Grav. **26**, 145005 (2009) [arXiv:0801.1847 [hep-th]].
[7] E. Poisson and W. Israel, Phys. Rev. Lett. **63**, 1663 (1989); Phys. Lett. B **233**, 74 (1989).
[8] S. W. Hawking, Phys. Rev. D **14**, 2460 (1976).

RELATIVITY

THE SPECIAL AND GENERAL THEORY

ALBERT EINSTEIN

DIGITAL REPRINT

Elegant Ebooks

COPYRIGHT INFORMATION

Book: *Relativity*: *The Special and General Theory*
Author: Albert Einstein, 1879–1955
First published: 1920

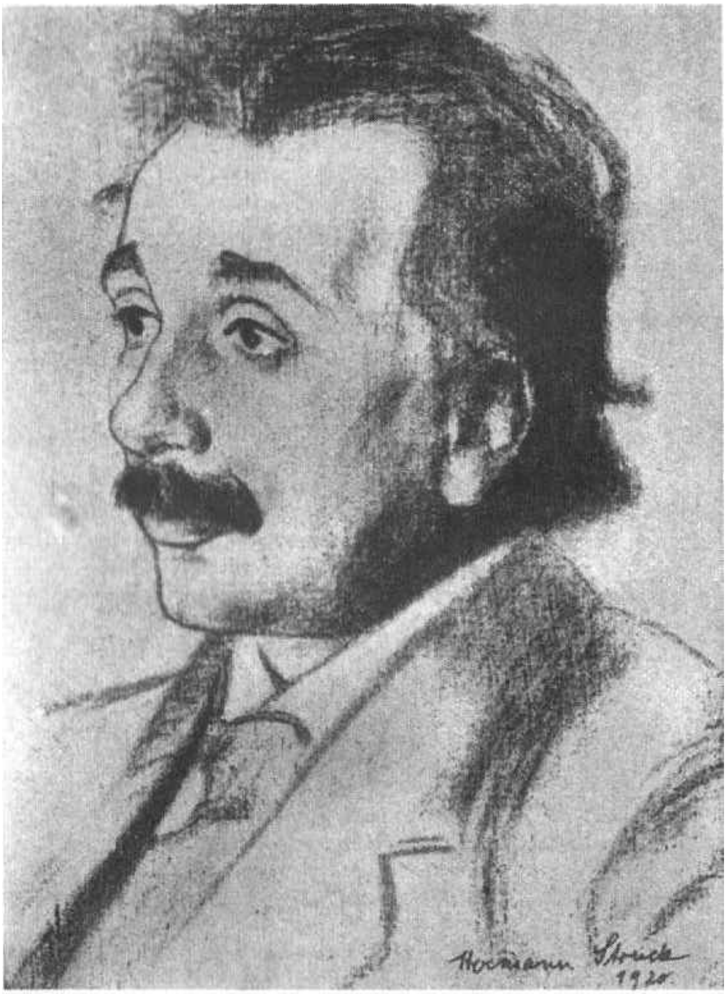
This PDF ebook was
created by José Menéndez.

NOTE ON THE TEXT

The text used in this ebook is from the first English translation, published in 1920, of *Über die spezielle und die allgemeine Relativitätstheorie*. A few misprints in the original text have been corrected, but to preserve all of the book, the original misprints are included in footnotes enclosed in square brackets and signed "J.M."

In addition, the line breaks and pagination of the original book have been reproduced, and the PDF's page numbers are linked to page images of the book from the Google Books Library Project for easy comparison. Please note: Google's images may not be available to people outside the U.S. and may be unavailable to U.S. residents at times.

Retaining the original pagination made it possible to include a fully functional (clickable) copy of the original index.



A. Einstein

RELATIVITY

THE SPECIAL AND GENERAL THEORY

BY

ALBERT EINSTEIN, Ph.D.
PROFESSOR OF PHYSICS IN THE UNIVERSITY OF BERLIN

TRANSLATED BY

ROBERT W. LAWSON, M.Sc.
UNIVERSITY OF SHEFFIELD

NEW YORK
HENRY HOLT AND COMPANY
1920

PREFACE

THE present book is intended, as far as possible, to give an exact insight into the theory of Relativity to those readers who, from a general scientific and philosophical point of view, are interested in the theory, but who are not conversant with the mathematical apparatus [1] of theoretical physics. The work presumes a standard of education corresponding to that of a university matriculation examination, and, despite the shortness of the book, a fair amount of patience and force of will on the part of the reader. The author has spared himself no pains in his endeavour to present the main ideas in the simplest and most intelligible form, and on the

[1] The mathematical fundaments of the special theory of relativity are to be found in the original papers of H. A. Lorentz, A. Einstein, H. Minkowski,[*] published under the title *Das Relativitätsprinzip* (The Principle of Relativity) in B. G. Teubner's collection of monographs *Fortschritte der mathematischen Wissenschaften* (Advances in the Mathematical Sciences), also in M. Laue's exhaustive book *Das Relativitäts prinzip* — published by Friedr. Vieweg & Son, Braunschweig. The general theory of relativity, together with the necessary parts of the theory of invariants, is dealt with in the author's book *Die Grundlagen der allgemeinen Relativitätstheorie* (The Foundations of the General Theory of Relativity) — Joh. Ambr. Barth, 1916; this book assumes some familiarity with the special theory of relativity.

[[*] Minkowski' — J.M.]

whole, in the sequence and connection in which they actually originated. In the interest of clearness, it appeared to me inevitable that I should repeat myself frequently, without paying the slightest attention to the elegance of the presentation. I adhered scrupulously to the precept of that brilliant theoretical physicist, L. Boltzmann, according to whom matters of elegance ought to be left to the tailor and to the cobbler. I make no pretence of having withheld from the reader difficulties which are inherent to the subject. On the other hand, I have purposely treated the empirical physical foundations of the theory in a "step-motherly" fashion, so that readers unfamiliar with physics may not feel like the wanderer who was unable to see the forest for trees. May the book bring some one a few happy hours of suggestive thought!

A. EINSTEIN

December, 1916

NOTE TO THE THIRD EDITION

IN the present year (1918) an excellent and detailed manual on the general theory of relativity, written by H. Weyl, was published by the firm Julius Springer (Berlin). This book, entitled *Raum — Zeit — Materie* (Space — Time — Matter), may be warmly recommended to mathematicians and physicists.

BIOGRAPHICAL NOTE

ALBERT EINSTEIN is the son of German-Jewish parents. He was born in 1879 in the town of Ulm, Würtemberg, Germany. His schooldays were spent in Munich, where he attended the *Gymnasium* until his sixteenth year. After leaving school at Munich, he accompanied his parents to Milan, whence he proceeded to Switzerland six months later to continue his studies.

From 1896 to 1900 Albert Einstein studied mathematics and physics at the Technical High School in Zurich, as he intended becoming a secondary school (*Gymnasium*) teacher. For some time afterwards he was a private tutor, and having meanwhile become naturalised, he obtained a post as engineer in the Swiss Patent Office in 1902, which position he occupied till 1909. The main ideas involved in the most important of Einstein's theories date back to this period. Amongst these may be mentioned: *The Special Theory of Relativity*, *Inertia of Energy*, *Theory of the Brownian Movement*, and the *Quantum-Law of the Emission and Absorption of Light* (1905). These were followed some years later by the

Theory of the Specific Heat of Solid Bodies, and the fundamental idea of the *General Theory of Relativity*.

During the interval 1909 to 1911 he occupied the post of Professor *Extraordinarius* at the University of Zurich, afterwards being appointed to the University of Prague, Bohemia, where he remained as Professor *Ordinarius* until 1912. In the latter year Professor Einstein accepted a similar chair at the *Polytechnikum*, Zurich, and continued his activities there until 1914, when he received a call to the Prussian Academy of Science, Berlin, as successor to Van't Hoff. Professor Einstein is able to devote himself freely to his studies at the Berlin Academy, and it was here that he succeeded in completing his work on the *General Theory of Relativity* (1915–17). Professor Einstein also lectures on various special branches of physics at the University of Berlin, and, in addition, he is Director of the Institute* for Physical Research of the *Kaiser Wilhelm Gesellschaft*.

Professor Einstein has been twice married. His first wife, whom he married at Berne in 1903, was a fellow-student from Serbia. There were two sons of this marriage, both of whom are living in Zurich, the elder being sixteen years of age. Recently Professor Einstein married a widowed cousin, with whom he is now living in Berlin.

R. W. L.

[* Institnte — J.M.]

TRANSLATOR'S NOTE

IN presenting this translation to the English-reading public, it is hardly necessary for me to enlarge on the Author's prefatory remarks, except to draw attention to those additions to the book which do not appear in the original.

At my request, Professor Einstein kindly supplied me with a portrait of himself, by one of Germany's most celebrated artists. Appendix III, on "The Experimental Confirmation of the General Theory of Relativity," has been written specially for this translation. Apart from these valuable additions to the book, I have included a biographical note on the Author, and, at the end of the book, an Index and a list of English references to the subject. This list, which is more suggestive than exhaustive, is intended as a guide to those readers who wish to pursue the subject farther.

I desire to tender my best thanks to my colleagues Professor S. R. Milner, D.Sc., and Mr. W. E. Curtis, A.R.C.Sc., F.R.A.S., also to my friend Dr. Arthur Holmes, A.R.C.Sc., F.G.S., of the Imperial College, for their kindness in reading through the manuscript, for helpful criticism, and for numerous suggestions. I owe an expression of thanks also to Messrs. Methuen for their ready counsel and advice, and for the care they have bestowed on the work during the course of its publication.

ROBERT W. LAWSON

THE PHYSICS LABORATORY
THE UNIVERSITY OF SHEFFIELD
June 12, 1920

CONTENTS

PART I

THE SPECIAL THEORY OF RELATIVITY

[* Relativ ty — J.M.]

[† "Theory" was changed to "Principle" in later editions. — J.M.]

RELATIVITY

PART I

THE SPECIAL THEORY OF RELATIVITY

I

PHYSICAL MEANING OF GEOMETRICAL PROPOSITIONS

IN your schooldays most of you who read this book made acquaintance with the noble building of Euclid's geometry, and you remember — perhaps with more respect than love — the magnificent structure, on the lofty staircase of which you were chased about for uncounted hours by conscientious teachers. By reason of your past experience, you would certainly regard every one with disdain who should pronounce even the most out-of-the-way proposition of this science to be untrue. But perhaps this feeling of proud certainty would leave you immediately if some one were to ask you: "What, then, do you mean by the assertion that these propositions are true?" Let us proceed to give this question a little consideration.

Geometry sets out from certain conceptions such as "plane," "point," and "straight line," with which we are able to associate more or less definite ideas, and from certain simple propositions (axioms) which, in virtue of these ideas, we are inclined to accept as "true." Then, on the basis of a logical process, the justification of which we feel ourselves compelled to admit, all remaining propositions are shown to follow from those axioms, *i.e.* they are proven. A proposition is then correct ("true") when it has been derived in the recognised manner from the axioms. The question of the "truth" of the individual geometrical propositions is thus reduced to one of the "truth" of the axioms. Now it has long been known that the last question is not only unanswerable by the methods of geometry, but that it is in itself entirely without meaning. We cannot ask whether it is true that only one straight line goes through two points. We can only say that Euclidean geometry deals with things called "straight lines," to each of which is ascribed the property of being uniquely determined by two points situated on it. The concept "true" does not tally with the assertions of pure geometry, because by the word "true" we are eventually in the habit of designating always the correspondence with a "real" object; geometry, however, is not concerned with the relation of the ideas involved in it to objects of experience, but only with the logical connection of these ideas among themselves.

It is not difficult to understand why, in spite of this, we feel constrained to call the propositions of geometry "true." Geometrical ideas correspond to more or less exact objects in nature, and these last are undoubtedly the exclusive cause of the genesis of those ideas. Geometry ought to refrain from such a course, in order to give to its structure the largest possible logical unity. The practice, for example, of seeing in a "distance" two marked positions on a practically rigid body is something which is lodged deeply in our habit of thought. We are accustomed further to regard three points as being situated on a straight line, if their apparent positions can be made to coincide for observation with one eye, under suitable choice of our place of observation.

If, in pursuance of our habit of thought, we now supplement the propositions of Euclidean geometry by the single proposition that two points on a practically rigid body always correspond to the same distance (line-interval), independently of any changes in position to which we may subject the body, the propositions of Euclidean geometry then resolve themselves into propositions on the possible relative position of practically rigid bodies.[1]

[1] It follows that a natural object is associated also with a straight line. Three points *A*, *B* and *C* on a rigid body thus lie in a straight line when, the points *A* and *C* being given, *B* is chosen such that the sum of the distances *AB* and *BC* is as short as possible. This incomplete suggestion will suffice for our present purpose.

Geometry which has been supplemented in this way is then to be treated as a branch of physics. We can now legitimately ask as to the "truth" of geometrical propositions interpreted in this way, since we are justified in asking whether these propositions are satisfied for those real things we have associated with the geometrical ideas. In less exact terms we can express this by saying that by the "truth" of a geometrical proposition in this sense we understand its validity for a construction with ruler and compasses.

Of course the conviction of the "truth" of geometrical propositions in this sense is founded exclusively on rather incomplete experience. For the present we shall assume the "truth" of the geometrical propositions, then at a later stage (in the general theory of relativity) we shall see that this "truth" is limited, and we shall consider the extent of its limitation.

II

THE SYSTEM OF CO–ORDINATES

ON the basis of the physical interpretation of distance which has been indicated, we are also in a position to establish the distance between two points on a rigid body by means of measurements. For this purpose we require a "distance" (rod *S*) which is to be used once and for all, and which we employ as a standard measure. If, now, *A* and *B* are two points on a rigid body, we can construct the line joining them according to the rules of geometry; then, starting from *A*, we can mark off the distance *S* time after time until we reach *B*. The number of these operations required is the numerical measure of the distance *AB*. This is the basis of all measurement of length.[1]

Every description of the scene of an event or of the position of an object in space is based on the specification of the point on a rigid body (body of reference) with which that event or object coin-

[1] Here we have assumed that there is nothing left over, *i.e.* that the measurement gives a whole number. This difficulty is got over by the use of divided measuring-rods, the introduction of which does not demand any fundamentally new method.

cides. This applies not only to scientific description, but also to everyday life. If I analyse the place specification "Trafalgar Square, London,"[1] I arrive at the following result. The earth is the rigid body to which the specification of place refers; "Trafalgar Square, London" is a well-defined point, to which a name has been assigned, and with which the event coincides in space.[2]

This primitive method of place specification deals only with places on the surface of rigid bodies, and is dependent on the existence of points on this surface which are distinguishable from each other. But we can free ourselves from both of these limitations without altering the nature of our specification of position. If, for instance, a cloud is hovering over Trafalgar Square, then we can determine its position relative to the surface of the earth by erecting a pole perpendicularly on the Square, so that it reaches the cloud. The length of the pole measured with the standard measuring-rod, combined with the specification of the position of the foot of the pole, supplies us with a complete place specification. On the basis

[1] I have chosen this as being more familiar to the English reader than the "Potsdamer Platz, Berlin," which is referred to in the original. (R. W. L.)

[2] It is not necessary here to investigate further the significance of the expression "coincidence in space." This conception is sufficiently obvious to ensure that differences of opinion are scarcely likely to arise as to its applicability in practice.

of this illustration, we are able to see the manner in which a refinement of the conception of position has been developed.

(*a*) We imagine the rigid body, to which the place specification is referred, supplemented in such a manner that the object whose position we require is reached by the completed rigid body.

(*b*) In locating the position of the object, we make use of a number (here the length of the pole measured with the measuring-rod) instead of designated points of reference.

(*c*) We speak of the height of the cloud even when the pole which reaches the cloud has not been erected. By means of optical observations of the cloud from different positions on the ground, and taking into account the properties of the propagation of light, we determine the length of the pole we should have required in order to reach the cloud.

From this consideration we see that it will be advantageous if, in the description of position, it should be possible by means of numerical measures to make ourselves independent of the existence of marked positions (possessing names) on the rigid body of reference. In the physics of measurement this is attained by the application of the Cartesian system of co-ordinates.

This consists of three plane surfaces perpendicular to each other and rigidly attached to a rigid body. Referred to a system of co-ordinates, the scene of any event will be determined (for the main part) by the specification of the lengths of the three perpendiculars or co-ordinates (x, y, z) which can be dropped from the scene of the event to those three plane surfaces. The lengths of these three perpendiculars can be determined by a series of manipulations with rigid measuring-rods performed according to the rules and methods laid down by Euclidean geometry.

In practice, the rigid surfaces which constitute the system of co-ordinates are generally not available; furthermore, the magnitudes of the co-ordinates are not actually determined by constructions with rigid rods, but by indirect means. If the results of physics and astronomy are to maintain their clearness, the physical meaning of specifications of position must always be sought in accordance with the above considerations.[1]

We thus obtain the following result: Every description of events in space involves the use of a rigid body to which such events have to be referred. The resulting relationship takes for granted that the laws of Euclidean geometry hold for "distances," the "distance" being represented physically by means of the convention of two marks on a rigid body.

[1] A refinement and modification of these views does not become necessary until we come to deal with the general theory of relativity, treated in the second part of this book.

III

SPACE AND TIME IN CLASSICAL MECHANICS

"THE purpose of mechanics is to describe how bodies change their position in space with time." I should load my conscience with grave sins against the sacred spirit of lucidity were I to formulate the aims of mechanics in this way, without serious reflection and detailed explanations. Let us proceed to disclose these sins.

It is not clear what is to be understood here by "position" and "space." I stand at the window of a railway carriage which is travelling uniformly, and drop a stone on the embankment, without throwing it. Then, disregarding the influence of the air resistance, I see the stone descend in a straight line. A pedestrian who observes the misdeed from the footpath notices that the stone falls to earth in a parabolic curve. I now ask: Do the "positions" traversed by the stone lie "in reality" on a straight line or on a parabola? Moreover, what is meant here by motion "in space"? From the considerations of the previous section the answer is self-evident. In the first place, we entirely shun the vague word "space," of which, we must honestly acknowledge, we cannot form the slightest conception, and we replace it by "motion relative to a practically rigid body of reference." The positions relative to the body of reference (railway carriage or embankment) have already been defined in detail in the preceding section. If instead of "body of reference" we insert "system of co-ordinates," which is a useful idea for mathematical description, we are in a position to say: The stone traverses a straight line relative to a system of co-ordinates rigidly attached to the carriage, but relative to a system of co-ordinates rigidly attached to the ground (embankment) it describes a parabola. With the aid of this example it is clearly seen that there is no such thing as an independently existing trajectory (lit. "path-curve" [1]), but only a trajectory relative to a particular body of reference.

In order to have a *complete* description of the motion, we must specify how the body alters its position *with time; i.e.* for every point on the trajectory it must be stated at what time the body is situated there. These data must be supplemented by such a definition of time that, in virtue of this definition, these time-values can be regarded essentially as magnitudes (results of measurements) capable of observation. If we take our stand on the ground of classical me-

[1] That is, a curve along which the body moves.

chanics, we can satisfy this requirement for our illustration in the following manner. We imagine two clocks of identical construction; the man at the railway-carriage window is holding one of them, and the man on the footpath the other. Each of the observers determines the position on his own reference-body occupied by the stone at each tick of the clock he is holding in his hand. In this connection we have not taken account of the inaccuracy involved by the finiteness of the velocity of propagation of light. With this and with a second difficulty prevailing here we shall have to deal in detail later.

IV

THE GALILEIAN SYSTEM OF CO–ORDINATES

AS is well known, the fundamental law of the mechanics of Galilei-Newton, which is known as the *law of inertia*, can be stated thus: A body removed sufficiently far from other bodies continues in a state of rest or of uniform motion in a straight line. This law not only says something about the motion of the bodies, but it also indicates the reference-bodies or systems of co-ordinates, permissible in mechanics, which can be used in mechanical description. The visible fixed stars are bodies for which the law of inertia certainly holds to a high degree of approximation. Now if we use a system of co-ordinates which is rigidly attached to the earth, then, relative to this system, every fixed star describes a circle of immense radius in the course of an astronomical day, a result which is opposed to the statement of the law of inertia. So that if we adhere to this law we must refer these motions only to systems of co-ordinates relative to which the fixed stars do not move in a circle. A system of co-ordinates of which the state of motion is such that the law of inertia holds relative to it is called a "Galileian system of co-ordinates." The laws of the mechanics of Galilei-Newton can be regarded as valid only for a Galileian system of co-ordinates.

V

THE PRINCIPLE OF RELATIVITY (IN THE RESTRICTED SENSE)

IN order to attain the greatest possible clearness, let us return to our example of the railway carriage supposed to be travelling uniformly. We call its motion a uniform translation ("uniform" because it is of constant velocity and direction, "translation" because although the carriage changes its position relative to the embankment yet it does not rotate in so doing). Let us imagine a raven flying through the air in such a manner that its motion, as observed from the embankment, is uniform and in a straight line. If we were to observe the flying raven from the moving railway carriage, we should find that the motion of the raven would be one of different velocity and direction, but that it would still be uniform and in a straight line. Expressed in an abstract manner we may say: If a mass m is moving uniformly in a straight line with respect to a co-ordinate system K, then it will also be moving uniformly and in a straight line relative to a second co-ordinate system K', provided that the latter is executing a uniform translatory motion with respect to *K*. In accordance with the discussion contained in the preceding section, it follows that:

If *K* is a Galileian co-ordinate system, then every other co-ordinate system *K'* is a Galileian one, when, in relation to *K*, it is in a condition of uniform motion of translation. Relative to *K'* the mechanical laws of Galilei-Newton hold good exactly as they do with respect to *K*.

We advance a step farther in our generalisation when we express the tenet thus: If, relative to *K*, *K'* is a uniformly moving co-ordinate system devoid of rotation, then natural phenomena run their course with respect to *K'* according to exactly the same general laws as with respect to *K*. This statement is called the *principle of relativity* (in the restricted sense).

As long as one was convinced that all natural phenomena were capable of representation with the help of classical mechanics, there was no need to doubt the validity of this principle of relativity. But in view of the more recent development of electrodynamics and optics it became more and more evident that classical mechanics affords an insufficient foundation for the physical description of all natural phenomena. At this juncture the question of the validity of the principle of relativity became ripe for discussion, and it did not appear impossible that the answer to this question might be in the negative.

Nevertheless, there are two general facts which at the outset speak very much in favour of the validity of the principle of relativity. Even though classical mechanics does not supply us with a sufficiently broad basis for the theoretical presentation of all physical phenomena, still we must grant it a considerable measure of "truth," since it supplies us with the actual motions of the heavenly bodies with a delicacy of detail little short of wonderful. The principle of relativity must therefore apply with great accuracy in the domain of *mechanics*. But that a principle of such broad generality should hold with such exactness in one domain of phenomena, and yet should be invalid for another, is *a priori* not very probable.

We now proceed to the second argument, to which, moreover, we shall return later. If the principle of relativity (in the restricted sense) does not hold, then the Galileian co-ordinate systems K, K', K'', etc., which are moving uniformly relative to each other, will not be *equivalent* for the description of natural phenomena. In this case we should be constrained to believe that natural laws are capable of being formulated in a particularly simple manner, and of course only on condition that, from amongst all possible Galileian co-ordinate systems, we should have chosen *one* (K_0) of a particular state of motion as our body of reference. We should then be justified (because of its merits for the description of natural phenomena) in calling this system "absolutely at rest," and all other Galileian systems K "in motion." If, for instance, our embankment were the system K_0, then our railway carriage would be a system K, relative to which less simple laws would hold than with respect to K_0. This diminished simplicity would be due to the fact that the carriage K would be in motion (*i.e.* "really") with respect to K_0. In the general laws of nature which have been formulated with reference to K, the magnitude and direction of the velocity of the carriage would necessarily play a part. We should expect, for instance, that the note emitted by an organ-pipe placed with its axis parallel to the direction of travel would be different from that emitted if the axis of the pipe were placed perpendicular to this direction. Now in virtue of its motion in an orbit round the sun, our earth is comparable with a railway carriage travelling with a velocity of about 30 kilometres per second. If the principle of relativity were not valid we should therefore expect that the direction of motion of the earth at any moment would enter into the laws of nature, and also that physical systems in their behaviour would be dependent on the orientation in space with respect to the earth. For owing to the alteration in direction of the velocity of rotation* of the earth in the course of a year, the earth cannot be at rest relative to the hypothetical system K_0 throughout the whole year. However, the most careful observations have never revealed such anisotropic properties in terrestrial physical space, *i.e.* a physical non-equivalence of different directions. This is a very powerful argument in favour of the principle of relativity.

[* The word "rotation" was correctly changed to "revolution" in later editions. — J.M.]

VI

THE THEOREM OF THE ADDITION OF VELOCITIES EMPLOYED IN CLASSICAL MECHANICS

LET us suppose our old friend the railway carriage to be travelling along the rails with a constant velocity v, and that a man traverses the length of the carriage in the direction of travel with a velocity w. How quickly, or, in other words, with what velocity W does the man advance relative to the embankment during the process? The only possible answer seems to result from the following consideration: If the man were to stand still for a second, he would advance relative to the embankment through a distance v equal numerically to the velocity of the carriage. As a consequence of his walking, however, he traverses an additional distance w relative to the carriage, and hence also relative to the embankment, in this second, the distance w being numerically equal to the velocity with which he is walking. Thus in total he covers the distance $W = v + w$ relative to the embankment in the second considered. We shall see later that this result, which expresses the theorem of the addition of velocities employed in classical mechanics, cannot be maintained; in other words, the law that we have just written down does not hold in reality. For the time being, however, we shall assume its correctness.

VII

THE APPARENT INCOMPATIBILITY OF THE LAW OF PROPAGATION OF LIGHT WITH THE PRINCIPLE OF RELATIVITY

THERE is hardly a simpler law in physics than that according to which light is propagated in empty space. Every child at school knows, or believes he knows, that this propagation takes place in straight lines with a velocity $c = 300,000$ km./sec. At all events we know with great exactness that this velocity is the same for all colours, because if this were not the case, the minimum of emission would not be observed simultaneously for different colours during the eclipse of a fixed star by its dark neighbour. By means of similar considerations based on observations of double stars, the Dutch astronomer De Sitter was also able to show that the velocity of propagation of light cannot depend on the velocity of motion of the body emitting the light. The assumption that this velocity of propagation is dependent on the direction "in space" is in itself improbable.

In short, let us assume that the simple law of the constancy of the velocity of light c (in vacuum) is justifiably believed by the child at school. Who would imagine that this simple law has plunged the conscientiously thoughtful physicist into the greatest intellectual difficulties? Let us consider how these difficulties arise.

Of course we must refer the process of the propagation of light (and indeed every other process) to a rigid reference-body (co-ordinate system). As such a system let us again choose our embankment. We shall imagine the air above it to have been removed. If a ray of light be sent along the embankment, we see from the above that the tip of the ray will be transmitted with the velocity c relative to the embankment. Now let us suppose that our railway carriage is again travelling along the railway lines with the velocity v, and that its direction is the same as that of the ray of light, but its velocity of course much less. Let us inquire about the velocity of propagation of the ray of light relative to the carriage. It is obvious that we can here apply the consideration of the previous section, since the ray of light plays the part of the man walking along relatively to the carriage. The velocity W of the man relative to the embankment is here replaced by the velocity of light relative to the embankment. w is the required velocity of light with respect to the carriage, and we have

$$w = c - v.$$

The velocity of propagation of a ray of light relative to the carriage thus comes out smaller than *c*.

But this result comes into conflict with the principle of relativity set forth in Section V. For, like every other general law of nature, the law of the transmission of light *in vacuo* must, according to the principle of relativity, be the same for the railway carriage as reference-body as when the rails are the body of reference. But, from our above consideration, this would appear to be impossible. If every ray of light is propagated relative to the embankment with the velocity *c*, then for this reason it would appear that another law of propagation of light must necessarily hold with respect to the carriage — a result contradictory to the principle of relativity.

In view of this dilemma there appears to be nothing else for it than to abandon either the principle of relativity or the simple law of the propagation of light *in vacuo*. Those of you who have carefully followed the preceding discussion are almost sure to expect that we should retain the principle of relativity, which appeals so convincingly to the intellect because it is so natural and simple. The law of the propagation of light *in vacuo* would then have to be replaced by a more complicated law conformable to the principle of relativity. The development of theoretical physics shows, however, that we cannot pursue this course. The epoch-making theoretical investigations of H. A. Lorentz on the electrodynamical and optical phenomena connected with moving bodies show that experience in this domain leads conclusively to a theory of electromagnetic phenomena, of which the law of the constancy of the velocity of light *in vacuo* is a necessary consequence. Prominent theoretical physicists were therefore more inclined to reject the principle of relativity, in spite of the fact that no empirical data had been found which were contradictory to this principle.

At this juncture the theory of relativity entered the arena. As a result of an analysis of the physical conceptions of time and space, it became evident that *in reality there is not the least incompatibility between the principle of relativity and the law of propagation of light*, and that by systematically holding fast to both these laws a logically rigid theory could be arrived at. This theory has been called the *special theory of relativity* to distinguish it from the extended theory, with which we shall deal later. In the following pages we shall present the fundamental ideas of the special theory of relativity.

VIII

ON THE IDEA OF TIME IN PHYSICS

LIGHTNING has struck the rails on our railway embankment at two places *A* and *B* far distant from each other. I make the additional assertion that these two lightning flashes occurred simultaneously. If now I ask you whether there is sense in this statement, you will answer my question with a decided "Yes." But if I now approach you with the request to explain to me the sense of the statement more precisely, you find after some consideration that the answer to this question is not so easy as it appears at first sight.

After some time perhaps the following answer would occur to you: "The significance of the statement is clear in itself and needs no further explanation; of course it would require some consideration if I were to be commissioned to determine by observations whether in the actual case the two events took place simultaneously or not." I cannot be satisfied with this answer for the following reason. Supposing that as a result of ingenious considerations an able meteorologist were to discover that the lightning must always strike the places *A* and *B* simultaneously, then we should be faced with the task of testing whether or not this theoretical result is in accordance with the reality. We encounter the same difficulty with all physical statements in which the conception "simultaneous" plays a part. The concept does not exist for the physicist until he has the possibility of discovering whether or not it is fulfilled in an actual case. We thus require a definition of simultaneity such that this definition supplies us with the method by means of which, in the present case, he can decide by experiment whether or not both the lightning strokes occurred simultaneously. As long as this requirement is not satisfied, I allow myself to be deceived as a physicist (and of course the same applies if I am not a physicist), when I imagine that I am able to attach a meaning to the statement of simultaneity. (I would ask the reader not to proceed farther until he is fully convinced on this point.)

After thinking the matter over for some time you then offer the following suggestion with which to test simultaneity. By measuring along the rails, the connecting line *AB* should be measured up and an observer placed at the mid-point *M* of the distance *AB*. This observer should be supplied with an arrangement (*e.g.* two mirrors inclined at 90°) which allows him visually to observe both places A and B at the same time. If the observer perceives the two flashes of lightning at the same time, then they are simultaneous.

I am very pleased with this suggestion, but for all that I cannot regard the matter as quite settled, because I feel constrained to raise the following objection: "Your definition would certainly be right, if I only knew that the light by means of which the observer at M perceives the lightning flashes travels along the length $A \longrightarrow M$ with the same velocity as along the length $B \longrightarrow M$. But an examination of this supposition would only be possible if we already had at our disposal the means of measuring time. It would thus appear as though we were moving here in a logical circle."

After further consideration you cast a somewhat disdainful glance at me — and rightly so — and you declare: "I maintain my previous definition nevertheless, because in reality it assumes absolutely nothing about light. There is only *one* demand to be made of the definition of simultaneity, namely, that in every real case it must supply us with an empirical decision as to whether or not the conception that has to be defined is fulfilled. That my definition satisfies this demand is indisputable. That light requires the same time to traverse the path $A \longrightarrow M$ as for the path $B \longrightarrow M$ is in reality neither a *supposition nor a hypothesis* about the physical nature of light, but a *stipulation* which I can make of my own freewill in order to arrive at a definition of simultaneity."

It is clear that this definition can be used to give an exact meaning not only to *two* events, but to as many events as we care to choose, and independently of the positions of the scenes of the events with respect to the body of reference [1] (here the railway embankment). We are thus led also to a definition of "time" in physics. For this purpose we suppose that clocks of identical construction are placed at the points *A*, *B* and *C* of the railway line (co-ordinate system), and that they are set in such a manner that the positions of their pointers are simultaneously (in the above sense) the same. Under these conditions we understand by the "time" of an event the reading (position of the hands) of that one of these clocks which is in the immediate vicinity (in space) of the event. In this manner a time-value is associated with every event which is essentially capable of observation.

This stipulation contains a further physical

[1] We suppose further that, when three events *A*, *B* and *C* take place in different places in such a manner that, if *A* is simultaneous with *B*, and *B* is simultaneous with *C* (simultaneous in the sense of the above definition), then the criterion for the simultaneity of the pair of events *A*, *C* is also satisfied. This assumption is a physical hypothesis about the law of propagation of light; it must certainly be fulfilled if we are to maintain the law of the constancy of the velocity of light *in vacuo*.

hypothesis, the validity of which will hardly be doubted without empirical evidence to the contrary. It has been assumed that all these clocks go *at the same rate* if they are of identical construction. Stated more exactly: When two clocks arranged at rest in different places of a reference-body are set in such a manner that a *particular* position of the pointers of the one clock is *simultaneous* (in the above sense) with the *same* position of the pointers of the other clock, then identical "settings" are always simultaneous (in the sense of the above definition).

IX

THE RELATIVITY OF SIMULTANEITY

UP to now our considerations have been referred to a particular body of reference, which we have styled a "railway embankment." We suppose a very long train travelling along the rails with the constant velocity *v* and in the direction indicated in Fig. 1. People travelling in this train will with advantage use the train as a rigid reference-body (co-ordinate system); they regard all events in reference to

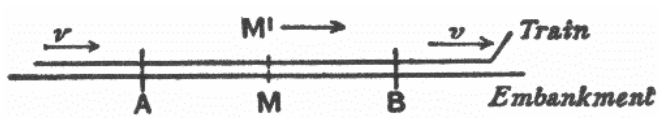


FIG. 1.

the train. Then every event which takes place along the line also takes place at a particular point of the train. Also the definition of simultaneity can be given relative to the train in exactly the same way as with respect to the embankment. As a natural consequence, however, the following question arises:

Are two events (*e.g.* the two strokes of lightning *A* and *B*) which are simultaneous *with reference to* *the railway embankment* also simultaneous *relatively to the train?* We shall show directly that the answer must be in the negative.

When we say that the lightning strokes A and B are simultaneous with respect to the embankment, we mean: the rays of light emitted at the places A and B, where the lightning occurs, meet each other at the mid-point M of the length $A \longrightarrow B$ of the embankment. But the events A and B also correspond to positions A and B on the train. Let M' be the mid-point of the distance $A \longrightarrow B$ on the travelling train. Just when the flashes [1] of lightning occur, this point M' naturally coincides with the point M, but it moves towards the right in the diagram with the velocity v of the train. If an observer sitting in the position M' in the train did not possess this velocity, then he would remain permanently at M, and the light rays emitted by the flashes of lightning A and B would reach him simultaneously, *i.e.* they would meet just where he is situated. Now in reality (considered with reference to the railway embankment) he is hastening towards the beam of light coming from B, whilst he is riding on ahead of the beam of light coming from A. Hence the observer will see the beam of light emitted from B earlier than he will see that emitted from A. Observers who take the railway train as their reference-body

[1] As judged from the embankment.

must therefore come to the conclusion that the lightning flash *B* took place earlier than the lightning flash *A*. We thus arrive at the important result:

Events which are simultaneous with reference to the embankment are not simultaneous with respect to the train, and *vice versa* (relativity of simultaneity). Every reference-body (co-ordinate system) has its own particular time; unless we are told the reference-body to which the statement of time refers, there is no meaning in a statement of the time of an event.

Now before the advent of the theory of relativity it had always tacitly been assumed in physics that the statement of time had an absolute significance, *i.e.* that it is independent of the state of motion of the body of reference. But we have just seen that this assumption is incompatible with the most natural definition of simultaneity; if we discard this assumption, then the conflict between the law of the propagation of light *in vacuo* and the principle of relativity (developed in Section VII) disappears.

We were led to that conflict by the considerations of Section VI, which are now no longer tenable. In that section we concluded that the man in the carriage, who traverses the distance *w per second* relative to the carriage, traverses the same distance also with respect to the embankment *in each second* of time. But, according to the foregoing considerations, the time required by a particular occurrence with respect to the carriage must not be considered equal to the duration of the same occurrence as judged from the embankment (as reference-body). Hence it cannot be contended that the man in walking travels the distance *w* relative to the railway line in a time which is equal to one second as judged from the embankment.

Moreover, the considerations of Section VI are based on yet a second assumption, which, in the light of a strict consideration, appears to be arbitrary, although it was always tacitly made even before the introduction of the theory of relativity.

X

ON THE RELATIVITY OF THE CONCEPTION OF DISTANCE

LET us consider two particular points on the train [1] travelling along the embankment with the velocity v, and inquire as to their distance apart. We already know that it is necessary to have a body of reference for the measurement of a distance, with respect to which body the distance can be measured up. It is the simplest plan to use the train itself as the reference-body (co-ordinate system). An observer in the train measures the interval by marking off his measuring-rod in a straight line (*e.g.* along the floor of the carriage) as many times as is necessary to take him from the one marked point to the other. Then the number which tells us how often the rod has to be laid down is the required distance.

It is a different matter when the distance has to be judged from the railway line. Here the following method suggests itself. If we call A' and B' the two points on the train whose distance apart is required, then both of these points are

[1] *e.g.* the middle of the first and of the hundredth carriage.

moving with the velocity v along the embankment. In the first place we require to determine the points A and B of the embankment which are just being passed by the two points A' and B' at a particular time t — judged from the embankment. These points A and B of the embankment can be determined by applying the definition of time given in Section VIII. The distance between these points A and B is then measured by repeated application of the measuring-rod along the embankment.

A priori it is by no means certain that this last measurement will supply us with the same result as the first. Thus the length of the train as measured from the embankment may be different from that obtained by measuring in the train itself. This circumstance leads us to a second objection which must be raised against the apparently obvious consideration of Section VI. Namely, if the man in the carriage covers the distance w in a unit of time — *measured from the train*, — then this distance — *as measured from the embankment* — is not necessarily also equal to w.

XI

THE LORENTZ TRANSFORMATION

THE results of the last three sections show that the apparent incompatibility of the law of propagation of light with the principle of relativity (Section VII) has been derived by means of a consideration which borrowed two unjustifiable hypotheses from classical mechanics; these are as follows:

(1) The time-interval (time) between two events is independent of the condition of motion of the body of reference.

(2) The space-interval (distance) between two points of a rigid body is independent of the condition of motion of the body of reference.

If we drop these hypotheses, then the dilemma of Section VII disappears, because the theorem of the addition of velocities derived in Section VI becomes invalid. The possibility presents itself that the law of the propagation of light *in vacuo* may be compatible with the principle of relativity, and the question arises: How have we to modify the considerations of Section VI in order to remove the apparent disagreement between these two fundamental results of experience? This question leads to a general one. In the discussion of Section VI we have to do with places and times relative both to the train and to the embankment. How are we to find the place and time of an event in relation to the train, when we know the place and time of the event with respect to the railway embankment? Is there a thinkable answer to this question of such a nature that the law of transmission of light *in vacuo* does not contradict the principle of relativity? In other words: Can we conceive of a relation between place and time of the individual events relative to both reference-bodies, such that every ray of light possesses the velocity of transmission *c* relative to the embankment and relative to the train? This question leads to a quite definite positive answer, and to a perfectly definite transformation law for the space-time magnitudes of an event when changing over from one body of reference to another.

Before we deal with this, we shall introduce the following incidental consideration. Up to the present we have only considered events taking place along the embankment, which had mathematically to assume the function of a straight line. In the manner indicated in Section II we can imagine this reference-body supplemented laterally and in a vertical direction by means of a framework of rods, so that an event which takes place anywhere can be localised with reference to this framework. Similarly, we can imagine the train travelling with the velocity *v* to be continued across the whole of space, so that every event, no matter how far off it may be, could also be localised with respect to the second framework. Without committing any fundamental error, we can disregard the fact that in reality these frameworks would continually interfere with each other, owing to the impenetrability of solid bodies. In every such framework we imagine three surfaces perpendicular to each other marked out, and designated as "co-ordinate planes" ("co-ordinate system"). A co-ordinate system K then corresponds to the embankment, and a co-ordinate system K' to the train. An event, wherever it may have taken place, would be fixed in space with respect to K by the three perpendiculars x, y, z on the co-ordinate planes, and with regard to time by a time-value t. Relative to K', *the same event* would be fixed in respect of space and time by corresponding values x', y', z', t', which of course are not identical with x, y, z, t. It has already been set forth in detail how these magnitudes are to be regarded as results of physical measurements.

Obviously our problem can be exactly formulated in the following manner. What are the values x', y', z', t' of an event with respect to K', when the magnitudes x, y, z, t, of the same event with respect to K are given? The relations must be so chosen that the law of the transmission of light *in vacuo* is satisfied for one and the same ray of light (and of course for every ray) with respect to K and K'. For the relative orientation in space of the co-ordinate systems indicated in the diagram (Fig. 2), this problem is solved by means of the equations:

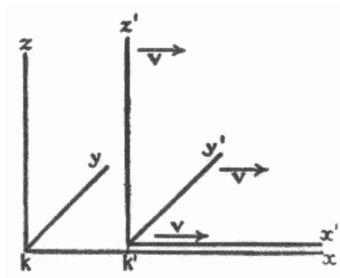


FIG. 2.

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$y' = y$$

$$z' = z$$

$$t' = \frac{t - \frac{v}{c^2} \cdot x}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

This system of equations is known as the "Lorentz transformation." [1]

If in place of the law of transmission of light we had taken as our basis the tacit assumptions of the older mechanics as to the absolute character

[1] A simple derivation of the Lorentz transformation is given in Appendix I.

of times and lengths, then instead of the above we should have obtained the following equations:

$$\begin{aligned} x' &= x - vt \\ y' &= y \\ z' &= z \\ t' &= t. \end{aligned}$$

This system of equations is often termed the "Galilei transformation." The Galilei transformation can be obtained from the Lorentz transformation by substituting an infinitely large value for the velocity of light *c* in the latter transformation.

Aided by the following illustration, we can readily see that, in accordance with the Lorentz transformation, the law of the transmission of light *in vacuo* is satisfied both for the reference-body *K* and for the reference-body *K'*. A light-signal is sent along the positive *x*-axis, and this light-stimulus advances in accordance with the equation

$$x = ct,$$

i.e. with the velocity *c*. According to the equations of the Lorentz transformation, this simple relation between *x* and *t* involves a relation between *x'* and *t'*. In point of fact, if we substitute for *x* the value *ct* in the first and fourth equations of the Lorentz transformation, we obtain:

$$x' = \frac{(c-v)t}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$t' = \frac{\left(1 - \frac{v}{c}\right)t}{\sqrt{1 - \frac{v^2}{c^2}}},$$

from which, by division, the expression

$$x' = ct'$$

immediately follows. If referred to the system K', the propagation of light takes place according to this equation. We thus see that the velocity of transmission relative to the reference-body K' is also equal to c. The same result is obtained for rays of light advancing in any other direction whatsoever. Of course this is not surprising, since the equations of the Lorentz transformation were derived conformably to this point of view.

XII

THE BEHAVIOUR OF MEASURING–RODS AND CLOCKS IN MOTION

I PLACE a metre-rod in the x'-axis of K' in such a manner that one end (the beginning) coincides with the point $x' = 0$, whilst the other end (the end of the rod) coincides with the point $x' = 1$. What is the length of the metre-rod relatively to the system K? In order to learn this, we need only ask where the beginning of the rod and the end of the rod lie with respect to K at a particular time t of the system K. By means of the first equation of the Lorentz transformation the values of these two points at the time $t = 0$ can be shown to be

$$x_{(\text{beginning of rod})} = 0 \cdot \sqrt{1 - \frac{v^2}{c^2}}$$

$$x_{(\text{end of rod})} = 1 \cdot \sqrt{1 - \frac{v^2}{c^2}},$$

the distance between the points being $\sqrt{1 - \frac{v^2}{c^2}}$. But the metre-rod is moving with the velocity v relative to K. It therefore follows that the length of a rigid metre-rod moving in the direction of its length with a velocity v is $\sqrt{1 - v^2/c^2}$ of a metre. The rigid rod is thus shorter when in motion than when at rest, and the more quickly it is moving, the shorter is the rod. For the velocity $v = c$ we should have $\sqrt{1 - v^2/c^2} = 0$, and for still greater velocities the square-root becomes imaginary. From this we conclude that in the theory of relativity the velocity c plays the part of a limiting velocity, which can neither be reached nor exceeded by any real body.

Of course this feature of the velocity c as a limiting velocity also clearly follows from the equations of the Lorentz transformation, for these become meaningless if we choose values of v greater than c.

If, on the contrary, we had considered a metre-rod at rest in the x-axis with respect to K, then we should have found that the length of the rod as judged from K' would have been $\sqrt{1 - v^2/c^2}$; this is quite in accordance with the principle of relativity which forms the basis of our considerations.

A priori it is quite clear that we must be able to learn something about the physical behaviour of measuring-rods and clocks from the equations of transformation, for the magnitudes x, y, z, t, are nothing more nor less than the results of measurements obtainable by means of measuring-rods and clocks. If we had based our considerations on the Galilei transformation we should not have obtained a contraction of the rod as a consequence of its motion.

Let us now consider a seconds-clock which is permanently situated at the origin ($x' = 0$) of K'. $t' = 0$ and $t' = 1$ are two successive ticks of this clock. The first and fourth equations of the Lorentz transformation give for these two ticks:

$$t = 0$$

and

$$t = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

As judged from K, the clock is moving with the velocity v; as judged from this reference-body, the time which elapses between two strokes of the clock is not one second, but $\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ seconds, *i.e.* a somewhat larger time. As a consequence of its motion the clock goes more slowly than when at rest. Here also the velocity c plays the part of an unattainable limiting velocity.

XIII

THEOREM OF THE ADDITION OF VELOCITIES. THE EXPERIMENT OF FIZEAU

NOW in practice we can move clocks and measuring-rods only with velocities that are small compared with the velocity of light; hence we shall hardly be able to compare the results of the previous section directly with the reality. But, on the other hand, these results must strike you as being very singular, and for that reason I shall now draw another conclusion from the theory, one which can easily be derived from the foregoing considerations, and which has been most elegantly confirmed by experiment.

In Section VI we derived the theorem of the addition of velocities in one direction in the form which also results from the hypotheses of classical mechanics. This theorem can also be deduced readily from the Galilei transformation (Section XI). In place of the man walking inside the carriage, we introduce a point moving relatively to the co-ordinate system K' in accordance with the equation

$$x' = wt'.$$

By means of the first and fourth equations of the

Galilei transformation we can express x' and t' in terms of x and t, and we then obtain

$$x = (v + w)t.$$

This equation expresses nothing else than the law of motion of the point with reference to the system K (of the man with reference to the embankment). We denote this velocity by the symbol W, and we then obtain, as in Section VI,

$$W = v + w \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (A).$$

But we can carry out this consideration just as well on the basis of the theory of relativity. In the equation

$$x' = wt'$$

we must then express x' and t' in terms of x and t, making use of the first and fourth equations of the *Lorentz transformation*. Instead of the equation (A) we then obtain the equation

$$W = \frac{v + w}{1 + \frac{vw}{c^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (B),$$

which corresponds to the theorem of addition for velocities in one direction according to the theory of relativity. The question now arises as to which of these two theorems is the better in accord with experience. On this point we are enlightened by a most important experiment which the brilliant physicist Fizeau performed more than half a century ago, and which has been repeated since then by some of the best experimental physicists, so that there can be no doubt about its result. The experiment is concerned with the following question. Light travels in a motionless liquid with a particular velocity *w*. How quickly does it travel in the direction of the arrow in the tube *T* (see the accompanying diagram, Fig. 3) when the liquid above mentioned is flowing through the tube with a velocity *v*?

In accordance with the principle of relativity we shall certainly have to take for granted that the propagation of light always takes place with the same velocity *w with respect to the liquid*, whether the latter is in motion with reference to other bodies or not. The velocity of light relative to the liquid and the velocity of the latter relative to the tube are thus known, and we require the velocity of light relative to the tube.

It is clear that we have the problem of Section VI again before us. The tube plays the part of

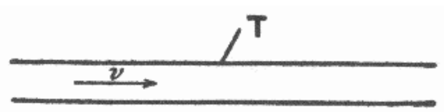


FIG. 3

the railway embankment or of the co-ordinate system *K*, the liquid plays the part of the carriage or of the co-ordinate system *K'*, and finally, the light plays the part of the man walking along the carriage, or of the moving point in the present section. If we denote the velocity of the light relative to the tube by W, then this is given by the equation (A) or (B), according as the Galilei transformation or the Lorentz transformation corresponds to the facts. Experiment [1] decides in favour of equation (B) derived from the theory of relativity, and the agreement is, indeed, very exact. According to recent and most excellent measurements by Zeeman, the influence of the velocity of flow v on the propagation of light is represented by formula (B) to within one per cent.

Nevertheless we must now draw attention to the fact that a theory of this phenomenon was given by H. A. Lorentz long before the statement of the theory of relativity. This theory was of a purely electrodynamical nature, and was obtained by the use of particular hypotheses as to the electromagnetic structure of matter. This circumstance, however, does not in the least diminish the conclusiveness of the experiment as a crucial test in favour of the theory of relativity, for the

[1] Fizeau found $W = w + v\left(1 - \frac{1}{n^2}\right)$, where $n = \frac{c}{w}$ is the index of refraction of the liquid. On the other hand, owing to the smallness of $\frac{vw}{c^2}$ as compared with 1, we can replace (B) in the first place by $W = (w + v)\left(1 - \frac{vw}{c^2}\right)$, or to the same order of approximation by $w + v\left(1 - \frac{1}{n^2}\right)$, which agrees with Fizeau's result.

electrodynamics of Maxwell-Lorentz, on which the original theory was based, in no way opposes the theory of relativity. Rather has the latter been developed from electrodynamics as an astoundingly simple combination and generalisation of the hypotheses, formerly independent of each other, on which electrodynamics was built.

XIV

THE HEURISTIC VALUE OF THE THEORY OF RELATIVITY

OUR train of thought in the foregoing pages can be epitomised in the following manner. Experience has led to the conviction that, on the one hand, the principle of relativity holds true, and that on the other hand the velocity of transmission of light *in vacuo* has to be considered equal to a constant c. By uniting these two postulates we obtained the law of transformation for the rectangular co-ordinates x, y, z and the time t of the events which constitute the processes of nature. In this connection we did not obtain the Galilei transformation, but, differing from classical mechanics, the *Lorentz transformation*.

The law of transmission of light, the acceptance of which is justified by our actual knowledge, played an important part in this process of thought. Once in possession of the Lorentz transformation, however, we can combine this with the principle of relativity, and sum up the theory thus:

Every general law of nature must be so constituted that it is transformed into a law of exactly the same form when, instead of the spacetime variables x, y, z, t of the original co-ordinate system K, we introduce new space-time variables x', y', z', t' of a co-ordinate system K'. In this connection the relation between the ordinary and the accented magnitudes is given by the Lorentz transformation. Or, in brief: General laws of nature are co-variant with respect to Lorentz transformations.

This is a definite mathematical condition that the theory of relativity demands of a natural law, and in virtue of this, the theory becomes a valuable heuristic aid in the search for general laws of nature. If a general law of nature were to be found which did not satisfy this condition, then at least one of the two fundamental assumptions of the theory would have been disproved. Let us now examine what general results the latter theory has hitherto evinced.

XV

GENERAL RESULTS OF THE THEORY

IT is clear from our previous considerations that the (special) theory of relativity has grown out of electrodynamics and optics. In these fields it has not appreciably altered the predictions of theory, but it has considerably simplified the theoretical structure, *i.e.* the derivation of laws, and — what is incomparably more important — it has considerably reduced the number of independent hypotheses forming the basis of theory. The special theory of relativity has rendered the Maxwell-Lorentz theory so plausible, that the latter would have been generally accepted by physicists even if experiment had decided less unequivocally in its favour.

Classical mechanics required to be modified before it could come into line with the demands of the special theory of relativity. For the main part, however, this modification affects only the laws for rapid motions, in which the velocities of matter v are not very small as compared with the velocity of light. We have experience of such rapid motions only in the case of electrons and ions; for other motions the variations from the laws of classical mechanics are too small to make themselves evident in practice. We shall not consider the motion of stars until we come to speak of the general theory of relativity. In accordance with the theory of relativity the kinetic energy of a material point of mass m is no longer given by the well-known expression

$$m\frac{v^2}{2},$$

but by the expression

$$\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}.^*$$

This expression approaches infinity as the velocity v approaches the velocity of light c. The velocity must therefore always remain less than c, however great may be the energies used to produce the acceleration. If we develop the expression for the kinetic energy in the form of a series, we obtain

$$mc^2 + m\frac{v^2}{2} + \frac{3}{8}m\frac{v^4}{c^2} + \ . \ . \ . \ .$$

When $\frac{v^2}{c^2}$ is small compared with unity, the third of these terms is always small in comparison with the second, which last is alone considered in classical mechanics. The first term mc^2 does not contain the velocity, and requires no consideration if we

[* $\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}$ — J.M.]

are only dealing with the question as to how the energy of a point-mass depends on the velocity. We shall speak of its essential significance later.

The most important result of a general character to which the special theory of relativity has led is concerned with the conception of mass. Before the advent of relativity, physics recognised two conservation laws of fundamental importance, namely, the law of the conservation of energy and the law of the conservation of mass; these two fundamental laws appeared to be quite independent of each other. By means of the theory of relativity they have been united into one law. We shall now briefly consider how this unification came about, and what meaning is to be attached to it.

The principle of relativity requires that the law of the conservation of energy should hold not only with reference to a co-ordinate system K, but also with respect to every co-ordinate system K' which is in a state of uniform motion of translation relative to K, or, briefly, relative to every "Galileian" system of co-ordinates. In contrast to classical mechanics, the Lorentz transformation is the deciding factor in the transition from one such system to another.

By means of comparatively simple considerations we are led to draw the following conclusion from these premises, in conjunction with the fundamental equations of the electrodynamics of Maxwell: A body moving with the velocity v, which absorbs[1] an amount of energy E_0 in the form of radiation without suffering an alteration in velocity in the process, has, as a consequence, its energy increased by an amount

$$\frac{E_0}{\sqrt{1-\frac{v^2}{c^2}}}.^*$$

In consideration of the expression given above for the kinetic energy of the body, the required energy of the body comes out to be

$$\frac{\left(m+\frac{E_0}{c^2}\right)c^2}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Thus the body has the same energy as a body of mass $\left(m+\frac{E_0}{c^2}\right)$ moving with the velocity v. Hence we can say: If a body takes up an amount of energy E_0, then its inertial mass increases by an amount $\frac{E_0}{c^2}$; the inertial mass of a body is not a constant, but varies according to the change in the energy of the body. The inertial mass of a system of bodies can even be regarded as a measure

[1] E_0 is the energy taken up, as judged from a co-ordinate system moving with the body.

[* $\frac{E_0}{\sqrt{1-\frac{v^2}{c^2}}}$ — J.M.]