is that this provides a framework in which to deal with the true particle states in the interacting theory through renormalization. Indeed, the formula (3.106), suitably interpreted, remains true even in the interacting theory, taking into account the swarm of virtual particles surrounding asymptotic states. This is the correct way to consider scattering. In this context, (3.106) is known as the LSZ reduction formula. You will derive it properly next term.

4. The Dirac Equation

> "A great deal more was hidden in the Dirac equation than the author had expected when he wrote it down in 1928. Dirac himself remarked in one of his talks that his equation was more intelligent than its author. It should be added, however, that it was Dirac who found most of the additional insights."
>
> *Weisskopf on Dirac*

So far we've only discussed scalar fields such that under a Lorentz transformation $x^\mu \to (x')^\mu = \Lambda^\mu{}_\nu x^\nu$, the field transforms as

$$\varphi(x) \to \varphi'(x) = \varphi(\Lambda^{-1}x) \tag{4.1}$$

We have seen that quantization of such fields gives rise to spin 0 particles. But most particles in Nature have an intrinsic angular momentum, or spin. These arise naturally in field theory by considering fields which themselves transform non-trivially under the Lorentz group. In this section we will describe the Dirac equation, whose quantization gives rise to fermionic spin 1/2 particles. To motivate the Dirac equation, we will start by studying the appropriate representation of the Lorentz group.

A familiar example of a field which transforms non-trivially under the Lorentz group is the vector field $A_\mu(x)$ of electromagnetism,

$$A^\mu(x) \to \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x) \tag{4.2}$$

We'll deal with this in Section 6. (It comes with its own problems!). In general, a field can transform as

$$\varphi^a(x) \to D[\Lambda]^a{}_b \varphi^b(\Lambda^{-1}x) \tag{4.3}$$

where the matrices $D[\Lambda]$ form a *representation* of the Lorentz group, meaning that

$$D[\Lambda_1]D[\Lambda_2] = D[\Lambda_1\Lambda_2] \tag{4.4}$$

and $D[\Lambda^{-1}] = D[\Lambda]^{-1}$ and $D[1] = 1$. How do we find the different representations? Typically, we look at infinitesimal transformations of the Lorentz group and study the resulting Lie algebra. If we write,

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu \tag{4.5}$$

for infinitesimal ω, then the condition for a Lorentz transformation $\Lambda^\mu{}_\sigma \Lambda^\nu{}_\rho \eta^{\sigma\rho} = \eta^{\mu\nu}$ becomes the requirement that ω is anti-symmetric:

$$\omega^{\mu\nu} + \omega^{\nu\mu} = 0 \tag{4.6}$$

Note that an antisymmetric 4 × 4 matrix has 4 × 3/2 = 6 independent components, which agrees with the 6 transformations of the Lorentz group: 3 rotations and 3 boosts. It's going to be useful to introduce a basis of these six 4 × 4 anti-symmetric matrices. We could call them $(M^A)^{\mu\nu}$, with $A = 1, \dots, 6$. But in fact it's better for us (although initially a little confusing) to replace the single index A with a pair of antisymmetric indices $[\rho\sigma]$, where $\rho, \sigma = 0, \dots, 3$, so we call our matrices $(M^{\rho\sigma})^{\mu}{}_{\nu}$. The antisymmetry on the ρ and σ indices means that, for example, $M^{01} = -M^{10}$, etc, so that ρ and σ again label six different matrices. Of course, the matrices are also antisymmetric on the $\mu\nu$ indices because they are, after all, antisymmetric matrices. With this notation in place, we can write a basis of six 4 × 4 antisymmetric matrices as

$$(M^{\rho\sigma})^{\mu\nu} = \eta^{\rho\mu}\eta^{\sigma\nu} - \eta^{\sigma\mu}\eta^{\rho\nu} \tag{4.7}$$

where the indices μ and ν are those of the 4 × 4 matrix, while ρ and σ denote which basis element we're dealing with. If we use these matrices for anything practical (for example, if we want to multiply them together, or act on some field) we will typically need to lower one index, so we have

$$(M^{\rho\sigma})^{\mu}{}_{\nu} = \eta^{\rho\mu}\delta^{\sigma}{}_{\nu} - \eta^{\sigma\mu}\delta^{\rho}{}_{\nu} \tag{4.8}$$

Since we lowered the index with the Minkowski metric, we pick up various minus signs which means that when written in this form, the matrices are no longer necessarily antisymmetric. Two examples of these basis matrices are,

$$(M^{01})^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad (M^{12})^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.9}$$

The first, M^{01}, generates boosts in the x^1 direction. It is real and symmetric. The second, M^{12}, generates rotations in the (x^1, x^2)-plane. It is real and antisymmetric. We can now write any $\omega^{\mu}{}_{\nu}$ as a linear combination of the $M^{\rho\sigma}$,

$$\omega^{\mu}{}_{\nu} = \frac{1}{2}\Omega_{\rho\sigma}(M^{\rho\sigma})^{\mu}{}_{\nu} \tag{4.10}$$

where $\Omega_{\rho\sigma}$ are just six numbers (again antisymmetric in the indices) that tell us what Lorentz transformation we're doing. The six basis matrices $M^{\rho\sigma}$ are called the *generators* of the Lorentz transformations. The generators obey the Lorentz Lie algebra relations,

$$[M^{\rho\sigma}, M^{\tau\nu}] = \eta^{\sigma\tau}M^{\rho\nu} - \eta^{\rho\tau}M^{\sigma\nu} + \eta^{\rho\nu}M^{\sigma\tau} - \eta^{\sigma\nu}M^{\rho\tau} \tag{4.11}$$

where we have suppressed the matrix indices. A finite Lorentz transformation can then be expressed as the exponential

$$\Lambda = \exp\left(\tfrac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma}\right) \tag{4.12}$$

Let me stress again what each of these objects are: the $M^{\rho\sigma}$ are six 4×4 basis elements of the Lorentz Lie algebra; the $\Omega_{\rho\sigma}$ are six numbers telling us what kind of Lorentz transformation we're doing (for example, they say things like rotate by $\vartheta = \pi/7$ about the x^3-direction and run at speed $v = 0.2$ in the x^1 direction).

4.1 The Spinor Representation

We're interested in finding other matrices which satisfy the Lorentz algebra commutation relations (4.11). We will construct the spinor representation. To do this, we start by defining something which, at first sight, has nothing to do with the Lorentz group. It is the *Clifford algebra*,

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}1 \tag{4.13}$$

where γ^μ, with $\mu = 0, 1, 2, 3$, are a set of four matrices and the 1 on the right-hand side denotes the unit matrix. This means that we must find four matrices such that

$$\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu \quad \text{when } \mu \neq \nu \tag{4.14}$$

and

$$(\gamma^0)^2 = 1 \quad , \quad (\gamma^i)^2 = -1 \quad i = 1, 2, 3 \tag{4.15}$$

It's not hard to convince yourself that there are no representations of the Clifford algebra using 2×2 or 3×3 matrices. The simplest representation of the Clifford algebra is in terms of 4×4 matrices. There are many such examples of 4×4 matrices which obey (4.13). For example, we may take

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{4.16}$$

where each element is itself a 2×2 matrix, with the σ^i the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.17}$$

which themselves satisfy $\{\sigma^i, \sigma^j\} = 2\delta^{ij}$.

One can construct many other representations of the Clifford algebra by taking $V\gamma^\mu V^{-1}$ for any invertible matrix V. However, up to this equivalence, it turns out that there is a unique irreducible representation of the Clifford algebra. The matrices (4.16) provide one example, known as the *Weyl* or *chiral representation* (for reasons that will soon become clear). We will soon restrict ourselves further, and consider only representations of the Clifford algebra that are related to the chiral representation by a unitary transformation V.

So what does the Clifford algebra have to do with the Lorentz group? Consider the commutator of two γ^μ,

$$S^{\rho\sigma} = \frac{1}{4}[\gamma^\rho, \gamma^\sigma] = \begin{cases} 0 & \rho = \sigma \\ \frac{1}{2}\gamma^\rho\gamma^\sigma & \rho \neq \sigma \end{cases} = \frac{1}{2}\gamma^\rho\gamma^\sigma - \frac{1}{2}\eta^{\rho\sigma} \tag{4.18}$$

Let's see what properties these matrices have:

Claim 4.1: $[S^{\mu\nu}, \gamma^\rho] = \gamma^\mu\eta^{\nu\rho} - \gamma^\nu\eta^{\rho\mu}$

Proof: When $\mu \neq \nu$ we have

$$\begin{aligned} [S^{\mu\nu}, \gamma^\rho] &= \tfrac{1}{2}[\gamma^\mu\gamma^\nu, \gamma^\rho] \\ &= \tfrac{1}{2}\gamma^\mu\gamma^\nu\gamma^\rho - \tfrac{1}{2}\gamma^\rho\gamma^\mu\gamma^\nu \\ &= \tfrac{1}{2}\gamma^\mu\{\gamma^\nu, \gamma^\rho\} - \tfrac{1}{2}\gamma^\mu\gamma^\rho\gamma^\nu - \tfrac{1}{2}\{\gamma^\rho, \gamma^\mu\}\gamma^\nu + \tfrac{1}{2}\gamma^\mu\gamma^\rho\gamma^\nu \\ &= \gamma^\mu\eta^{\nu\rho} - \gamma^\nu\eta^{\rho\mu} \end{aligned}$$

Claim 4.2: The matrices $S^{\mu\nu}$ form a representation of the Lorentz algebra (4.11), meaning

$$[S^{\mu\nu}, S^{\rho\sigma}] = \eta^{\nu\rho}S^{\mu\sigma} - \eta^{\mu\rho}S^{\nu\sigma} + \eta^{\mu\sigma}S^{\nu\rho} - \eta^{\nu\sigma}S^{\mu\rho} \tag{4.19}$$

Proof: Taking $\rho \neq \sigma$, and using Claim 4.1 above, we have

$$\begin{aligned} [S^{\mu\nu}, S^{\rho\sigma}] &= \tfrac{1}{2}[S^{\mu\nu}, \gamma^\rho\gamma^\sigma] \\ &= \tfrac{1}{2}[S^{\mu\nu}, \gamma^\rho]\gamma^\sigma + \tfrac{1}{2}\gamma^\rho[S^{\mu\nu}, \gamma^\sigma] \\ &= \tfrac{1}{2}\gamma^\mu\gamma^\sigma\eta^{\nu\rho} - \tfrac{1}{2}\gamma^\nu\gamma^\sigma\eta^{\rho\mu} + \tfrac{1}{2}\gamma^\rho\gamma^\mu\eta^{\nu\sigma} - \tfrac{1}{2}\gamma^\rho\gamma^\nu\eta^{\sigma\mu} \end{aligned} \tag{4.20}$$

Now using the expression (4.18) to write $\gamma^\mu\gamma^\sigma = 2S^{\mu\sigma} + \eta^{\mu\sigma}$, we have

$$[S^{\mu\nu}, S^{\rho\sigma}] = S^{\mu\sigma}\eta^{\nu\rho} - S^{\nu\sigma}\eta^{\rho\mu} + S^{\rho\mu}\eta^{\nu\sigma} - S^{\rho\nu}\eta^{\sigma\mu} \tag{4.21}$$

which is our desired expression.

4.1.1 Spinors

The $S^{\mu\nu}$ are 4×4 matrices, because the γ^μ are 4×4 matrices. So far we haven't given an index name to the rows and columns of these matrices: we're going to call them $\alpha, \beta = 1, 2, 3, 4$.

We need a field for the matrices $(S^{\mu\nu})^\alpha{}_\beta$ to act upon. We introduce the Dirac *spinor* field $\psi^\alpha(x)$, an object with four complex components labelled by $\alpha = 1, 2, 3, 4$. Under Lorentz transformations, we have

$$\psi^\alpha(x) \to S[\Lambda]^\alpha{}_\beta \psi^\beta(\Lambda^{-1}x) \tag{4.22}$$

where

$$\Lambda = \exp\left(\tfrac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma}\right) \tag{4.23}$$

$$S[\Lambda] = \exp\left(\tfrac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) \tag{4.24}$$

Although the basis of generators $M^{\rho\sigma}$ and $S^{\rho\sigma}$ are different, we use the same six numbers $\Omega_{\rho\sigma}$ in both Λ and $S[\Lambda]$: this ensures that we're doing the same Lorentz transformation on x and ψ. Note that we denote both the generator $S^{\rho\sigma}$ and the full Lorentz transformation $S[\Lambda]$ as "S". To avoid confusion, the latter will always come with the square brackets $[\Lambda]$.

Both Λ and $S[\Lambda]$ are 4×4 matrices. So how can we be sure that the spinor representation is something new, and isn't equivalent to the familiar representation $\Lambda^\mu{}_\nu$? To see that the two representations are truly different, let's look at some specific transformations.

Rotations

$$S^{ij} = \frac{1}{2}\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} = \frac{i}{2}\epsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \qquad (\text{for } i \neq j) \tag{4.25}$$

If we write the rotation parameters as $\Omega_{ij} = -\epsilon_{ijk}\phi^k$ (meaning $\Omega_{12} = -\phi^3$, etc) then the rotation matrix becomes

$$S[\Lambda] = \exp\left(\tfrac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) = \begin{pmatrix} e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} \end{pmatrix} \tag{4.26}$$

where we need to remember that $\Omega_{12} = -\Omega_{21} = -\phi^3$ when following factors of 2.

Consider now a rotation by 2π about, say, the x^3-axis. This is achieved by $\vec{\phi} = (0, 0,$

and the spinor rotation matrix becomes,

$$S[\Lambda] = \begin{pmatrix} e^{+i\pi\sigma^3} & 0 \\ 0 & e^{+i\pi\sigma^3} \end{pmatrix} = -1 \qquad (4.27)$$

Therefore under a 2π rotation

$$\psi^\alpha(x) \to -\psi^\alpha(x) \qquad (4.28)$$

which is definitely not what happens to a vector! To check that we haven't been cheating with factors of 2, let's see how a vector would transform under a rotation by $\vec{\phi} = (0, 0, \phi^3)$. We have

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma}\right) = \exp\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \phi^3 & 0 \\ 0 & -\phi^3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad (4.29)$$

So when we rotate a vector by $\phi^3 = 2\pi$, we learn that $\Lambda = 1$ as you would expect. So $S[\Lambda]$ is definitely a different representation from the familiar vector representation $\Lambda^\mu{}_\nu$.

Boosts

$$S_{[0i]} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \sigma^i & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -\sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} \qquad (4.30)$$

$2 \quad 0$

σ^i

Writing the boost parameter as $\Omega_{i0} = -\Omega_{0i} = \chi_i$, we have

$$S[\Lambda] = \begin{pmatrix} e^{+\vec{\chi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{-\vec{\chi}\cdot\vec{\sigma}/2} \end{pmatrix} \qquad (4.31)$$

Representations of the Lorentz Group are not Unitary

Note that for rotations given in (4.26), $S[\Lambda]$ is unitary, satisfying $S[\Lambda]^\dagger S[\Lambda] = 1$. But for boosts given in (4.31), $S[\Lambda]$ is not unitary. In fact, there are *no* finite dimensional unitary representations of the Lorentz group. We have demonstrated this explicitly for the spinor representation using the chiral representation (4.16) of the Clifford algebra. We can get a feel for why it is true for a spinor representation constructed from any representation of the Clifford algebra. Recall that

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) \qquad (4.32)$$

so the representation is unitary if $S^{\mu\nu}$ are anti-hermitian, i.e. $(S^{\mu\nu})^\dagger = -S^{\mu\nu}$. But we have

$$(S^{\mu\nu})^\dagger = -\frac{1}{4}[(\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger] \qquad (4.33)$$

which can be anti-hermitian if all γ^μ are hermitian or all are anti-hermitian. However, we can never arrange for this to happen since

$$(\gamma^0)^2 = 1 \Rightarrow \text{Real Eigenvalues}$$
$$(\gamma^i)^2 = -1 \Rightarrow \text{Imaginary Eigenvalues} \tag{4.34}$$

So we could pick γ^0 to be hermitian, but we can only pick γ^i to be anti-hermitian. Indeed, in the chiral representation (4.16), the matrices have this property: $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. In general there is no way to pick γ^μ such that $S^{\mu\nu}$ are anti-hermitian.

4.2 Constructing an Action

We now have a new field to work with, the Dirac spinor ψ. We would like to construct a Lorentz invariant equation of motion. We do this by constructing a Lorentz invariant action.

We will start in a naive way which won't work, but will give us a clue how to proceed. Define

$$\psi^\dagger(x) = (\psi^\star)^T(x) \tag{4.35}$$

which is the usual adjoint of a multi-component object. We could then try to form a Lorentz scalar by taking the product $\psi^\dagger\psi$, with the spinor indices summed over. Let's see how this transforms under Lorentz transformations,

$$\psi(x) \rightarrow S[\Lambda]\,\psi(\Lambda^{-1}x)$$
$$\psi^\dagger(x) \rightarrow \psi^\dagger(\Lambda^{-1}x)\,S[\Lambda]^\dagger \tag{4.36}$$

So $\psi^\dagger(x)\psi(x) \rightarrow \psi^\dagger(\Lambda^{-1}x)S[\Lambda]^\dagger S[\Lambda]\psi(\Lambda^{-1}x)$. But, as we have seen, for some Lorentz transformation $S[\Lambda]^\dagger S[\Lambda] \neq 1$ since the representation is not unitary. This means that $\psi^\dagger\psi$ isn't going to do it for us: it doesn't have any nice transformation under the Lorentz group, and certainly isn't a scalar. But now we see why it fails, we can also see how to proceed. Let's pick a representation of the Clifford algebra which, like the chiral representation (4.16), satisfies $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. Then for all $\mu = 0, 1, 2, 3$ we have

$$\gamma^0\gamma^\mu\gamma^0 = (\gamma^\mu)^\dagger \tag{4.37}$$

which, in turn, means that

$$(S^{\mu\nu})^\dagger = \frac{1}{4}[(\gamma^\nu)^\dagger, (\gamma^\mu)^\dagger] = -\gamma^0 S^{\mu\nu}\gamma^0 \tag{4.38}$$

so that

$$S[\Lambda]^\dagger = \exp\left(\tfrac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\dagger\right) = \gamma^0 S[\Lambda]^{-1}\gamma^0 \tag{4.39}$$

With this in mind, we now define the *Dirac adjoint*

$$\bar{\psi}(x) = \psi^\dagger(x)\,\gamma^0 \tag{4.40}$$

Let's now see what Lorentz covariant objects we can form out of a Dirac spinor ψ and its adjoint $\bar{\psi}$.

Claim 4.3: $\bar{\psi}\psi$ is a Lorentz scalar.

Proof: Under a Lorentz transformation,

$$\begin{aligned}\bar{\psi}(x)\,\psi(x) = \psi^\dagger(x)\,\gamma^0\,\psi(x) &\to \psi^\dagger(\Lambda^{-1}x)\,S[\Lambda]^\dagger\,\gamma^0\,S[\Lambda]\psi(\Lambda^{-1}x)\\ &= \psi^\dagger(\Lambda^{-1}x)\,\gamma^0\,\psi(\Lambda^{-1}x)\\ &= \bar{\psi}(\Lambda^{-1}x)\,\psi(\Lambda^{-1}x)\end{aligned} \tag{4.41}$$

which is indeed the transformation law for a Lorentz scalar.

Claim 4.4: $\bar{\psi}\gamma^\mu\psi$ is a Lorentz vector, which means that

$$\bar{\psi}(x)\,\gamma^\mu\,\psi(x) \to \Lambda^\mu{}_\nu\bar{\psi}(\Lambda^{-1}x)\,\gamma^\nu\,\psi(\Lambda^{-1}x) \tag{4.42}$$

This equation means that we can treat the $\mu = 0, 1, 2, 3$ index on the γ^μ matrices as a true vector index. In particular we can form Lorentz scalars by contracting it with other Lorentz indices.

Proof: Suppressing the x argument, under a Lorentz transformation we have,

$$\bar{\psi}\gamma^\mu\psi \to \bar{\psi}\,S[\Lambda]^{-1}\gamma^\mu S[\Lambda]\psi \tag{4.43}$$

If $\bar{\psi}\gamma^\mu\psi$ is to transform as a vector, we must have

$$S[\Lambda]^{-1}\gamma^\mu S[\Lambda] = \Lambda^\mu{}_\nu\gamma^\nu \tag{4.44}$$

We'll now show this. We work infinitesimally, so that

$$\Lambda = \exp\left(\tfrac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right) \approx 1 + \tfrac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma} + \ldots \tag{4.45}$$

$$S[\Lambda] = \exp\left(\tfrac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) \approx 1 + \tfrac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma} + \ldots \tag{4.46}$$

so the requirement (4.44) becomes

$$-[S^{\rho\sigma}, \gamma^\mu] = (\mathcal{M}^{\rho\sigma})^\mu{}_\nu \gamma^\nu \tag{4.47}$$

where we've suppressed the α, β indices on γ^μ and $S^{\mu\nu}$, but otherwise left all other indices explicit. In fact equation (4.47) follows from Claim 4.1 where we showed that $[S^{\rho\sigma}, \gamma^\mu] = \gamma^\rho\eta^{\sigma\mu} - \gamma^\sigma\eta^{\mu\rho}$. To see this, we write the right-hand side of (4.47) by expanding out $\mathcal{M}$,

$$\begin{aligned}(\mathcal{M}^{\rho\sigma})^\mu{}_\nu \gamma^\nu &= (\eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho_\nu)\gamma^\nu \\ &= \eta^{\rho\mu}\gamma^\sigma - \eta^{\sigma\mu}\gamma^\rho \end{aligned}\tag{4.48}$$

which means that the proof follows if we can show

$$-[S^{\rho\sigma}, \gamma^\mu] = \eta^{\rho\mu}\gamma^\sigma - \eta^{\sigma\mu}\gamma^\rho \tag{4.49}$$

which is exactly what we proved in Claim 4.1.

Claim 4.5: $\bar{\psi}\gamma^\mu\gamma^\nu\psi$ transforms as a Lorentz tensor. More precisely, the symmet- ric part is a Lorentz scalar, proportional to $\eta^{\mu\nu}\bar{\psi}\psi$, while the antisymmetric part is a Lorentz tensor, proportional to $\bar{\psi}S^{\mu\nu}\psi$.

Proof: As above.

We are now armed with three bilinears of the Dirac field, $\bar{\psi}\psi$, $\bar{\psi}\gamma^\mu\psi$ and $\bar{\psi}\gamma^\mu\gamma^\nu\psi$, each of which transforms covariantly under the Lorentz group. We can try to build a Lorentz invariant action from these. In fact, we need only the first two. We choose

$$S = \int d^4x\, \bar{\psi}(x)\,(i\gamma^\mu\partial_\mu - m)\,\psi(x) \tag{4.50}$$

This is the Dirac action. The factor of "i" is there to make the action real; upon complex conjugation, it cancels a minus sign that comes from integration by parts. (Said another way, it's there for the same reason that the Hermitian momentum operator $-i\nabla$ in quantum mechanics has a factor i). As we will see in the next section, after quantization this theory describes particles and anti-particles of mass $|m|$ and spin 1/2. Notice that the Lagrangian is first order, rather than the second order Lagrangians we were working with for scalar fields. Also, the mass appears in the Lagrangian as m, which can be positive or negative.

The equation of motion follows from the action (4.50) by varying with respect to ψ and $\bar{\psi}$ independently. Varying with respect to $\bar{\psi}$, we have

$$(i\gamma^\mu\partial_\mu - m)\,\psi = 0 \tag{4.51}$$

This is the *Dirac equation*. It's completely gorgeous. Varying with respect to ψ gives the conjugate equation

$$i\partial_\mu\bar{\psi}\,\gamma^\mu + m\bar{\psi} = 0 \tag{4.52}$$

The Dirac equation is first order in derivatives, yet miraculously Lorentz invariant. If we tried to write down a first order equation of motion for a scalar field, it would look like $v^\mu\partial_\mu\varphi = \ldots$, which necessarily includes a privileged vector in spacetime v^μ and is not Lorentz invariant. However, for spinor fields, the magic of the γ^μ matrices means that the Dirac Lagrangian is Lorentz invariant.

The Dirac equation mixes up different components of ψ through the matrices γ^μ. However, each individual component itself solves the Klein-Gordon equation. To see this, write

$$(i\gamma^\nu\partial_\nu + m)(i\gamma^\mu\partial_\mu - m)\,\psi = -\left(\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + m^2\right)\psi = 0 \tag{4.53}$$

But $\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu = \frac{1}{2}\{\gamma^\mu,\gamma^\nu\}\partial_\mu\partial_\nu = \partial_\mu\partial^\mu$, so we get

$$-(\partial_\mu\partial^\mu + m^2)\psi = 0 \tag{4.54}$$

where this last equation has no γ^μ matrices, and so applies to each component ψ^α, with $\alpha = 1, 2, 3, 4$.

The Slash

Let's introduce some useful notation. We will often come across 4-vectors contracted with γ^μ matrices. We write

$$A_\mu\gamma^\mu \equiv \not{A} \tag{4.55}$$

so the Dirac equation reads

$$(i\not{\partial} - m)\psi = 0 \tag{4.56}$$

4.4 Chiral Spinors

When we've needed an explicit form of the γ^μ matrices, we've used the chiral representation

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{4.57}$$

In this representation, the spinor rotation transformation $S[\Lambda_{\rm rot}]$ and boost transformation $S[\Lambda_{\rm boost}]$ were computed in (4.26) and (4.31). Both are block diagonal,

$$S[\Lambda_{\rm rot}] = \begin{pmatrix} e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} \end{pmatrix} \quad \text{and} \quad S[\Lambda_{\rm boost}] = \begin{pmatrix} e^{+\vec{\chi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{-\vec{\chi}\cdot\vec{\sigma}/2} \end{pmatrix} \tag{4.58}$$

This means that the Dirac spinor representation of the Lorentz group is *reducible*. It decomposes into two irreducible representations, acting only on two-component spinors $u_\pm$ which, in the chiral representation, are defined by

$$\psi = \begin{pmatrix} u_+ \\ u_- \end{pmatrix} \tag{4.59}$$

The two-component objects $u_\pm$ are called *Weyl spinors* or *chiral spinors*. They transform in the same way under rotations,

$$u_\pm \to e^{i\vec{\varphi}\cdot\vec{\sigma}/2}\, u_\pm \tag{4.60}$$

but oppositely under boosts,

$$u_\pm \to e^{\pm\vec{\chi}\cdot\vec{\sigma}/2}\, u_\pm \tag{4.61}$$

In group theory language, u_+ is in the $(\frac{1}{2}, 0)$ representation of the Lorentz group, while u_- is in the $(0, \frac{1}{2})$ representation. The Dirac spinor ψ lies in the $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ representation. (Strictly speaking, the spinor is a representation of the double cover of the Lorentz group $SL(2, \mathbf{C})$).

4.4.1 The Weyl Equation

Let's see what becomes of the Dirac Lagrangian under the decomposition (4.59) into Weyl spinors. We have

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi = iu_-^\dagger \sigma^\mu\partial_\mu u_- + iu_+^\dagger \bar{\sigma}^\mu \partial_\mu u_+ - m(u_+^\dagger u_- + u_-^\dagger u_+) = 0 \tag{4.62}$$

where we have introduced some new notation for the Pauli matrices with a $\mu = 0, 1, 2, 3$ index,

$$\sigma^\mu = (1, \sigma^i) \quad \text{and} \quad \bar{\sigma}^\mu = (1, -\sigma^i) \tag{4.63}$$

From (4.62), we see that a massive fermion requires both u_+ and u_-, since they couple through the mass term. However, a massless fermion can be described by u_+ (or u_-) alone, with the equation of motion

$$i\bar{\sigma}^\mu \partial_\mu u_+ = 0$$

$$\text{or} \quad i\sigma^\mu \partial_\mu u_- = 0 \tag{4.64}$$

These are the *Weyl equations*.

Degrees of Freedom

Let me comment here on the degrees of freedom in a spinor. The Dirac fermion has 4 complex components = 8 real components. How do we count degrees of freedom? In classical mechanics, the number of degrees of freedom of a system is equal to the dimension of the configuration space or, equivalently, half the dimension of the phase space. In field theory we have an infinite number of degrees of freedom, but it makes sense to count the number of degrees of freedom per spatial point: this should at least be finite. For example, in this sense a real scalar field φ has a single degree of freedom. At the quantum level, this translates to the fact that it gives rise to a single type of particle. A classical complex scalar field has two degrees of freedom, corresponding to the particle and the anti-particle in the quantum theory.

But what about a Dirac spinor? One might think that there are 8 degrees of freedom. But this isn't right. Crucially, and in contrast to the scalar field, the equation of motion is first order rather than second order. In particular, for the Dirac Lagrangian, the momentum conjugate to the spinor ψ is given by

$$\pi_\psi = \partial \mathcal{L}/\partial \dot{\psi} = i\psi^\dagger \tag{4.65}$$

It is not proportional to the time derivative of ψ. This means that the phase space for a spinor is therefore parameterized by ψ and $\psi^\dagger$, while for a scalar it is parameterized by φ and $\pi = \dot{\varphi}$. So the *phase space* of the Dirac spinor ψ has 8 real dimensions and correspondingly the number of real degrees of freedom is 4. We will see in the next section that, in the quantum theory, this counting manifests itself as two degrees of freedom (spin up and down) for the particle, and a further two for the anti-particle.

A similar counting for the Weyl fermion tells us that it has two degrees of freedom.

4.4.2 γ^5

The Lorentz group matrices $S[\Lambda]$ came out to be block diagonal in (4.58) because we chose the specific representation (4.57). In fact, this is why the representation (4.57) is called the chiral representation: it's because the decomposition of the Dirac spinor ψ is simply given by (4.59). But what happens if we choose a different representation γ^μ of the Clifford algebra, so that

$$\gamma^\mu \rightarrow U\gamma^\mu U^{-1} \quad \text{and} \quad \psi \rightarrow U\psi \quad ? \tag{4.66}$$

Now $S[\Lambda]$ will not be block diagonal. Is there an invariant way to define chiral spinors? We can do this by introducing the "fifth" gamma-matrix

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{4.67}$$

You can check that this matrix satisfies

$$\{\gamma^5, \gamma^\mu\} = 0 \quad \text{and} \quad (\gamma^5)^2 = +1 \tag{4.68}$$

The reason that this is called γ^5 is because the set of matrices $\tilde{\gamma}^A = (\gamma^\mu, i\gamma^5)$, with $A = 0, 1, 2, 3, 4$ satisfy the $d = 4 + 1$ Clifford algebra $\{\tilde{\gamma}^A, \tilde{\gamma}^B\} = 2\eta^{AB}$. (You might think that γ^4 would be a better name! But γ^5 is the one everyone chooses - it's a more sensible name in Euclidean space, where $A = 1, 2, 3, 4, 5$). You can also check that $[S_{\mu\nu}, \gamma^5] = 0$, which means that γ^5 is a scalar under rotations and boosts. Since $(\gamma^5)^2 = 1$, this means we may form the Lorentz invariant projection operators

$$P_\pm = \frac{1}{2}\left(1 \pm \gamma^5\right) \tag{4.69}$$

such that $P_+^2 = P_+$ and $P_-^2 = P_-$ and $P_+P_- = 0$. One can check that for the chiral representation (4.57),

$$\gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.70}$$

from which we see that the operators $P_\pm$ project onto the Weyl spinors $u_\pm$. However, for an arbitrary representation of the Clifford algebra, we may use γ^5 to define the chiral spinors,

$$\psi_\pm = P_\pm\psi \tag{4.71}$$

which form the irreducible representations of the Lorentz group. ψ_+ is often called a "left-handed" spinor, while ψ_- is "right-handed". The name comes from the way the spin precesses as a massless fermion moves: we'll see this in Section 4.7.2.

4.4.3 Parity

The spinors $\psi_\pm$ are related to each other by *parity*. Let's pause to define this concept. The Lorentz group is defined by $x^\mu \to \Lambda^\mu{}_\nu x^\nu$ such that

$$\Lambda^\mu{}_\nu \Lambda^\rho{}_\sigma \eta^{\nu\sigma} = \eta^{\mu\rho} \tag{4.72}$$

So far we have only considered transformations Λ which are continuously connected to the identity; these are the ones which have an infinitesimal form. However there are also two discrete symmetries which are part of the Lorentz group. They are

$$\begin{aligned} \text{Time Reversal} \quad T &: x^0 \to -x^0 \ ; \ x^i \to x^i \\ \text{Parity} \quad P &: x^0 \to x^0 \ ; \ x^i \to -x^i \end{aligned} \tag{4.73}$$

We won't discuss time reversal too much in this course. (It turns out to be represented by an anti-unitary transformation on states. See, for example the book by Peskin and Schroeder). But parity has an important role to play in the standard model and, in particular, the theory of the weak interaction.

Under parity, the left and right-handed spinors are exchanged. This follows from the transformation of the spinors under the Lorentz group. In the chiral representation, we saw that the rotation (4.60) and boost (4.61) transformations for the Weyl spinors $u_\pm$ are

$$u_\pm \xrightarrow{\text{rot}} e^{i\vec{\varphi}\cdot\vec{\sigma}/2} u_\pm \quad \text{and} \quad u_\pm \xrightarrow{\text{boost}} e^{\pm\vec{\chi}\cdot\vec{\sigma}/2} u_\pm \tag{4.74}$$

Under parity, rotations don't change sign. But boosts do flip sign. This confirms that parity exchanges right-handed and left-handed spinors, $P : u_\pm \to u_\mp$, or in the notation $\psi_\pm = \frac{1}{2}(1 \pm \gamma^5)\psi$, we have

$$P : \psi_\pm(\vec{x}, t) \to \psi_\mp(-\vec{x}, t) \tag{4.75}$$

Using this knowledge of how chiral spinors transform, and the fact that $P^2 = 1$, we see that the action of parity on the Dirac spinor itself can be written as

$$P : \psi(\vec{x}, t) \to \gamma^0\psi(-\vec{x}, t) \tag{4.76}$$

Notice that if $\psi(\vec{x}, t)$ satisfies the Dirac equation, then the parity transformed spinor $\gamma^0\psi(-\vec{x}, t)$ also satisfies the Dirac equation, meaning

$$(i\gamma^0\partial_t + i\gamma^i\partial_i - m)\gamma^0\psi(-\vec{x}, t) = \gamma^0(i\gamma^0\partial_t - i\gamma^i\partial_i - m)\psi(-\vec{x}, t) = 0 \tag{4.77}$$

where the extra minus sign from passing γ^0 through γ^i is compensated by the derivative acting on $-\vec{x}$ instead of $+\vec{x}$.

4.4.4 Chiral Interactions

Let's now look at how our interaction terms change under parity. We can look at each of our spinor bilinears from which we built the action,

$$P: \bar{\psi}\psi(\vec{x},t) \to \bar{\psi}\psi(-\vec{x},t) \tag{4.78}$$

which is the transformation of a scalar. For the vector $\bar{\psi}\gamma^\mu\psi$, we can look at the temporal and spatial components separately,

$$\begin{aligned} P: \bar{\psi}\gamma^0\psi(\vec{x},t) &\to \bar{\psi}\gamma^0\psi(-\vec{x},t) \\ P: \bar{\psi}\gamma^i\psi(\vec{x},t) &\to \bar{\psi}\gamma^0\gamma^i\gamma^0\psi(-\vec{x},t) = -\bar{\psi}\gamma^i\psi(-\vec{x},t) \end{aligned} \tag{4.79}$$

which tells us that $\bar{\psi}\gamma^\mu\psi$ transforms as a vector, with the spatial part changing sign. You can also check that $\bar{\psi}S^{\mu\nu}\psi$ transforms as a suitable tensor.

However, now we've discovered the existence of γ^5, we can form another Lorentz scalar and another Lorentz vector,

$$\bar{\psi}\gamma^5\psi \quad \text{and} \quad \bar{\psi}\gamma^5\gamma^\mu\psi \tag{4.80}$$

How do these transform under parity? We can check:

$$P: \bar{\psi}\gamma^5\psi(\vec{x},t) \to \bar{\psi}\gamma^0\gamma^5\gamma^0\psi(-\vec{x},t) = -\bar{\psi}\gamma^5\psi(-\vec{x},t) \tag{4.81}$$

$$P: \bar{\psi}\gamma^5\gamma^\mu\psi(\vec{x},t) \to \bar{\psi}\gamma^0\gamma^5\gamma^\mu\gamma^0\psi(-\vec{x},t) = \begin{cases} -\bar{\psi}\gamma^5\gamma^0\psi(-\vec{x},t) & \mu = 0 \\ +\bar{\psi}\gamma^5\gamma^i\psi(-\vec{x},t) & \mu = i \end{cases}$$

which means that $\bar{\psi}\gamma^5\psi$ transforms as a *pseudoscalar*, while $\bar{\psi}\gamma^5\gamma^\mu\psi$ transforms as an *axial vector*. To summarize, we have the following spinor bilinears,

$$\begin{aligned} \bar{\psi}\psi &: \quad \text{scalar} \\ \bar{\psi}\gamma^\mu\psi &: \quad \text{vector} \\ \bar{\psi}S^{\mu\nu}\psi &: \quad \text{tensor} \\ \bar{\psi}\gamma^5\psi &: \quad \text{pseudoscalar} \\ \bar{\psi}\gamma^5\gamma^\mu\psi &: \quad \text{axial vector} \end{aligned} \tag{4.82}$$

The total number of bilinears is $1 + 4 + (4 \times 3/2) + 4 + 1 = 16$ which is all we could hope for from a 4-component object.

We're now armed with new terms involving γ^5 that we can start to add to our Lagrangian to construct new theories. Typically such terms will break parity invariance of the theory, although this is not always true. (For example, the term $\varphi\bar{\psi}\gamma^5\psi$ doesn't break parity if φ is itself a pseudoscalar). Nature makes use of these parity violating interactions by using γ^5 in the weak force. A theory which treats $\psi_\pm$ on an equal footing is called a *vector-like theory*. A theory in which ψ_+ and ψ_- appear differently is called a *chiral theory*.

4.5 Majorana Fermions

Our spinor ψ^α is a complex object. It has to be because the representation $S[\Lambda]$ is typically also complex. This means that if we were to try to make ψ real, for example by imposing $\psi = \psi^\star$, then it wouldn't stay that way once we make a Lorentz transformation. However, there is a way to impose a reality condition on the Dirac spinor ψ. To motivate this possibility, it's simplest to look at a novel basis for the Clifford algebra, known as the *Majorana basis*.

$$\gamma^0 = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} i\sigma^3 & 0 \\ 0 & i\sigma^3 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} -i\sigma^1 & 0 \\ 0 & -i\sigma^1 \end{pmatrix}$$

These matrices satisfy the Clifford algebra. What is special about them is that they are all pure imaginary $(\gamma^\mu)^\star = -\gamma^\mu$. This means that the generators of the Lorentz group $S^{\mu\nu} = \frac{1}{4}[\gamma^\mu, \gamma^\nu]$, and hence the matrices $S[\Lambda]$ are real. So with this basis of the Clifford algebra, we can work with a real spinor simply by imposing the condition,

$$\psi = \psi^\star \tag{4.83}$$

which is preserved under Lorentz transformation. Such spinors are called *Majorana spinors*.

So what's the story if we use a general basis for the Clifford algebra? We'll ask only that the basis satisfies $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. We then define the *charge conjugate* of a Dirac spinor ψ as

$$\psi^{(c)} = C\psi^\star \tag{4.84}$$

Here C is a 4×4 matrix satisfying

$$C^\dagger C = 1 \quad \text{and} \quad C^\dagger\gamma^\mu C = -(\gamma^\mu)^\star \tag{4.85}$$

Let's firstly check that (4.84) is a good definition, meaning that $\psi^{(c)}$ transforms nicely under a Lorentz transformation. We have

$$\psi^{(c)} \to CS[\Lambda]^\star\psi^\star = S[\Lambda]C\psi^\star = S[\Lambda]\psi^{(c)} \tag{4.86}$$

where we've made use of the properties (4.85) in taking the matrix C through $S[\Lambda]^\star$. In fact, not only does $\psi^{(c)}$ transform nicely under the Lorentz group, but if ψ satisfies the Dirac equation, then $\psi^{(c)}$ does too. This follows from,

$$(i\partial\!\!\!/ - m)\psi = 0 \;\Rightarrow\; (-i\partial\!\!\!/^\star - m)\psi^\star = 0$$
$$\Rightarrow\; C(-i\partial\!\!\!/^\star - m)\psi^\star = (+i\partial\!\!\!/ - m)\psi^{(c)} = 0$$

Finally, we can now impose the Lorentz invariant reality condition on the Dirac spinor, to yield a Majorana spinor,

$$\psi^{(c)} = \psi \tag{4.87}$$

After quantization, the Majorana spinor gives rise to a fermion that is its own anti-particle. This is exactly the same as in the case of scalar fields, where we've seen that a real scalar field gives rise to a spin 0 boson that is its own anti-particle. (Be aware: In many texts an extra factor of γ^0 is absorbed into the definition of C).

So what is this matrix C? Well, for a given representation of the Clifford algebra, it is something that we can find fairly easily. In the Majorana basis, where the gamma matrices are pure imaginary, we have simply $C_{\text{Maj}} = 1$ and the Majorana condition $\psi = \psi^{(c)}$ becomes $\psi = \psi^\star$. In the chiral basis (4.16), only γ^2 is imaginary, and we may take $C_{\text{chiral}} = i\gamma^2 = \begin{pmatrix} 0 & i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix}$. (The matrix $i\sigma^2$ that appears here is simply the anti-symmetric matrix $\epsilon^{\alpha\beta}$). It is interesting to see how the Majorana condition (4.87) looks in terms of the decomposition into left and right handed Weyl spinors (4.59). Plugging in the various definitions, we find that $u_+ = i\sigma^2 u_-^\star$ and $u_- = -i\sigma^2 u_+^\star$. In other words, a Majorana spinor can be written in terms of Weyl spinors as

$$\psi = \begin{pmatrix} u_+ \\ -i\sigma^2 u_+^\star \end{pmatrix} \tag{4.88}$$

Notice that it's not possible to impose the Majorana condition $\psi = \psi^{(c)}$ at the same time as the Weyl condition ($u_- = 0$ or $u_+ = 0$). Instead the Majorana condition relates u_- and u_+.

An Aside: Spinors in Different Dimensions: The ability to impose Majorana or Weyl conditions on Dirac spinors depends on both the dimension and the signature of spacetime. One can always impose the Weyl condition on a spinor in even dimensional Minkowski space, basically because you can always build a suitable "γ^5" projection matrix by multiplying together all the other γ-matrices. The pattern for when the Majorana condition can be imposed is a little more sporadic. Interestingly, although the Majorana condition and Weyl condition cannot be imposed simultaneously in four dimensions, you can do this in Minowski spacetimes of dimension $2, 10, 18, \ldots$.

The Dirac Lagrangian enjoys a number of symmetries. Here we list them and compute the associated conserved currents.

Spacetime Translations

Under spacetime translations the spinor transforms as

$$\delta\psi = \epsilon^\mu \partial_\mu \psi \tag{4.89}$$

The Lagrangian depends on $\partial_\mu\psi$, but not $\partial_\mu\bar{\psi}$, so the standard formula (1.41) gives us the energy-momentum tensor

$$T^{\mu\nu} = i\bar{\psi}\gamma^\mu\partial^\nu\psi - \eta^{\mu\nu}\mathcal{L} \tag{4.90}$$

Since a current is conserved only when the equations of motion are obeyed, we don't lose anything by imposing the equations of motion already on $T^{\mu\nu}$. In the case of a scalar field this didn't really buy us anything because the equations of motion are second order in derivatives, while the energy-momentum is typically first order. However, for a spinor field the equations of motion are first order: $(i\partial\!\!\!/ - m)\psi = 0$. This means we can set $\mathcal{L} = 0$ in $T^{\mu\nu}$, leaving

$$T^{\mu\nu} = i\bar{\psi}\gamma^\mu\partial^\nu\psi \tag{4.91}$$

In particular, we have the total energy

$$E = \int d^3x\, T^{00} = \int d^3x\, i\bar{\psi}\gamma^0\dot{\psi} = \int d^3x\,\psi^\dagger\gamma^0(-i\gamma^i\partial_i + m)\psi \tag{4.92}$$

where, in the last equality, we have again used the equations of motion.

Lorentz Transformations

Under an infinitesimal Lorentz transformation, the Dirac spinor transforms as (4.22) which, in infinitesimal form, reads

$$\delta\psi^\alpha = -\omega^\mu_{\ \nu}\, x^\nu\partial_\mu\psi^\alpha + \tfrac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\alpha_{\ \beta}\psi^\beta \tag{4.93}$$

where, following (4.10), we have $\omega^\mu_{\ \nu} = \frac{1}{2}\Omega_{\rho\sigma}(\mathcal{M}^{\rho\sigma})^\mu_{\ \nu}$, and $\mathcal{M}^{\rho\sigma}$ are the generators of the Lorentz algebra given by (4.8)

$$(\mathcal{M}^{\rho\sigma})^\mu_{\ \nu} = \eta^{\rho\mu}\delta^\sigma_{\ \nu} - \eta^{\sigma\mu}\delta^\rho_{\ \nu} \tag{4.94}$$

which, after direct substitution, tells us that $\omega^{\mu\nu} = \Omega^{\mu\nu}$. So we get

$$\delta\psi^\alpha = -\omega^{\mu\nu}\, x_\nu \partial_\mu \psi^\alpha - \tfrac{1}{2}(S_{\mu\nu})^\alpha{}_\beta \psi^\beta \qquad (4.95)$$

The conserved current arising from Lorentz transformations now follows from the same calculation we saw for the scalar field (1.54) with two differences: firstly, as we saw above, the spinor equations of motion set $\mathcal{L} = 0$; secondly, we pick up an extra piece in the current from the second term in (4.95). We have

$$(\mathcal{J}^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} + i\bar{\psi}\gamma^\mu S^{\rho\sigma}\psi \qquad (4.96)$$

After quantization, when $(\mathcal{J}^\mu)^{\rho\sigma}$ is turned into an operator, this extra term will be responsible for providing the single particle states with internal angular momentum, telling us that the quantization of a Dirac spinor gives rise to a particle carrying spin 1/2.

Internal Vector Symmetry

The Dirac Lagrangian is invariant under rotating the phase of the spinor, $\psi \to e^{-i\alpha}\psi$. This gives rise to the current

$$j^\mu_V = \bar{\psi}\gamma^\mu\psi \qquad (4.97)$$

where "V" stands for *vector*, reflecting the fact that the left and right-handed components $\psi_\pm$ transform in the same way under this symmetry. We can easily check that j^μ_V is conserved under the equations of motion,

$$\partial_\mu j^\mu_V = (\partial_\mu\bar{\psi})\gamma^\mu\psi + \bar{\psi}\gamma^\mu(\partial_\mu\psi) = im\bar{\psi}\psi - im\bar{\psi}\psi = 0 \qquad (4.98)$$

where, in the last equality, we have used the equations of motion $i\partial\!\!\!/\,\psi = m\psi$ and $i\partial_\mu\bar{\psi}\gamma^\mu = -m\bar{\psi}$. The conserved quantity arising from this symmetry is

$$Q = \int d^3x\, \bar{\psi}\gamma^0\psi = \int d^3x\, \psi^\dagger\psi \qquad (4.99)$$

We will see shortly that this has the interpretation of electric charge, or particle number, for fermions.

Axial Symmetry

When $m = 0$, the Dirac Lagrangian admits an extra internal symmetry which rotates left and right-handed fermions in opposite directions,

$$\psi \to e^{i\alpha\gamma^5}\psi \quad \text{and} \quad \bar{\psi} \to \bar{\psi}e^{i\alpha\gamma^5} \qquad (4.100)$$

Here the second transformation follows from the first after noting that $e^{-i\alpha\gamma^5}\gamma^0 = \gamma^0 e^{+i\alpha\gamma^5}$. This gives the conserved current,

$$j^\mu_A = \bar\psi\gamma^\mu\gamma^5\psi \tag{4.101}$$

where A is for "axial" since j^μ_A is an axial vector. This is conserved only when $m = 0$. Indeed, with the full Dirac Lagrangian we may compute

$$\partial_\mu j^\mu_A = (\partial_\mu\bar\psi)\gamma^\mu\gamma^5\psi + \bar\psi\gamma^\mu\gamma^5\partial_\mu\psi = 2im\bar\psi\gamma^5\psi \tag{4.102}$$

which vanishes only for $m = 0$. However, in the quantum theory things become more interesting for the axial current. When the theory is coupled to gauge fields (in a manner we will discuss in Section 6), the axial transformation remains a symmetry of the classical Lagrangian. But it doesn't survive the quantization process. It is the archetypal example of an *anomaly*: a symmetry of the classical theory that is not preserved in the quantum theory.

4.7 Plane Wave Solutions

Let's now study the solutions to the Dirac equation

$$(i\gamma^\mu\partial_\mu - m)\psi = 0 \tag{4.103}$$

We start by making a simple ansatz:

$$\psi = u(\vec p)\, e^{-ip\cdot x} \tag{4.104}$$

where $u(\vec p)$ is a four-component spinor, independent of spacetime x which, as the notation suggests, can depend on the 3-momentum $\vec p$. The Dirac equation then becomes

$$(\gamma^\mu p_\mu - m)u(\vec p) = \begin{pmatrix} -m & p_\mu\sigma^\mu \\ p_\mu\bar\sigma^\mu & -m \end{pmatrix} u(\vec p) = 0 \tag{4.105}$$

where we're again using the definition,

$$\sigma^\mu = (1, \sigma^i) \quad \text{and} \quad \bar\sigma^\mu = (1, -\sigma^i) \tag{4.106}$$

Claim: The solution to (4.105) is

$$u(\vec p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi \\ \sqrt{p\cdot\bar\sigma}\,\xi \end{pmatrix} \tag{4.107}$$

for any 2-component spinor ξ which we will normalize to $\xi^\dagger \xi = 1$.

Proof: Let's write $u(\vec{p})^T = (u_1, u_2)$. Then equation (4.105) reads

$$(p \cdot \sigma)\, u_2 = m u_1 \quad \text{and} \quad (p \cdot \bar{\sigma})\, u_1 = m u_2 \tag{4.108}$$

Either one of these equations implies the other, a fact which follows from the identity $(p \cdot \sigma)(p \cdot \bar{\sigma}) = p_0^2 - p_i p_j \sigma^i \sigma^j = p_0^2 - p_i p_j \delta^{ij} = p_\mu p^\mu = m^2$. To start with, let's try the ansatz $u_1 = (p \cdot \sigma)\xi'$ for some spinor ξ'. Then the second equation in (4.108) immediately tells us that $u_2 = m\xi'$. So we learn that any spinor of the form

$$u(\vec{p}) = A \begin{pmatrix} (p \cdot \sigma)\, \xi' \\ m\xi' \end{pmatrix} \tag{4.109}$$

with constant A is a solution to (4.105). To make this more symmetric, we choose $A = 1/m$ and $\xi' = \sqrt{p \cdot \bar{\sigma}}\, \xi$ with constant ξ. Then $u_1 = (p \cdot \sigma)\sqrt{p \cdot \bar{\sigma}}\, \xi = m\sqrt{p \cdot \sigma}\, \xi$. So we get the promised result (4.107)

Negative Frequency Solutions

We get further solutions to the Dirac equation from the ansatz

$$\psi = v(\vec{p})\, e^{+ip \cdot x} \tag{4.110}$$

Solutions of the form (4.104), which oscillate in time as $\psi \sim e^{-iEt}$, are called positive frequency solutions. If we compute the energy of these solutions using (4.92), we find that it is positive. Those of the form (4.110), which oscillate as $\psi \sim e^{+iEt}$, are negative frequency solutions. Now if we compute the energy using (4.92), it is negative.

The Dirac equation requires that the 4-component spinor $v(\vec{p})$ satisfies

$$(\gamma^\mu p_\mu + m)\, v(\vec{p}) = \begin{pmatrix} m & p_\mu \sigma^\mu \\ p_\mu \bar{\sigma}^\mu & m \end{pmatrix} v(\vec{p}) = 0 \tag{4.111}$$

which is solved by

$$v(\vec{p}) = \begin{pmatrix} \sqrt{p \cdot \sigma}\, \eta \\ -\sqrt{p \cdot \bar{\sigma}}\, \eta \end{pmatrix} \tag{4.112}$$

for some 2-component spinor η which we take to be constant and normalized to $\eta^\dagger \eta = 1$.

4.7.1 Some Examples

Consider the positive frequency solution with mass m and 3-momentum $\vec{p} = 0$,

$$u(\vec{p}) = \sqrt{m}\begin{pmatrix}\xi\\ \xi\end{pmatrix} \tag{4.113}$$

where ξ is any 2-component spinor. Spatial rotations of the field act on ξ by (4.26),

$$\xi \to e^{+i\vec{\varphi}\cdot\vec{\sigma}/2}\,\xi \tag{4.114}$$

The 2-component spinor ξ defines the *spin* of the field. This should be familiar from quantum mechanics. A field with spin up (down) along a given direction is described by the eigenvector of the corresponding Pauli matrix with eigenvalue +1 (-1 respectively). For example, $\xi^T = (1, 0)$ describes a field with spin up along the z-axis. After quantization, this will become the spin of the associated particle. In the rest of this section, we'll indulge in an abuse of terminology and refer to the classical solutions to the Dirac equations as "particles", even though they have no such interpretation before quantization.

Consider now boosting the particle with spin $\xi^T = (1, 0)$ along the x^3 direction, with $p^\mu = (E, 0, 0, p^3)$. The solution to the Dirac equation becomes

$$u(\vec{p}) = \begin{pmatrix}\sqrt{p\cdot\sigma}\begin{pmatrix}1\\0\end{pmatrix}\\ \sqrt{p\cdot\bar{\sigma}}\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} = \begin{pmatrix}\sqrt{E-p^3}\begin{pmatrix}1\\0\end{pmatrix}\\ \sqrt{E+p^3}\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} \tag{4.115}$$

In fact, this expression also makes sense for a massless field, for which $E = p^3$. (We picked the normalization (4.107) for the solutions so that this would be the case). For a massless particle we have

$$u(\vec{p}) = \sqrt{2E}\begin{pmatrix}0\\0\\1\\0\end{pmatrix} \tag{4.116}$$

Similarly, for a boosted solution of the spin down $\xi^T = (0, 1)$ field, we have

$$u(\vec{p}) = \begin{pmatrix}\sqrt{p\cdot\sigma}\begin{pmatrix}0\\1\end{pmatrix}\\ \sqrt{p\cdot\bar{\sigma}}\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix} = \begin{pmatrix}\sqrt{E+p^3}\begin{pmatrix}0\\1\end{pmatrix}\\ \sqrt{E-p^3}\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix} \xrightarrow{m\to 0} \sqrt{2E}\begin{pmatrix}0\\1\\0\\0\end{pmatrix} \tag{4.117}$$

4.7.2 Helicity

The helicity operator is the projection of the angular momentum along the direction of momentum,

$$h = \frac{i}{2}\epsilon_{ijk}\hat{p}^i S^{jk} = \frac{1}{2}\hat{p}^i \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} \tag{4.118}$$

where S^{ij} is the rotation generator given in (4.25). The massless field with spin $\xi^T = (1, 0)$ in (4.116) has helicity $h = 1/2$: we say that it is *right-handed*. Meanwhile, the field (4.117) has helicity $h = -1/2$: it is *left-handed*.

4.7.3 Some Useful Formulae: Inner and Outer Products

There are a number of identities that will be very useful in the following section, regarding the inner (and outer) products of the spinors $u(\vec{p})$ and $v(\vec{p})$. It's firstly convenient to introduce a basis ξ^s and η^s, $s = 1, 2$ for the two-component spinors such that

$$\xi^{r\dagger}\xi^s = \delta^{rs} \quad \text{and} \quad \eta^{r\dagger}\eta^s = \delta^{rs} \tag{4.119}$$

for example,

$$\xi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \xi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.120}$$

and similarly for η^s. Let's deal first with the positive frequency plane waves. The two independent solutions are now written as

$$u^s(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} \tag{4.121}$$

We can take the inner product of four-component spinors in two different ways: either as $u^\dagger \cdot u$, or as $\bar{u} \cdot u$. Of course, only the latter will be Lorentz invariant, but it turns out that the former is needed when we come to quantize the theory. Here we state both:

$$u^{r\dagger}(\vec{p})\cdot u^s(\vec{p}) = \left(\xi^{r\dagger}\sqrt{p\cdot\sigma}, \xi^{r\dagger}\sqrt{p\cdot\bar{\sigma}}\right)\begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix}$$

$$= \xi^{r\dagger}p\cdot\sigma\xi^s + \xi^{r\dagger}p\cdot\bar{\sigma}\xi^s = 2\xi^{r\dagger}p_0\xi^s = 2p_0\delta^{rs} \tag{4.122}$$

while the Lorentz invariant inner product is

$$\bar{u}^r(\vec{p})\cdot u^s(\vec{p}) = \left(\xi^{r\dagger}\sqrt{p\cdot\sigma}, \xi^{r\dagger}\sqrt{p\cdot\bar{\sigma}}\right)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} = 2m\delta^{rs} \tag{4.123}$$

We have analogous results for the negative frequency solutions, which we may write as

$$v^s(\vec p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\eta^s \\ -\sqrt{p\cdot\bar\sigma}\,\eta^s \end{pmatrix} \quad \text{with} \quad v^{r\dagger}(\vec p)\cdot v^s(\vec p) = 2p_0\delta^{rs} \ \text{and} \ \bar v^r(\vec p)\cdot v^s(\vec p) = -2m\delta^{rs} \tag{4.124}$$

We can also compute the inner product between u and v. We have

$$\bar u^r(\vec p)\cdot v^s(\vec p) = \left(\xi^{r\dagger}\sqrt{p\cdot\sigma},\ \xi^{r\dagger}\sqrt{p\cdot\bar\sigma}\right)\gamma^0 \begin{pmatrix} \sqrt{p\cdot\sigma}\,\eta^s \\ -\sqrt{p\cdot\bar\sigma}\,\eta^s \end{pmatrix}$$

$$= \xi^{r\dagger}\sqrt{(p\cdot\bar\sigma)(p\cdot\sigma)}\,\eta^s - \xi^{r\dagger}\sqrt{(p\cdot\bar\sigma)(p\cdot\sigma)}\,\eta^s = 0 \tag{4.125}$$

and similarly, $\bar v^r(\vec p)\cdot u^s(\vec p) = 0$. However, when we come to $u^\dagger\cdot v$, it is a slightly different combination that has nice properties (and this same combination appears when we quantize the theory). We look at $u^{r\dagger}(\vec p)\cdot v^s(-\vec p)$, with the 3-momentum in the spinor v taking the opposite sign. Defining the 4-momentum $(p^r)^\mu = (p^0, -\vec p)$, we have

$$u^{r\dagger}(\vec p)\cdot v^s(-\vec p) = \left(\xi^{r\dagger}\sqrt{p\cdot\sigma},\ \xi^{r\dagger}\sqrt{p\cdot\bar\sigma}\right) \begin{pmatrix} \sqrt{p^r\cdot\sigma}\,\eta^s \\ -\sqrt{p^r\cdot\bar\sigma}\,\eta^s \end{pmatrix}$$

$$= \xi^{r\dagger}\sqrt{(p\cdot\sigma)(p^r\cdot\sigma)}\,\eta^s - \xi^{r\dagger}\sqrt{(p\cdot\bar\sigma)(p^r\cdot\bar\sigma)}\,\eta^s \tag{4.126}$$

Now the terms under the square-root are given by $(p\cdot\sigma)(p^r\cdot\sigma) = (p_0 + p_i\sigma^i)(p_0 - p_i\sigma^i) = p_0^2 - \vec p^{\,2} = m^2$. The same expression holds for $(p\cdot\bar\sigma)(p^r\cdot\bar\sigma)$, and the two terms cancel. We learn

$$u^{r\dagger}(\vec p)\cdot v^s(-\vec p) = v^{r\dagger}(\vec p)\cdot u^s(-\vec p) = 0 \tag{4.127}$$

Outer Products

There's one last spinor identity that we need before we turn to the quantum theory. It is:

Claim:

$$\sum_{s=1}^{2} u^s(\vec p)\,\bar u^s(\vec p) = p\!\!\!/ + m \tag{4.128}$$

where the two spinors are not now contracted, but instead placed back to back to give a 4×4 matrix. Also,

$$\sum_{s=1}^{2} v^s(\vec p)\,\bar v^s(\vec p) = p\!\!\!/ - m \tag{4.129}$$

Proof:

$$\sum_{s=1}^{2} u^s(\vec{p})\,\bar{u}^s(\vec{p}) = \sum_{s=1}^{2} \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} \quad \sqrt{p\cdot\bar{\sigma}}\,\xi^{s\dagger}, \quad \sqrt{p\cdot\sigma}\,\xi^{s\dagger} \tag{4.130}$$

But $\sum_s \xi^s \xi^{s\dagger} = \mathbf{1}$, the 2 × 2 unit matrix, which then gives us

$$\sum_{s=1} u^s(\vec{p})\,\bar{u}^s(\vec{p}) \quad m \quad p\cdot\sigma \quad p\cdot\bar{\sigma} \quad m \tag{4.131}$$

which is the desired result. A similar proof works for $\sum_s v^s(\vec{p})\,\bar{v}^s(\vec{p})$.

5. Quantizing the Dirac Field

We would now like to quantize the Dirac Lagrangian,

$$\mathcal{L} = \bar{\psi}(x)\left(i\partial\!\!\!/ - m\right)\psi(x) \tag{5.1}$$

We will proceed naively and treat ψ as we did the scalar field. But we'll see that things go wrong and we will have to reconsider how to quantize this theory.

5.1 A Glimpse at the Spin-Statistics Theorem

We start in the usual way and define the momentum,

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\bar{\psi}\gamma^0 = i\psi^\dagger \tag{5.2}$$

For the Dirac Lagrangian, the momentum conjugate to ψ is $i\psi^\dagger$. It does not involve the time derivative of ψ. This is as it should be for an equation of motion that is first order in time, rather than second order. This is because we need only specify ψ and $\psi^\dagger$ on an initial time slice to determine the full evolution.

To quantize the theory, we promote the field ψ and its momentum $\psi^\dagger$ to operators, satisfying the canonical commutation relations, which read

$$[\psi_\alpha(\vec{x}), \psi_\beta(\vec{y})] = [\psi^\dagger_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{y})] = 0$$
$$[\psi_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{y})] = \delta_{\alpha\beta}\,\delta^{(3)}(\vec{x} - \vec{y}) \tag{5.3}$$

It's this step that we'll soon have to reconsider.

Since we're dealing with a free theory, where any classical solution is a sum of plane waves, we may write the quantum operators as

$$\psi(\vec{x}) = \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left[b^s_{\vec{p}}\,u^s(\vec{p})\,e^{+i\vec{p}\cdot\vec{x}} + c^{s\,\dagger}_{\vec{p}}\,v^s(\vec{p})\,e^{-i\vec{p}\cdot\vec{x}}\right]$$
$$\psi^\dagger(\vec{x}) = \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left[b^{s\,\dagger}_{\vec{p}}\,u^s(\vec{p})^\dagger\,e^{-i\vec{p}\cdot\vec{x}} + c^s_{\vec{p}}\,v^s(\vec{p})^\dagger\,e^{+i\vec{p}\cdot\vec{x}}\right] \tag{5.4}$$

where the operators $b^{s\,\dagger}_{\vec{p}}$ create particles associated to the spinors $u^s(\vec{p})$, while $c^{s\,\dagger}_{\vec{p}}$ create particles associated to $v^s(\vec{p})$. As with the scalars, the commutation relations of the fields imply commutation relations for the annihilation and creation operators.

Claim: The field commutation relations (5.3) are equivalent to

$$[b^r_{\vec{p}}, b^{s\,\dagger}_{\vec{q}}] = (2\pi)^3 \delta^{rs}\, \delta^{(3)}(\vec{p} - \vec{q})$$
$$[c^r_{\vec{p}}, c^{s\,\dagger}_{\vec{q}}] = -(2\pi)^3 \delta^{rs}\, \delta^{(3)}(\vec{p} - \vec{q}) \tag{5.5}$$

with all other commutators vanishing.

Note the strange minus sign in the $[c, c^\dagger]$ term. This means that we can't define the ground state $|0\rangle$ as something annihilated by $c^r_{\vec{p}}|0\rangle = 0$, because then the excited states $c^{s\,\dagger}_{\vec{p}}|0\rangle$ would have negative norm. To avoid this, we will have to flip the interpretation of c and $c^\dagger$, with the vacuum defined by $c^{s\,\dagger}_{\vec{p}}|0\rangle = 0$ and the excited states by $c^r_{\vec{p}}|0\rangle$. This, as we will see, will be our undoing.

Proof: Let's show that the $[b, b^\dagger]$ and $[c, c^\dagger]$ commutators reproduce the field commutators (5.3),

$$[\psi(\vec{x}), \psi^\dagger(\vec{y})] = \sum_{r,s} \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{\sqrt{4E_{\vec{p}}E_{\vec{q}}}} \left([b^r_{\vec{p}}, b^{s\,\dagger}_{\vec{q}}]\, u^r(\vec{p}) u^s(\vec{q})^\dagger e^{i(\vec{x}\cdot\vec{p} - \vec{y}\cdot\vec{q})} + [c^{r\,\dagger}_{\vec{p}}, c^s_{\vec{q}}]\, v^r(\vec{p}) v^s(\vec{q})^\dagger e^{-i(\vec{x}\cdot\vec{p} - \vec{y}\cdot\vec{q})} \right)$$

$$= \sum_s \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \left(u^s(\vec{p}) \bar{u}^s(\vec{p}) \gamma^0 e^{i\vec{p}\cdot(\vec{x}-\vec{y})} + v^s(\vec{p}) \bar{v}^s(\vec{p}) \gamma^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \right) \tag{5.6}$$

At this stage we use the outer product formulae (4.128) and (4.129) which tell us $\sum_s u^s(\vec{p})\bar{u}^s(\vec{p}) = \not{p} + m$ and $\sum_s v^s(\vec{p})\bar{v}^s(\vec{p}) = \not{p} - m$, so that

$$[\psi(\vec{x}), \psi^\dagger(\vec{y})] = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \left((\not{p} + m)\gamma^0 e^{i\vec{p}\cdot(\vec{x}-\vec{y})} + (\not{p} - m)\gamma^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \right)$$

$$= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \left((p_0\gamma^0 + p_i\gamma^i + m)\gamma^0 + (p_0\gamma^0 - p_i\gamma^i - m)\gamma^0 \right) e^{+i\vec{p}\cdot(\vec{x}-\vec{y})}$$

where, in the second term, we've changed $\vec{p} \to -\vec{p}$ under the integration sign. Now, using $p_0 = E_{\vec{p}}$ we have

$$[\psi(\vec{x}), \psi^\dagger(\vec{y})] = \int \frac{d^3p}{(2\pi)^3} e^{+i\vec{p}\cdot(\vec{x}-\vec{y})} = \delta^{(3)}(\vec{x} - \vec{y}) \tag{5.7}$$

as promised. Notice that it's a little tricky in the middle there, making sure that the $p_i\gamma^i$ terms cancel. This was the reason we needed the minus sign in the $[c, c^\dagger]$ commutator terms in (5.5).

5.1.1 The Hamiltonian

To proceed, let's construct the Hamiltonian for the theory. Using the momentum $\pi = i\psi^\dagger$, we have

$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L} = \bar{\psi}\,(-i\gamma^i\partial_i + m)\psi \tag{5.8}$$

which means that $H = \int d^3x\, \mathcal{H}$ agrees with the conserved energy computed using Noether's theorem (4.92). We now wish to turn the Hamiltonian into an operator. Let's firstly look at

$$(-i\gamma^i\partial_i + m)\psi = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left[b^s_{\vec{p}}(-\gamma^i p_i + m)u^s(\vec{p})\,e^{+i\vec{p}\cdot\vec{x}} + c^{s\dagger}_{\vec{p}}(\gamma^i p_i + m)v^s(\vec{p})\,e^{-i\vec{p}\cdot\vec{x}}\right]$$

where, for once we've left the sum over $s = 1, 2$ implicit. There's a small subtlety with the minus signs in deriving this equation that arises from the use of the Minkowski metric in contracting indices, so that $\vec{p}\cdot\vec{x} \equiv \sum_i x^i p^i = -x^i p_i$. Now we use the defining equations for the spinors $u^s(\vec{p})$ and $v^s(\vec{p})$ given in (4.105) and (4.111), to replace

$$(-\gamma^i p_i + m)u^s(\vec{p}) = \gamma^0 p_0\, u^s(\vec{p}) \quad\text{and}\quad (\gamma^i p_i + m)v^s(\vec{p}) = -\gamma^0 p_0\, v^s(\vec{p}) \tag{5.9}$$

so we can write

$$(-i\gamma^i\partial_i + m)\psi = \int \frac{d^3p}{(2\pi)^3}\sqrt{\frac{E_{\vec{p}}}{2}}\,\gamma^0\left[b^s_{\vec{p}}\,u^s(\vec{p})\,e^{+i\vec{p}\cdot\vec{x}} - c^{s\dagger}_{\vec{p}}\,v^s(\vec{p})\,e^{-i\vec{p}\cdot\vec{x}}\right] \tag{5.10}$$

We now use this to write the operator Hamiltonian

$$\begin{aligned}
H &= \int d^3x\,\psi^\dagger\gamma^0(-i\gamma^i\partial_i + m)\psi \\
&= \int \frac{d^3x\,d^3p\,d^3q}{(2\pi)^6}\sqrt{\frac{E_{\vec{p}}}{4E_{\vec{q}}}}\left[b^{r\dagger}_{\vec{q}}\,u^r(\vec{q})^\dagger e^{-i\vec{q}\cdot\vec{x}} + c^r_{\vec{q}}\,v^r(\vec{q})^\dagger e^{+i\vec{q}\cdot\vec{x}}\right]\cdot \\
&\qquad\qquad \left[b^s_{\vec{p}}\,u^s(\vec{p})\,e^{+i\vec{p}\cdot\vec{x}} - c^{s\dagger}_{\vec{p}}\,v^s(\vec{p})\,e^{-i\vec{p}\cdot\vec{x}}\right] \\
&= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2}\left[b^{r\dagger}_{\vec{p}}b^s_{\vec{p}}\,[u^r(\vec{p})^\dagger\cdot u^s(\vec{p})] - c^r_{\vec{p}}c^{s\dagger}_{\vec{p}}\,[v^r(\vec{p})^\dagger\cdot v^s(\vec{p})]\right. \\
&\qquad\qquad \left. - b^{r\dagger}_{\vec{p}}c^{s\dagger}_{-\vec{p}}\,[u^r(\vec{p})^\dagger\cdot v^s(-\vec{p})] + c^r_{\vec{p}}b^s_{-\vec{p}}\,[v^r(\vec{p})^\dagger\cdot u^s(-\vec{p})]\right]
\end{aligned}$$

where, in the last two terms we have relabelled $\vec{p} \to -\vec{p}$. We now use our inner product formulae (4.122), (4.124) and (4.127) which read

$$u^r(\vec{p})^\dagger\cdot u^s(\vec{p}) = v^r(\vec{p})^\dagger\cdot v^s(\vec{p}) = 2p_0\delta^{rs} \qquad\text{and}\qquad u^r(\vec{p})^\dagger\cdot v^s(-\vec{p}) = v^r(\vec{p})^\dagger\cdot u^s(-\vec{p}) = 0$$

giving us

$$H = \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} \left(b^{s\,\dagger}_{\vec{p}} b^s_{\vec{p}} - c^s_{\vec{p}} c^{s\,\dagger}_{\vec{p}} \right) \tag{5.11}$$

$$= \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} \left(b^{s\,\dagger}_{\vec{p}} b^s_{\vec{p}} - c^{s\,\dagger}_{\vec{p}} c^s_{\vec{p}} + (2\pi)^3 \delta^{(3)}(0) \right) \tag{5.12}$$

The $\delta^{(3)}$ term is familiar and easily dealt with by normal ordering. The $b^\dagger b$ term is familiar and we can check that $b^\dagger$ create positive energy states as expected,

$[H, b$ ... $\vec{p}$ ... $\vec{p}$

The minus sign in front of the $c^\dagger c$ term should make us nervous. If we think of $c^\dagger$ as creation operators then there's no problem since, using the commutation relation (5.5), we still find that $c^\dagger$ creates positive energy states,

$\vec{p}$ $\vec{p}$ [

However, as we noted after (5.5), these states have negative norm. To have a sensible Hilbert space, we need to interpret c as the creation operator. But then the Hamiltonian is not bounded below because

$$[H, c^s_{\vec{p}}] = -E_{\vec{p}}\, c^s_{\vec{p}}$$

This is a disaster. Taken seriously it would tell us that we could tumble to states of lower and lower energy by continually producing c particles. As the English would say, it's all gone a bit Pete Tong. (No relation).

Since the above calculation was a little tricky, you might think that it's possible to rescue the theory to get the minus signs to work out right. You can play around with different things, but you'll always find this minus sign cropping up somewhere. And, in fact, it's telling us something important that we missed.

5.2 Fermionic Quantization

The key piece of physics that we missed is that spin 1/2 particles are fermions, meaning that they obey Fermi-Dirac statistics with the quantum state picking up a minus sign upon the interchange of any two particles. This fact is embedded into the structure of relativistic quantum field theory: the spin-statistics theorem says that integer spin fields must be quantized as bosons, while half-integer spin fields must be quantized as fermions. Any attempt to do otherwise will lead to an inconsistency, such as the unbounded Hamiltonian we saw in (5.12).

So how do we go about quantizing a field as a fermion? Recall that when we quantized the scalar field, the resulting particles obeyed bosonic statistics because the creation and annihilation operators satisfied the commutation relations,

$$[a^\dagger_{\vec{p}}, a^\dagger_{\vec{q}}] = 0 \quad \Rightarrow \quad a^\dagger_{\vec{p}} a^\dagger_{\vec{q}} |0\rangle \equiv |\vec{p}, \vec{q}\rangle = |\vec{q}, \vec{p}\rangle \tag{5.13}$$

To have states obeying fermionic statistics, we need anti-commutation relations, $\{A, B\} \equiv AB + BA$. Rather than (5.3), we will ask that the spinor fields satisfy

$$\begin{aligned} \{\psi_\alpha(\vec{x}), \psi_\beta(\vec{y})\} &= \{\psi^\dagger_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{y})\} = 0 \\ \{\psi_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{y})\} &= \delta_{\alpha\beta}\, \delta^{(3)}(\vec{x} - \vec{y}) \end{aligned} \tag{5.14}$$

We still have the expansion (5.4) of ψ and $\psi^\dagger$ in terms of b, $b^\dagger$, c and $c^\dagger$. But now the same proof that led us to (5.5) tells us that

$$\begin{aligned} \{b^r_{\vec{p}}, b^{s\,\dagger}_{\vec{q}}\} &= (2\pi)^3 \delta^{rs} \delta^{(3)}(\vec{p} - \vec{q}) \\ \{c^r_{\vec{p}}, c^{s\,\dagger}_{\vec{q}}\} &= (2\pi)^3 \delta^{rs} \delta^{(3)}(\vec{p} - \vec{q}) \end{aligned} \tag{5.15}$$

with all other *anti-commutators* vanishing,

$$\{b^r_{\vec{p}}, b^s_{\vec{q}}\} = \{c^r_{\vec{p}}, c^s_{\vec{q}}\} = \{b^r_{\vec{p}}, c^{s\,\dagger}_{\vec{q}}\} = \{b^r_{\vec{p}}, c^s_{\vec{q}}\} = \ldots = 0 \tag{5.16}$$

The calculation of the Hamiltonian proceeds as before, all the way through to the penultimate line (5.11). At that stage, we get

$$\begin{aligned} H &= \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} \left[b^{s\,\dagger}_{\vec{p}} b^s_{\vec{p}} - c^s_{\vec{p}} c^{s\,\dagger}_{\vec{p}} \right] \\ &= \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} \left[b^{s\,\dagger}_{\vec{p}} b^s_{\vec{p}} + c^{s\,\dagger}_{\vec{p}} c^s_{\vec{p}} - (2\pi)^3 \delta^{(3)}(0) \right] \end{aligned} \tag{5.17}$$

The anti-commutators have saved us from the indignity of a Hamiltonian unbounded below. Note that when normal ordering the Hamiltonian we now throw away a negative contribution $-(2\pi)^3\delta^{(3)}(0)$. In principle, this could partially cancel the positive contribution from bosonic fields. Cosmological constant problem anyone?!

5.2.1 Fermi-Dirac Statistics

Just as in the bosonic case, we define the vacuum $|0\rangle$ to satisfy,

$$b^s_{\vec{p}}\,|0\rangle = c^s_{\vec{p}}\,|0\rangle = 0 \tag{5.18}$$

Although b and c obey anti-commutation relations, the Hamiltonian (5.17) has nice commutation relations with them. You can check that

$$[H, b^r_{\vec{p}}] = -E_{\vec{p}}\, b^r_{\vec{p}} \quad \text{and} \quad [H, b^{r\,\dagger}_{\vec{p}}] = E_{\vec{p}}\, b^{r\,\dagger}_{\vec{p}}$$
$$[H, c^r_{\vec{p}}] = -E_{\vec{p}}\, c^r_{\vec{p}} \quad \text{and} \quad [H, c^{r\,\dagger}_{\vec{p}}] = E_{\vec{p}}\, c^{r\,\dagger}_{\vec{p}} \tag{5.19}$$

This means that we can again construct a tower of energy eigenstates by acting on the vacuum by $b^{r\,\dagger}_{\vec{p}}$ and $c^{r,\dagger}_{\vec{p}}$ to create particles and antiparticles, just as in the bosonic case.

For example, we have the one-particle states

$$|\vec{p}, r\rangle = b^{r\,\dagger}_{\vec{p}}|0\rangle \tag{5.20}$$

The two particle states now satisfy

$$|\vec{p}_1, r_1; \vec{p}_2, r_2\rangle \equiv b^{r_1\,\dagger}_{\vec{p}_1} b^{r_2\,\dagger}_{\vec{p}_2}|0\rangle = -|\vec{p}_2, r_2; \vec{p}_1, r_1\rangle \tag{5.21}$$

confirming that the particles do indeed obey Fermi-Dirac statistics. In particular, we have the Pauli-Exclusion principle $|\vec{p}, r; \vec{p}, r\rangle = 0$. Finally, if we wanted to be sure about the spin of the particle, we could act with the angular momentum operator (4.96) to confirm that a stationary particle $|\vec{p} = 0, r\rangle$ does indeed carry intrinsic angular momentum 1/2 as expected.

5.3 Dirac's Hole Interpretation

> "In this attempt, the success seems to have been on the side of Dirac rather than logic"
>
> *Pauli on Dirac*

Let's pause our discussion to make a small historical detour. Dirac originally viewed his equation as a relativistic version of the Schrödinger equation, with ψ interpreted as the wavefunction for a single particle with spin. To reinforce this interpretation, he wrote $(i\partial\!\!\!/ - m)\psi = 0$ as

$$i\frac{\partial\psi}{\partial t} = -i\,\vec{\alpha}\cdot\vec{\nabla}\psi + m\beta\psi \equiv \hat{H}\psi \tag{5.22}$$

where $\vec{\alpha} = -\gamma^0\vec{\gamma}$ and $\beta = \gamma^0$. Here the operator $\hat{H}$ is interpreted as the one-particle Hamiltonian. This is a very different viewpoint from the one we now have, where ψ is a classical field that should be quantized. In Dirac's view, the Hamiltonian of the system is $\hat{H}$ defined above, while for us the Hamiltonian is the field operator (5.17).

Let's see where Dirac's viewpoint leads.

With the interpretation of ψ as a single-particle wavefunction, the plane-wave solutions (4.104) and (4.110) to the Dirac equation are thought of as energy eigenstates, with

$$\psi = u(\vec{p})\, e^{-ip\cdot x} \quad \Rightarrow \quad i\frac{\partial\psi}{\partial t} = E_{\vec{p}}\,\psi$$
$$\psi = v(\vec{p})\, e^{+ip\cdot x} \quad \Rightarrow \quad i\frac{\partial\psi}{\partial t} = -E_{\vec{p}}\,\psi \qquad (5.23)$$

which look like positive and negative energy solutions. The spectrum is once again unbounded below; there are states $v(\vec{p})$ with arbitrarily low energy $-E_{\vec{p}}$. At first glance this is disastrous, just like the unbounded field theory Hamiltonian (5.12). Dirac postulated an ingenious solution to this problem: since the electrons are fermions (a fact which is put in by hand to Dirac's theory) they obey the Pauli-exclusion principle. So we could simply stipulate that in the true vacuum of the universe, all the negative energy states are filled. Only the positive energy states are accessible. These filled negative energy states are referred to as the *Dirac sea*. Although you might worry about the infinite negative charge of the vacuum, Dirac argued that only charge differences would be observable (a trick reminiscent of the normal ordering prescription we used for field operators).

Having avoided disaster by floating on an infinite sea comprised of occupied negative energy states, Dirac realized that his theory made a shocking prediction. Suppose that a negative energy state is excited to a positive energy state, leaving behind a hole. The hole would have all the properties of the electron, except it would carry positive charge. After flirting with the idea that it may be the proton, Dirac finally concluded that the hole is a new particle: the positron. Moreover, when a positron comes across an electron, the two can annihilate. Dirac had predicted anti-matter, one of the greatest achievements of theoretical physics. It took only a couple of years before the positron was discovered experimentally in 1932.

Although Dirac's physical insight led him to the right answer, we now understand that the interpretation of the Dirac spinor as a single-particle wavefunction is not really correct. For example, Dirac's argument for anti-matter relies crucially on the particles being fermions while, as we have seen already in this course, anti-particles exist for both fermions and bosons. What we really learn from Dirac's analysis is that there is no consistent way to interpret the Dirac equation as describing a single particle. It is instead to be thought of as a classical field which has only positive energy solutions because the Hamiltonian (4.92) is positive definite. Quantization of this field then gives rise to both particle and anti-particle excitations.

This from Julian Schwinger:

> *"Until now, everyone thought that the Dirac equation referred directly to physical particles. Now, in field theory, we recognize that the equations refer to a sublevel. Experimentally we are concerned with particles, yet the old equations describe fields.... When you begin with field equations, you operate on a level where the particles are not there from the start. It is when you solve the field equations that you see the emergence of particles."*

5.4 Propagators

Let's now move to the Heisenberg picture. We define the spinors $\psi(\vec{x}, t)$ at every point in spacetime such that they satisfy the operator equation

$$\frac{\partial \psi}{\partial t} = i[H, \psi] \tag{5.24}$$

We solve this by the expansion

$$\psi(x) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\vec{p}}}} \left[b^s_{\vec{p}} u^s(\vec{p}) e^{-ip\cdot x} + c^{s\dagger}_{\vec{p}} v^s(\vec{p}) e^{+ip\cdot x} \right]$$

$$\psi^\dagger(x) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\vec{p}}}} \left[b^{s\dagger}_{\vec{p}} u^{s\dagger}(\vec{p}) e^{+ip\cdot x} + c^s_{\vec{p}} v^{s\dagger}(\vec{p}) e^{-ip\cdot x} \right] \tag{5.25}$$

Let's now look at the anti-commutators of these fields. We define the fermionic propagator to be

$$iS_{\alpha\beta} = \{\psi_\alpha(x), \bar{\psi}_\beta(y)\} \tag{5.26}$$

In what follows we will often drop the indices and simply write $iS(x-y) = \{\psi(x), \bar{\psi}(y)\}$, but you should remember that $S(x-y)$ is a 4×4 matrix. Inserting the expansion (5.25), we have

$$\begin{aligned} iS(x-y) &= \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{\sqrt{4E_{\vec{p}}E_{\vec{q}}}} \Big[\{b^s_{\vec{p}}, b^{r\dagger}_{\vec{q}}\} u^s(\vec{p}) \bar{u}^r(\vec{q}) e^{-i(p\cdot x - q\cdot y)} \\ &\qquad + \{c^{s\dagger}_{\vec{p}}, c^r_{\vec{q}}\} v^s(\vec{p}) \bar{v}^r(\vec{q}) e^{+i(p\cdot x - q\cdot y)} \Big] \\ &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \Big[u^s(\vec{p}) \bar{u}^s(\vec{p}) e^{-ip\cdot(x-y)} + v^s(\vec{p}) \bar{v}^s(\vec{p}) e^{+ip\cdot(x-y)} \Big] \\ &\quad \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E} \end{aligned}$$

$$\not\!\! \qquad \text{———} \ \text{——} \quad (p\!\!/ + m)e^{-ip\cdot(x-} \quad \not\!\! (p\!\!/ - m)e^{+ip\cdot(x-} \qquad (5.27)$$

$(2\pi)^3 \ 2E$

where to reach the final line we have used the outer product formulae (4.128) and (4.129). We can then write

$$iS(x-y) = (i\partial\!\!\!/_x + m)(D(x-y) - D(y-x)) \tag{5.28}$$

in terms of the propagator for a real scalar field $D(x-y)$ which, recall, can be written as (2.90)

$$D(x-y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} e^{-ip\cdot(x-y)} \tag{5.29}$$

Some comments:

- For spacelike separated points $(x-y)^2 < 0$, we have already seen that $D(x-y) - D(y-x) = 0$. In the bosonic theory, we made a big deal of this since it ensured that

$$[\varphi(x), \varphi(y)] = 0 \qquad (x-y)^2 < 0 \tag{5.30}$$

  outside the lightcone, which we trumpeted as proof that our theory was causal. However, for fermions we now have

$$\{\psi_\alpha(x), \psi_\beta(y)\} = 0 \qquad (x-y)^2 < 0 \tag{5.31}$$

  outside the lightcone. What happened to our precious causality? The best that we can say is that all our observables are bilinear in fermions, for example the Hamiltonian (5.17). These still commute outside the lightcone. The theory remains causal as long as fermionic operators are not observable. If you think this is a little weak, remember that no one has ever seen a physical measuring apparatus come back to minus itself when you rotate by 360 degrees!

- At least away from singularities, the propagator satisfies

$$(i\partial\!\!\!/_x - m)S(x-y) = 0 \tag{5.32}$$

  which follows from the fact that $(\partial\!\!\!/_x^2 + m^2)D(x-y) = 0$ using the mass shell condition $p^2 = m^2$.

5.5 The Feynman Propagator

By a similar calculation to that above, we can determine the vacuum expectation value,

$$\begin{aligned} \langle 0|\, \psi_\alpha(x)\bar{\psi}_\beta(y)\, |0\rangle &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} (p\!\!\!/ + m)_{\alpha\beta}\, e^{-ip\cdot(x-y)} \\ \langle 0|\, \bar{\psi}_\beta(y)\psi_\alpha(x)\, |0\rangle &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} (p\!\!\!/ - m)_{\alpha\beta}\, e^{+ip\cdot(x-y)} \end{aligned} \tag{5.33}$$

We now define the Feynman propagator $S_F(x-y)$, which is again a 4 × 4 matrix, as the time ordered product,

$$S_F(x-y) = \langle 0|\, T\psi(x)\bar{\psi}(y)\,|0\rangle \equiv \begin{cases} \langle 0|\psi(x)\bar{\psi}(y)|0\rangle & x^0 > y^0 \\ \langle 0| -\bar{\psi}(y)\psi(x)\,|0\rangle & y^0 > x^0 \end{cases} \quad (5.34)$$

Notice the minus sign! It is necessary for Lorentz invariance. When $(x-y)^2 < 0$, there is no invariant way to determine whether $x^0 > y^0$ or $y^0 > x^0$. In this case the minus sign is necessary to make the two definitions agree since $\{\psi(x), \bar{\psi}(y)\} = 0$ outside the lightcone.

We have the 4-momentum integral representation for the Feynman propagator,

$$S_F(x-y) = i\int \frac{d^4p}{(2\pi)^4}\, e^{-ip\cdot(x-y)}\, \frac{\gamma\cdot p + m}{p^2 - m^2 + i\epsilon} \quad (5.35)$$

which satisfies $(i\partial\!\!\!/_x - m)S_F(x-y) = i\delta^{(4)}(x-y)$, so that S_F is a Green's function for the Dirac operator.

The minus sign that we see in (5.34) also occurs for any string of operators inside a time ordered product $T(\ldots)$. While bosonic operators commute inside T, fermionic operators anti-commute. We have this same behaviour for normal ordered products as well, with fermionic operators obeying $:\psi_1\psi_2: = -:\psi_2\psi_1:$. With the understanding that all fermionic operators anti-commute inside T and $::$, Wick's theorem proceeds just as in the bosonic case. We define the contraction

$$\overbrace{\psi(x)\bar{\psi}(y)} = T(\psi(x)\bar{\psi}(y)) - :\psi(x)\bar{\psi}(y): = S_F(x-y) \quad (5.36)$$

5.6 Yukawa Theory

The interaction between a Dirac fermion of mass m and a real scalar field of mass μ is governed by the Yukawa theory,

$$\mathcal{L} = \tfrac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \tfrac{1}{2}\mu^2\varphi^2 + \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \lambda\varphi\bar{\psi}\psi \quad (5.37)$$

which is the proper version of the baby scalar Yukawa theory we looked at in Section 3. Couplings of this type appear in the standard model, between fermions and the Higgs boson. In that context, the fermions can be leptons (such as the electron) or quarks.

Yukawa originally proposed an interaction of this type as an effective theory of nuclear forces. With an eye to this, we will again refer to the φ particles as mesons, and the ψ particles as nucleons. Except, this time, the nucleons have spin. (This is still not a particularly realistic theory of nucleon interactions, not least because we're omitting isospin. Moreover, in Nature the relevant mesons are pions which are pseudoscalars, so a coupling of the form $\varphi\bar{\psi}\gamma^5\psi$ would be more appropriate. We'll turn to this briefly in Section 5.7.3).

Note the dimensions of the various fields. We still have $[\phi] = 1$, but the kinetic terms require that $[\psi] = 3/2$. Thus, unlike in the case with only scalars, the coupling is dimensionless: $[\lambda] = 0$.

We'll proceed as we did in Section 3, firstly computing the amplitude of a particular scattering process then, with that calculation as a guide, writing down the Feynman rules for the theory. We start with:

5.6.1 An Example: Putting Spin on Nucleon Scattering

Let's study $\psi\psi \to \psi\psi$ scattering. This is the same calculation we performed in Section (3.3.3) except now the fermions have spin. Our initial and final states are

$$
\begin{aligned}
|i\rangle &= \sqrt{4E_{\vec p}E_{\vec q}}\, b^{s\dagger}_{\vec p}\, b^{r\dagger}_{\vec q}|0\rangle \equiv |\vec p, s; \vec q, r\rangle \\
|f\rangle &= \sqrt{4E_{\vec p'}E_{\vec q'}}\, b^{s'\dagger}_{\vec p'}\, b^{r'\dagger}_{\vec q'}|0\rangle \equiv |\vec p', s'; \vec q', r'\rangle
\end{aligned} \tag{5.38}
$$

We need to be a little cautious about minus signs, because the $b^\dagger$'s now anti-commute. In particular, we should be careful when we take the adjoint. We have

$$
\langle f| = \sqrt{4E_{\vec p'}E_{\vec q'}}\, \langle 0|\, b^{r'}_{\vec q'}\, b^{s'}_{\vec p'} \tag{5.39}
$$

We want to calculate the order λ^2 terms from the S-matrix element $\langle f|\, S - 1\, |i\rangle$.

$$
\frac{(-i\lambda)^2}{2}\int d^4x_1 d^4x_2\; T\left(\bar\psi(x_1)\psi(x_1)\phi(x_1)\,\bar\psi(x_2)\psi(x_2)\phi(x_2)\right) \tag{5.40}
$$

where, as usual, all fields are in the interaction picture. Just as in the bosonic calculation, the contribution to nucleon scattering comes from the contraction

$$
:\bar\psi(x_1)\psi(x_1)\bar\psi(x_2)\psi(x_2):\; \overbrace{\phi(x_1)\,\phi(x_2)} \tag{5.41}
$$

We just have to be careful about how the spinor indices are contracted. Let's start by looking at how the fermionic operators act on $|i\rangle$. We expand out the ψ fields, leaving the $\bar\psi$ fields alone for now. We may ignore the $c^\dagger$ pieces in ψ since they give no contribution at order λ^2. We have

$$
\begin{aligned}
:\bar\psi(x_1)\psi(x_1)\,\bar\psi(x_2)\psi(x_2):\; b^{s\dagger}_{\vec p}\, b^{r\dagger}_{\vec q}|0\rangle = -\int \frac{d^3k_1\, d^3k_2}{(2\pi)^6}\, [\bar\psi(x_1)\cdot u^m(\vec k_1)]\,[\bar\psi(x_2)\cdot u^n(\vec k_2)] \\
\frac{e^{-ik_1\cdot x_1 - ik_2\cdot x_2}}{\sqrt{4E_{\vec k_1}E_{\vec k_2}}}\, b^m_{\vec k_1}\, b^n_{\vec k_2}\, b^{s\dagger}_{\vec p}\, b^{r\dagger}_{\vec q}|0\rangle
\end{aligned} \tag{5.42}
$$

where we've used square brackets $[\cdot]$ to show how the spinor indices are contracted. The minus sign that sits out front came from moving $\psi(x_1)$ past $\bar\psi(x_2)$. Now anti-commuting the b's past the $b^\dagger$'s, we get

$$= \frac{1}{\sqrt{E_{\vec p}E_{\vec q}}}\Big([\bar\psi(x_1)\cdot u^r(\vec q)]\,[\bar\psi(x_2)\cdot u^s(\vec p)]\,e^{-ip\cdot x_2-iq\cdot x_1} - [\bar\psi(x_1)\cdot u^s(\vec p)]\,[\bar\psi(x_2)\cdot u^r(\vec q)]\,e^{-ip\cdot x_1-iq\cdot x_2}\Big)|0\rangle \quad (5.43)$$

Note, in particular, the relative minus sign that appears between these two terms. Now let's see what happens when we hit this with $\langle f|$. We look at

$$\langle 0|\, b^{r'}_{\vec q'} b^{s'}_{\vec p'}\,[\bar\psi(x_1)\cdot u^r(\vec q)]\,[\bar\psi(x_2)\cdot u^s(\vec p)]\,|0\rangle = \frac{e^{+ip'\cdot x_1+iq'\cdot x_2}}{2\sqrt{E_{\vec p'}E_{\vec q'}}}[\bar u^{s'}(\vec p')\cdot u^r(\vec q)]\,[\bar u^{r'}(\vec q')\cdot u^s(\vec p)] - \frac{e^{+ip'\cdot x_2+iq'\cdot x_1}}{2\sqrt{E_{\vec p'}E_{\vec q'}}}[\bar u^{r'}(\vec q')\cdot u^r(\vec q)]\,[\bar u^{s'}(\vec p')\cdot u^s(\vec p)]$$

The $[\bar\psi(x_1)\cdot u^s(\vec p)]\,[\bar\psi(x_2)\cdot u^r(\vec q)]$ term in (5.43) doubles up with this, cancelling the factor of 1/2 in front of (5.40). Meanwhile, the $1/\sqrt{E}$ terms cancel the relativistic state normalization. Putting everything together, we have the following expression for $\langle f|\,S-1\,|i\rangle$

$$(-i\lambda)^2\int \frac{d^4x_1\,d^4x_2\,d^4k}{(2\pi)^4}\,\frac{ie^{ik\cdot(x_1-x_2)}}{k^2-\mu^2+i\epsilon}\Big([\bar u^{s'}(\vec p')\cdot u^s(\vec p)]\,[\bar u^{r'}(\vec q')\cdot u^r(\vec q)]\,e^{+ix_1\cdot(q'-q)+ix_2\cdot(p'-p)} - [\bar u^{s'}(\vec p')\cdot u^r(\vec q)]\,[\bar u^{r'}(\vec q')\cdot u^s(\vec p)]\,e^{ix_1\cdot(p'-q)+ix_2\cdot(q'-p)}\Big)$$

where we've put the φ propagator back in. Performing the integrals over x_1 and x_2, this becomes,

$$\int d^4k\,\frac{(2\pi)^4 i(-i\lambda)^2}{k^2-\mu^2+i\epsilon}\Big([\bar u^{s'}(\vec p')\cdot u^s(\vec p)]\,[\bar u^{r'}(\vec q')\cdot u^r(\vec q)]\,\delta^{(4)}(q'-q+k)\,\delta^{(4)}(p'-p-k) - [\bar u^{s'}(\vec p')\cdot u^r(\vec q)]\,[\bar u^{r'}(\vec q')\cdot u^s(\vec p)]\,\delta^{(4)}(p'-q+k)\,\delta^{(4)}(q'-p-k)\Big)$$

And we're almost there! Finally, writing the S-matrix element in terms of the amplitude in the usual way, $\langle f|\,S-1\,|i\rangle = i\mathcal{A}(2\pi)^4\delta^{(4)}(p+q-p'-q')$, we have

$$\mathcal{A} = (-i\lambda)^2\left(\frac{[\bar u^{s'}(\vec p')\cdot u^s(\vec p)]\,[\bar u^{r'}(\vec q')\cdot u^r(\vec q)]}{(p'-p)^2-\mu^2+i\epsilon} - \frac{[\bar u^{s'}(\vec p')\cdot u^r(\vec q)]\,[\bar u^{r'}(\vec q')\cdot u^s(\vec p)]}{(q'-p)^2-\mu^2+i\epsilon}\right)$$

which is our final answer for the amplitude.

5.7 Feynman Rules for Fermions

It's important to bear in mind that the calculation we just did kind of blows. Thankfully the Feynman rules will once again encapsulate the combinatoric complexities and make life easier for us. The rules to compute amplitudes are the following

- To each incoming fermion with momentum p and spin r, we associate a spinor $u^r(\vec{p})$. For outgoing fermions we associate $\bar{u}^r(\vec{p})$.

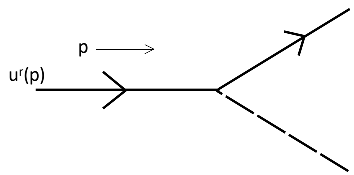


Figure 21: An incoming fermion

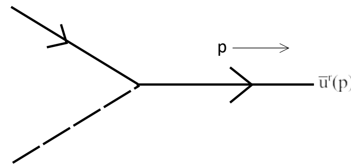


Figure 22: An outgoing fermion

- To each incoming anti-fermion with momentum p and spin r, we associate a spinor $\bar{v}^r(\vec{p})$. For outgoing anti-fermions we associate $v^r(\vec{p})$.

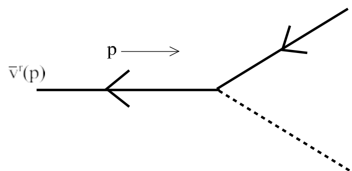


Figure 23: An incoming anti-fermion

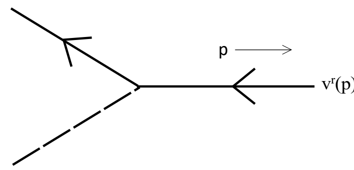
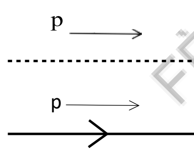


Figure 24: An outgoing anti-fermion

- Each vertex gets a factor of $-i\lambda$.
- Each internal line gets a factor of the relevant propagator.



$$\frac{i}{p^2-\mu^2+i\epsilon} \quad \text{for scalars}$$

$$\frac{i(\not{p}+m)}{p^2-m^2+i\epsilon} \quad \text{for fermions} \tag{5.44}$$

The arrows on the fermion lines must flow consistently through the diagram (this ensures fermion number conservation). Note that the fermionic propagator is a 4×4 matrix. The matrix indices are contracted at each vertex, either with further propagators, or with external spinors u, $\bar{u}$, v or $\bar{v}$.

- Impose momentum conservation at each vertex, and integrate over undetermined loop momenta.
- Add extra minus signs for statistics. Some examples will be given below.

5.7.1 Examples

Let's run through the same examples we did for the scalar Yukawa theory. Firstly, we have

Nucleon Scattering

For the example we worked out previously, the two lowest order Feynman diagrams are shown in Figure 25. We've drawn the second Feynman diagram with the legs crossed

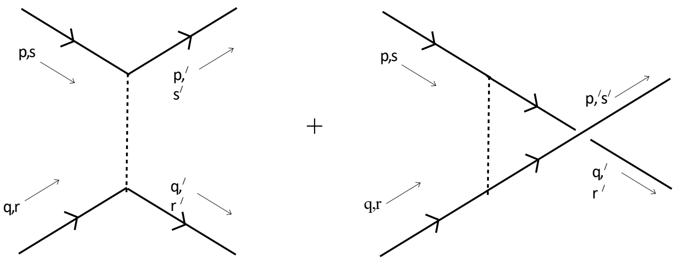


Figure 25: The two Feynman diagrams for nucleon scattering

to emphasize the fact that it picks up a minus sign due to statistics. (Note that the way the legs point in the Feynman diagram doesn't tell us the direction in which the particles leave the scattering event: the momentum label does that. The two diagrams above are different because the incoming legs are attached to different outgoing legs). Using the Feynman rules we can read off the amplitude.

$$\mathcal{A} = (-i\lambda)^2 \left(\frac{[\bar{u}^{s'}(\vec{p}\,') \cdot u^{s}(\vec{p})]\,[\bar{u}^{r'}(\vec{q}\,') \cdot u^{r}(\vec{q})]}{(p-p')^2-\mu^2} - \frac{[\bar{u}^{s'}(\vec{p}\,') \cdot u^{r}(\vec{q})]\,[\bar{u}^{r'}(\vec{q}\,') \cdot u^{s}(\vec{p})]}{(p-q')^2-\mu^2} \right) \quad (5.45)$$

The denominators in each term are due to the meson propagator, with the momentum determined by conservation at each vertex. This agrees with the amplitude we computed earlier using Wick's theorem.

Nucleon to Meson Scattering

Let's now look at $\psi\bar{\psi} \to \phi\phi$. The two lowest order Feynman diagrams are shown in Figure 26. Applying the Feynman rules, we have

$$\mathcal{A} = (-i\lambda)^2 \left[\frac{\bar{v}^{r}(\vec{q})[\gamma^\mu(p_\mu - p'_\mu) + m]u^s(\vec{p})}{(p-p')^2 - m^2} + \frac{\bar{v}^{r}(\vec{q})[\gamma^\mu(p_\mu - q'_\mu) + m]u^s(\vec{p})}{(p-q')^2 - m^2} \right]$$

Since the internal line is now a fermion, the propagator contains $\gamma_\mu(p_\mu - p'_\mu) + m$ factors. This is a 4×4 matrix which sits on the top, sandwiched between the two external spinors. Now the exchange statistics applies to the final meson states. These are bosons and, correspondingly, there is no relative minus sign between the two diagrams.

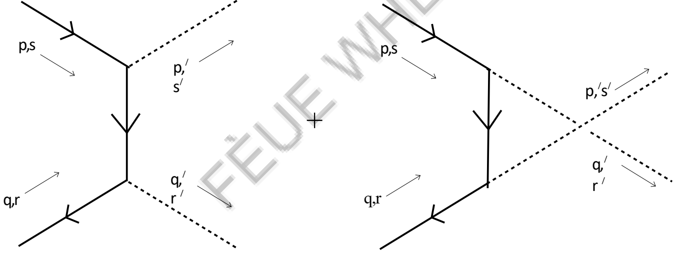


Figure 26: The two Feynman diagrams for nucleon to meson scattering

Nucleon-Anti-Nucleon Scattering

For $\psi\bar{\psi} \to \psi\bar{\psi}$, the two lowest order Feynman diagrams are of two distinct types, just like in the bosonic case. They are shown in Figure 27.

The corresponding amplitude is given by,

$$\mathcal{A} = (-i\lambda)^2 \left(-\frac{[\bar{u}^{s'}_{\vec{p}\,'} \cdot u^{s}_{\vec{p}}]\,[\bar{v}^{r}_{\vec{q}} \cdot v^{r'}_{\vec{q}\,'}]}{(p-p')^2 - \mu^2} + \frac{[\bar{v}^{r}_{\vec{q}} \cdot u^{s}_{\vec{p}}]\,[\bar{u}^{s'}_{\vec{p}\,'} \cdot v^{r'}_{\vec{q}\,'}]}{(p+q)^2 - \mu^2 + i\epsilon} \right) \qquad (5.46)$$

As in the bosonic diagrams, there is again the difference in the momentum dependence in the denominator. But now the difference in the diagrams is also reflected in the spinor contractions in the numerator.

More subtle are the minus signs. The fermionic statistics mean that the first diagram has an extra minus sign relative to the $\psi\psi$ scattering of Figure 25. Since this minus sign will be important when we come to figure out whether the Yukawa force is attractive or repulsive, let's go back to basics and see where it comes from. The initial and final states for this scattering process are

$$|i\rangle = \sqrt{4E_{\vec{p}}E_{\vec{q}}}\, b^{s\,\dagger}_{\vec{p}} c^{r\,\dagger}_{\vec{q}} |0\rangle \equiv |\vec{p}, s; \vec{q}, r\rangle$$
$$|f\rangle = \sqrt{4E_{\vec{p}\,'}E_{\vec{q}\,'}}\, b^{s'\,\dagger}_{\vec{p}\,'} c^{r'\,\dagger}_{\vec{q}\,'} |0\rangle \equiv |\vec{p}\,', s'; \vec{q}\,', r'\rangle \qquad (5.47)$$

The ordering of $b^\dagger$ and $c^\dagger$ in these states is crucial and reflects the scattering $\psi\bar{\psi} \to \psi\bar{\psi}$, as opposed to $\psi\bar{\psi} \to \bar{\psi}\psi$ which would differ by a minus sign. The first diagram in Figure 27 comes from the term in the perturbative expansion,

$$\langle f|:\bar{\psi}(x_1)\psi(x_1)\bar{\psi}(x_2)\psi(x_2): b^{s\,\dagger}_{\vec{p}} c^{r\,\dagger}_{\vec{q}}|0\rangle \sim \langle f|\,[\bar{v}^m_{\vec{k}_1}\cdot\psi(x_1)]\,[\bar{\psi}(x_2)\cdot u^n_{\vec{k}_2}]\,c^m_{\vec{k}_1} b^n_{\vec{k}_2} b^{s\,\dagger}_{\vec{p}} c^{r\,\dagger}_{\vec{q}}|0\rangle$$

where we've neglected a bunch of objects in this equation like $\int d^4k_i$ and exponential factors because we only want to keep track of the minus signs. Moving the annihilation operators past the creation operators, we have

$$+\langle f|\,[\bar{v}^r_{\vec{q}}\cdot\psi(x_1)]\,[\bar{\psi}(x_2)\cdot u^s_{\vec{p}}]\,|0\rangle \qquad (5.48)$$

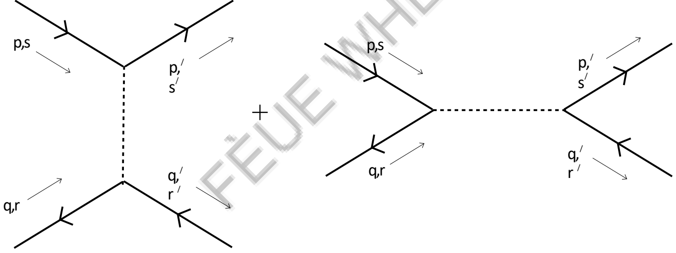


Figure 27: The two Feynman diagrams for nucleon-anti-nucleon scattering

Repeating the process by expanding out the $\psi(x_1)$ and $\bar{\psi}(x_2)$ fields and moving them to the left to annihilate $\langle f|$, we have

$$\langle 0| c^{r'}_{\vec{q}'} b^{s'}_{\vec{p}'} c^{m\dagger}_{\vec{l}_1} b^{n\dagger}_{\vec{l}_2} [\bar{v}^r(\vec{q}) \cdot v^m(\vec{l}_1)] [\bar{u}^n(\vec{l}_2) \cdot u^s(\vec{p})] |0\rangle \sim -[\bar{v}^r(\vec{q}) \cdot v^{r'}(\vec{q}')] [\bar{u}^{s'}(\vec{p}') \cdot u^s(\vec{p})]$$

where the minus sign has appeared from anti-commuting $c^{m\dagger}_{\vec{l}_1}$ past $b^{s'}_{\vec{p}'}$. This is the overall minus sign found in (5.46). One can also follow similar contractions to compute the second diagram in Figure 27.

Meson Scattering

Finally, we can also compute the scattering of $\phi\phi \to \phi\phi$ which, as in the bosonic case, picks up its leading contribution at one-loop. The amplitude for the diagram shown in the figure is

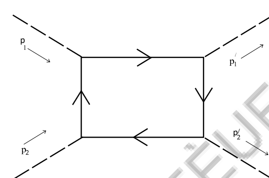
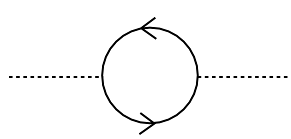


Figure 28:

$$iA = -(-i\lambda)^4 \int \frac{d^4k}{(2\pi)^4} \mathrm{Tr}\, \frac{\not{k} + m}{(k^2 - m^2 + i\epsilon)} \frac{\not{k} + \not{p}'_1 + m}{((k + p'_1)^2 - m^2 + i\epsilon)} \times \frac{\not{k} + \not{p}'_1 - \not{p}_1 + m}{((k + p'_1 - p_1)^2 - m^2 + i\epsilon)} \frac{\not{k} - \not{p}'_2 + m}{((k - p'_2)^2 - m^2 + i\epsilon)}$$

Notice that the high momentum limit of the integral is $\int d^4k/k^4$, which is no longer finite, but diverges logarithmically. You will have to wait until next term to make sense of this integral.

There's an overall minus sign sitting in front of this amplitude. This is a generic feature of diagrams with fermions running in loops: each fermionic loop in a diagram gives rise to an extra minus sign. We can see this rather simply in the diagram



which involves the expression

$$\bar{\psi}_\alpha(x)\,\psi_\alpha(x)\bar{\psi}_\beta(y)\,\psi_\beta(y) = -\,\psi_\beta(y)\bar{\psi}_\alpha(x)\,\psi_\alpha(x)\bar{\psi}_\beta(y)$$
$$= -\mathrm{Tr}\,(S_F(y-x)S_F(x-y))$$

After passing the fermion fields through each other, a minus sign appears, sitting in front of the two propagators.

5.7.2 The Yukawa Potential Revisited

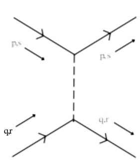
We saw in Section 3.5.2, that the exchange of a real scalar particle gives rise to a universally attractive Yukawa potential between two spin zero particles. Does the same hold for the spin 1/2 particles?

Recall that the strategy to compute the potential is to take the non-relativistic limit of the scattering amplitude, and compare with the analogous result from quantum mechanics. Our new amplitude now also includes the spinor degrees of freedom $u(\vec{p})$ and $v(\vec{p})$. In the non-relativistic limit, $p \to (m, \vec{p})$, and

$$u(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi \\ \sqrt{p\cdot\bar{\sigma}}\,\xi \end{pmatrix} \to \sqrt{m}\begin{pmatrix} \xi \\ \xi \end{pmatrix}$$

$$v(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi \\ -\sqrt{p\cdot\bar{\sigma}}\,\xi \end{pmatrix} \to \sqrt{m}\begin{pmatrix} \xi \\ -\xi \end{pmatrix} \tag{5.49}$$

In this limit, the spinor contractions in the amplitude for $\psi\psi \to \psi\psi$ scattering (5.45) become $\bar{u}^{s'} \cdot u^s = 2m\delta^{ss'}$ and the amplitude is



$$= -i(-i\lambda)^2\,(2m)\left(\frac{\delta^{s's}\delta^{r'r}}{(\vec{p}-\vec{p}')^2+\mu^2} - \frac{\delta^{s'r}\delta^{r's}}{(\vec{p}-\vec{q}')^2+\mu^2}\right) \tag{5.50}$$

The δ symbols tell us that spin is conserved in the non-relativistic limit, while the momentum dependence is the same as in the bosonic case, telling us that once again the particles feel an attractive Yukawa potential,

$$U(\vec{r}) = -\frac{\lambda^2 e^{-\mu r}}{4\pi r} \tag{5.51}$$

Repeating the calculation for $\psi\bar{\psi} \to \psi\bar{\psi}$, there are two minus signs which cancel each other. The first is the extra overall minus sign in the scattering amplitude (5.46), due to the fermionic nature of the particles. The second minus sign comes from the non-relativistic limit of the spinor contraction for anti-particles in (5.46), which is $\bar{v}^{s'} \cdot v^{s} = -2m\delta^{ss'}$. These two signs cancel, giving us once again an attractive Yukawa potential (5.51).

5.7.3 Pseudo-Scalar Coupling

Rather than the standard Yukawa coupling, we could instead consider

$$\mathcal{L}_{\text{Yuk}} = -\lambda \phi \bar{\psi} \gamma^5 \psi \tag{5.52}$$

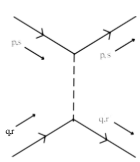
This still preserves parity if ϕ is a pseudoscalar, i.e.

$$P : \phi(\vec{x}, t) \to -\phi(-\vec{x}, t) \tag{5.53}$$

We can compute in this theory very simply: the Feynman rule for the interaction vertex is now changed to a factor of $-i\lambda\gamma^5$. For example, the Feynman diagrams for $\psi\psi \to \psi\psi$ scattering are again given by Figure 25, with the amplitude now

$$\mathcal{A} = (-i\lambda)^2 \left(\frac{[\bar{u}^{s'}(\vec{p}\,')\gamma^5 u^s(\vec{p})]\,[\bar{u}^{r'}(\vec{q}\,')\gamma^5 u^r(\vec{q})]}{(p-p')^2 - \mu^2} - \frac{[\bar{u}^{s'}(\vec{q}\,')\gamma^5 u^r(\vec{q})]\,[\bar{u}^{r'}(\vec{p}\,')\gamma^5 u^s(\vec{p})]}{(p-q')^2 - \mu^2} \right)$$

We could again try to take the non-relativistic limit for this amplitude. But this time, things work a little differently. Using the expressions for the spinors (5.49), we have $\bar{u}^s \gamma^5 u^s \to 0$ in the non-relativistic limit. To find the non-relativistic amplitude, we must go to next to leading order. One can easily check that $\bar{u}^{s'}(\vec{p}\,')\gamma^5 u^s(\vec{p}) \to m\, \xi^{s'T}(\vec{p} - \vec{p}\,')\cdot\vec{\sigma}\,\xi^s$. So, in the non-relativistic limit, the leading order amplitude arising from pseudoscalar exchange is given by a spin-spin coupling,



$$\to +im(-i\lambda)^2 \frac{[\xi^{s'T}(\vec{p} - \vec{p}\,')\cdot\vec{\sigma}\,\xi^s]\,[\xi^{r'T}(\vec{p} - \vec{p}\,')\cdot\vec{\sigma}\,\xi^r]}{(\vec{p} - \vec{p}\,')^2 + \mu^2} \tag{5.54}$$

6. Quantum Electrodynamics

In this section we finally get to quantum electrodynamics (QED), the theory of light interacting with charged matter. Our path to quantization will be as before: we start with the free theory of the electromagnetic field and see how the quantum theory gives rise to a photon with two polarization states. We then describe how to couple the photon to fermions and to bosons.

6.1 Maxwell's Equations

The Lagrangian for Maxwell's equations in the absence of any sources is simply

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} \tag{6.1}$$

where the field strength is defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{6.2}$$

The equations of motion which follow from this Lagrangian are

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) = -\partial_\mu F^{\mu\nu} = 0 \tag{6.3}$$

Meanwhile, from the definition of $F_{\mu\nu}$, the field strength also satisfies the Bianchi identity

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \tag{6.4}$$

To make contact with the form of Maxwell's equations you learn about in high school, we need some 3-vector notation. If we define $A^\mu = (\varphi, \vec{A})$, then the electric field $\vec{E}$ and magnetic field $\vec{B}$ are defined by

$$\vec{E} = -\nabla\varphi - \frac{\partial \vec{A}}{\partial t} \quad \text{and} \quad \vec{B} = \nabla \times \vec{A} \tag{6.5}$$

which, in terms of $F_{\mu\nu}$, becomes

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \tag{6.6}$$

The Bianchi identity (6.4) then gives two of Maxwell's equations,

$$\nabla \cdot \vec{B} = 0 \quad \text{and} \quad \frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E} \tag{6.7}$$

These remain true even in the presence of electric sources. Meanwhile, the equations of motion give the remaining two Maxwell equations,

$$\nabla \cdot \vec{E} = 0 \quad \text{and} \quad \frac{\partial \vec{E}}{\partial t} = \nabla \times \vec{B} \tag{6.8}$$

As we will see shortly, in the presence of charged matter these equations pick up extra terms on the right-hand side.

6.1.1 Gauge Symmetry

The massless vector field A_μ has 4 components, which would naively seem to tell us that the gauge field has 4 degrees of freedom. Yet we know that the photon has only two degrees of freedom which we call its polarization states. How are we going to resolve this discrepancy? There are two related comments which will ensure that quantizing the gauge field A_μ gives rise to 2 degrees of freedom, rather than 4.

- The field A_0 has no kinetic term $\dot{A}_0$ in the Lagrangian: it is not dynamical. This means that if we are given some initial data A_i and $\dot{A}_i$ at a time t_0, then the field A_0 is fully determined by the equation of motion $\nabla \cdot \vec{E} = 0$ which, expanding out, reads

$$\nabla^2 A_0 + \nabla \cdot \frac{\partial \vec{A}}{\partial t} = 0 \tag{6.9}$$

This has the solution

$$A_0(\vec{x}) = \int d^3x' \, \frac{(\nabla \cdot \partial \vec{A}/\partial t)(\vec{x}\,')}{4\pi|\vec{x} - \vec{x}\,'|} \tag{6.10}$$

So A_0 is not independent: we don't get to specify A_0 on the initial time slice. It looks like we have only 3 degrees of freedom in A_μ rather than 4. But this is still one too many.

- The Lagrangian (6.3) has a *very* large symmetry group, acting on the vector potential as

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \lambda(x) \tag{6.11}$$

for any function $\lambda(x)$. We'll ask only that $\lambda(x)$ dies off suitably quickly at spatial $\vec{x} \to \infty$. We call this a *gauge symmetry*. The field strength is invariant under the gauge symmetry:

$$F_{\mu\nu} \to \partial_\mu(A_\nu + \partial_\nu \lambda) - \partial_\nu(A_\mu + \partial_\mu \lambda) = F_{\mu\nu} \tag{6.12}$$

So what are we to make of this? We have a theory with an infinite number of symmetries, one for each function $\lambda(x)$. Previously we only encountered symmetries which act the same at all points in spacetime, for example $\psi \to e^{i\alpha}\psi$ for a complex scalar field. Noether's theorem told us that these symmetries give rise to conservation laws. Do we now have an infinite number of conservation laws?

The answer is no! Gauge symmetries have a very different interpretation than the global symmetries that we make use of in Noether's theorem. While the latter take a physical state to another physical state with the same properties, the gauge symmetry is to be viewed as a redundancy in our description. That is, two states related by a gauge symmetry are to be identified: they are the same physical state. (There is a small caveat to this statement which is explained in Section 6.3.1). One way to see that this interpretation is necessary is to notice that Maxwell's equations are not sufficient to specify the evolution of A_μ. The equations read,

$$[\eta_{\mu\nu}(\partial^\rho\partial_\rho) - \partial_\mu\partial_\nu]\, A^\nu = 0 \tag{6.13}$$

But the operator $[\eta_{\mu\nu}(\partial^\rho\partial_\rho) - \partial_\mu\partial_\nu]$ is not invertible: it annihilates any function of the form $\partial_\mu\lambda$. This means that given any initial data, we have no way to uniquely determine A_μ at a later time since we can't distinguish between A_μ and $A_\mu + \partial_\mu\lambda$. This would be problematic if we thought that A_μ is a physical object. However, if we're happy to identify A_μ and $A_\mu + \partial_\mu\lambda$ as corresponding to the same physical state, then our problems disappear.

Since gauge invariance is a redundancy of the system, we might try to formulate the theory purely in terms of the local, physical, gauge invariant objects $\vec{E}$ and $\vec{B}$. This is fine for the free classical theory: Maxwell's equations were, after all, first written in terms of $\vec{E}$ and $\vec{B}$. But it is not possible to describe certain quantum phenomena, such as the Aharonov-Bohm effect, without using the gauge potential A_μ. We will see shortly that we also require the gauge potential to describe classically charged fields. To describe Nature, it appears that we have to introduce quantities A_μ that we can never measure.

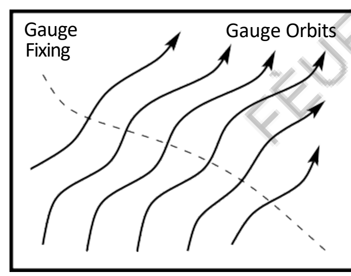


Figure 29:

The picture that emerges for the theory of electromagnetism is of an enlarged phase space, foliated by gauge orbits as shown in the figure. All states that lie along a given line can be reached by a gauge transformation and are identified. To make progress, we pick a representative from each gauge orbit. It doesn't matter which representative we pick — after all, they're all physically equivalent. But we should make sure that we pick a "good" gauge, in which we cut the orbits.

Different representative configurations of a physical state are called different *gauges*. There are many possibilities, some of which will be more useful in different situations. Picking a gauge is rather like picking coordinates that are adapted to a particular problem. Moreover, different gauges often reveal slightly different aspects of a problem. Here we'll look at two different gauges:

- **Lorentz Gauge:** $\partial_\mu A^\mu = 0$

  To see that we can always pick a representative configuration satisfying $\partial_\mu A^\mu = 0$, suppose that we're handed a gauge field ${}_\mu A^r$ satisfying $\partial_\mu (A^r)^\mu = f(x)$. Then we choose $A_\mu = A^r + \partial_\mu \lambda$, where

$$\partial_\mu \partial^\mu \lambda = -f \tag{6.14}$$

  This equation always has a solution. In fact this condition doesn't pick a unique representative from the gauge orbit. We're always free to make further gauge transformations with $\partial_\mu \partial^\mu \lambda = 0$, which also has non-trivial solutions. As the name suggests, the Lorentz gauge[3] has the advantage that it is Lorentz invariant.

- **Coulomb Gauge:** $\nabla \cdot \vec{A} = 0$

  We can make use of the residual gauge transformations in Lorentz gauge to pick $\nabla \cdot \vec{A} = 0$. (The argument is the same as before). Since A_0 is fixed by (6.10), we have as a consequence

$$A_0 = 0 \tag{6.15}$$

  (This equation will no longer hold in Coulomb gauge in the presence of charged matter). Coulomb gauge breaks Lorentz invariance, so may not be ideal for some purposes. However, it is very useful to exhibit the physical degrees of freedom: the 3 components of $\vec{A}$ satisfy a single constraint: $\nabla \cdot \vec{A} = 0$, leaving behind just 2 degrees of freedom. These will be identified with the two polarization states of the photon. Coulomb gauge is sometimes called radiation gauge.

[3]Named after Lorenz who had the misfortune to be one letter away from greatness.

6.2 The Quantization of the Electromagnetic Field

In the following we shall quantize free Maxwell theory twice: once in Coulomb gauge, and again in Lorentz gauge. We'll ultimately get the same answers and, along the way, see that each method comes with its own subtleties.

The first of these subtleties is common to both methods and comes when computing the momentum π^μ conjugate to A_μ,

$$\pi^0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = 0$$
$$\pi^i = \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = -F^{0i} \equiv E^i \tag{6.16}$$

so the momentum π^0 conjugate to A_0 vanishes. This is the mathematical consequence of the statement we made above: A_0 is not a dynamical field. Meanwhile, the momentum conjugate to A_i is our old friend, the electric field. We can compute the Hamiltonian,

$$H = \int d^3x\, \pi^i \dot{A}_i - \mathcal{L}$$
$$= \int d^3x\, \tfrac{1}{2}\vec{E}\cdot\vec{E} + \tfrac{1}{2}\vec{B}\cdot\vec{B} - A_0(\nabla\cdot\vec{E}) \tag{6.17}$$

So A_0 acts as a Lagrange multiplier which imposes Gauss' law

$$\nabla\cdot\vec{E} = 0 \tag{6.18}$$

which is now a constraint on the system in which $\vec{A}$ are the physical degrees of freedom. Let's now see how to treat this system using different gauge fixing conditions.

6.2.1 Coulomb Gauge

In Coulomb gauge, the equation of motion for $\vec{A}$ is

$$\partial_\mu \partial^\mu \vec{A} = 0 \tag{6.19}$$

which we can solve in the usual way,

$$\vec{A} = \int \frac{d^3p}{(2\pi)^3}\, \vec{\xi}(\vec{p})\, e^{ip\cdot x} \tag{6.20}$$

with $p_0^2 = |\vec{p}|^2$. The constraint $\nabla\cdot\vec{A} = 0$ tells us that $\vec{\xi}$ must satisfy

$$\vec{\xi}\cdot\vec{p} = 0 \tag{6.21}$$

which means that $\vec{\xi}$ is perpendicular to the direction of motion $\vec{p}$. We can pick $\vec{\xi}(\vec{p})$ to be a linear combination of two orthonormal vectors $\vec{\epsilon}_r$, $r = 1, 2$, each of which satisfies $\vec{\epsilon}_r(\vec{p}) \cdot \vec{p} = 0$ and

$$\vec{\epsilon}_r(\vec{p}) \cdot \vec{\epsilon}_s(\vec{p}) = \delta_{rs} \qquad r, s = 1, 2 \tag{6.22}$$

These two vectors correspond to the two polarization states of the photon. It's worth pointing out that you can't consistently pick a continuous basis of polarization vectors for every value of $\vec{p}$ because you can't comb the hair on a sphere. But this topological fact doesn't cause any complications in computing QED scattering processes.

To quantize we turn the Poisson brackets into commutators. Naively we would write

$$[A_i(\vec{x}), A_j(\vec{y})] = [E^i(\vec{x}), E^j(\vec{y})] = 0$$
$$[A_i(\vec{x}), E^j(\vec{y})] = i\delta^j_i\, \delta^{(3)}(\vec{x} - \vec{y}) \tag{6.23}$$

But this can't quite be right, because it's not consistent with the constraints. We still want to have $\nabla \cdot \vec{A} = \nabla \cdot \vec{E} = 0$, now imposed on the operators. But from the commutator relations above, we see

$$[\nabla \cdot \vec{A}(\vec{x}), \nabla \cdot \vec{E}(\vec{y})] = i\nabla^2 \delta^{(3)}(\vec{x} - \vec{y}) \neq 0 \tag{6.24}$$

What's going on? In imposing the commutator relations (6.23) we haven't correctly taken into account the constraints. In fact, this is a problem already in the classical theory, where the Poisson bracket structure is already altered[4]. The correct Poisson bracket structure leads to an alteration of the last commutation relation,

$$[A_i(\vec{x}), E_j(\vec{y})] = i\left(\delta_{ij} - \frac{\partial_i\partial_j}{\nabla^2}\right)\delta^{(3)}(\vec{x} - \vec{y}) \tag{6.25}$$

To see that this is now consistent with the constraints, we can rewrite the right-hand side of the commutator in momentum space,

$$[A_i(\vec{x}), E_j(\vec{y})] = i\int \frac{d^3p}{(2\pi)^3}\left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2}\right) e^{i\vec{p}\cdot(\vec{x}-\vec{y})} \tag{6.26}$$

which is now consistent with the constraints, for example

$$[\partial_i A_i(\vec{x}), E_j(\vec{y})] = i\int \frac{d^3p}{(2\pi)^3}\left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2}\right) i p_i\, e^{i\vec{p}\cdot(\vec{x}-\vec{y})} = 0 \tag{6.27}$$

[4]For a nice discussion of the classical and quantum dynamics of constrained systems, see the small book by Paul Dirac, "*Lectures on Quantum Mechanics*"

We now write $\vec{A}$ in the usual mode expansion,

$$\vec{A}(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{r=1}^{2} \vec{\epsilon}_r(\vec{p}) \left[a^r_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} + a^{r\,\dagger}_{\vec{p}}\, e^{-i\vec{p}\cdot\vec{x}} \right]$$

$$\vec{E}(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} (-i) \sqrt{\frac{|\vec{p}|}{2}} \sum_{r=1}^{2} \vec{\epsilon}_r(\vec{p}) \left[a^r_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} - a^{r\,\dagger}_{\vec{p}}\, e^{-i\vec{p}\cdot\vec{x}} \right] \qquad (6.28)$$

where, as before, the polarization vectors satisfy

$$\vec{\epsilon}_r(\vec{p}) \cdot \vec{p} = 0 \quad \text{and} \quad \vec{\epsilon}_r(\vec{p}) \cdot \vec{\epsilon}_s(\vec{p}) = \delta_{rs} \qquad (6.29)$$

It is not hard to show that the commutation relations (6.25) are equivalent to the usual commutation relations for the creation and annihilation operators,

$$[a^r_{\vec{p}}, a^s_{\vec{q}}] = [a^{r\,\dagger}_{\vec{p}}, a^{s\,\dagger}_{\vec{q}}] = 0$$

$$[a^r_{\vec{p}}, a^{s\,\dagger}_{\vec{q}}] = (2\pi)^3 \delta^{rs}\, \delta^{(3)}(\vec{p} - \vec{q}) \qquad (6.30)$$

where, in deriving this, we need the completeness relation for the polarization vectors,

$$\sum_{r=1}^{2} \epsilon^i_r(\vec{p})\, \epsilon^j_r(\vec{p}) = \delta^{ij} - \frac{p^i p^j}{|\vec{p}|^2} \qquad (6.31)$$

You can easily check that this equation is true by acting on both sides with a basis of vectors $(\vec{\epsilon}_1(\vec{p}), \vec{\epsilon}_2(\vec{p}), \vec{p})$.

We derive the Hamiltonian by substituting (6.28) into (6.17). The last term vanishes in Coulomb gauge. After normal ordering, and playing around with $\vec{\epsilon}_r$ polarization vectors, we get the simple expression

$$H = \int \frac{d^3p}{(2\pi)^3} |\vec{p}| \sum_{r=1}^{2} a^{r\,\dagger}_{\vec{p}}\, a^r_{\vec{p}} \qquad (6.32)$$

The Coulomb gauge has the advantage that the physical degrees of freedom are manifest. However, we've lost all semblance of Lorentz invariance. One place where this manifests itself is in the propagator for the fields $A_i(x)$ (in the Heisenberg picture). In Coulomb gauge the propagator reads

$$D^{tr}_{ij}(x - y) \equiv \langle 0 |\, T A_i(x) A_j(y)\, | 0 \rangle = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 + i\epsilon} \left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2} \right) e^{-ip\cdot(x-y)} \qquad (6.33)$$

The *tr* superscript on the propagator refers to the "transverse" part of the photon. When we turn to the interacting theory, we will have to fight to massage this propagator into something a little nicer.

6.2.2 Lorentz Gauge

We could try to work in a Lorentz invariant fashion by imposing the Lorentz gauge condition $\partial_\mu A^\mu = 0$. The equations of motion that follow from the action are then

$$\partial_\mu \partial^\mu A^\nu = 0 \tag{6.34}$$

Our approach to implementing Lorentz gauge will be a little different from the method we used in Coulomb gauge. We choose to change the theory so that (6.34) arises directly through the equations of motion. We can achieve this by taking the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu)^2 \tag{6.35}$$

The equations of motion coming from this action are

$$\partial_\mu F^{\mu\nu} + \partial^\nu (\partial_\mu A^\mu) = \partial_\mu \partial^\mu A^\nu = 0 \tag{6.36}$$

(In fact, we could be a little more general than this, and consider the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\alpha} (\partial_\mu A^\mu)^2 \tag{6.37}$$

with arbitrary α and reach similar conclusions. The quantization of the theory is independent of α and, rather confusingly, different choices of α are sometimes also referred to as different "gauges". We will use $\alpha = 1$, which is called "Feynman gauge". The other common choice, $\alpha = 0$, is called "Landau gauge".)

Our plan will be to quantize the theory (6.36), and only later impose the constraint $\partial_\mu A^\mu = 0$ in a suitable manner on the Hilbert space of the theory. As we'll see, we will also have to deal with the residual gauge symmetry of this theory which will prove a little tricky. At first, we can proceed very easily, because both π^0 and π^i are dynamical:

$$\begin{aligned} \pi^0 &= \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = -\partial_\mu A^\mu \\ \pi^i &= \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = \partial^i A^0 - \dot{A}^i \end{aligned} \tag{6.38}$$

Turning these classical fields into operators, we can simply impose the usual commutation relations,

$$\begin{aligned} [A_\mu(\vec{x}), A_\nu(\vec{y})] &= [\pi^\mu(\vec{x}), \pi^\nu(\vec{y})] = 0 \\ [A_\mu(\vec{x}), \pi_\nu(\vec{y})] &= i\eta_{\mu\nu}\, \delta^{(3)}(\vec{x} - \vec{y}) \end{aligned} \tag{6.39}$$

and we can make the usual expansion in terms of creation and annihilation operators and 4 polarization vectors $(\epsilon_\mu)^\lambda$, with $\lambda = 0, 1, 2, 3$.

$$\begin{aligned} A_\mu(\vec{x}) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda(\vec{p}) \left[a_{\vec{p}}^\lambda e^{i\vec{p}\cdot\vec{x}} + a_{\vec{p}}^{\lambda\,\dagger} e^{-i\vec{p}\cdot\vec{x}} \right] \\ \pi^\mu(\vec{x}) &= \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{|\vec{p}|}{2}}\,(+i) \sum_{\lambda=0}^{3} (\epsilon^\mu)^\lambda(\vec{p}) \left[a_{\vec{p}}^\lambda e^{i\vec{p}\cdot\vec{x}} - a_{\vec{p}}^{\lambda\,\dagger} e^{-i\vec{p}\cdot\vec{x}} \right] \end{aligned} \tag{6.40}$$

Note that the momentum π^μ comes with a factor of $(+i)$, rather than the familiar $(-i)$ that we've seen so far. This can be traced to the fact that the momentum (6.38) for the classical fields takes the form $\pi^\mu = -\dot{A}^\mu +$ In the Heisenberg picture, it becomes clear that this descends to $(+i)$ in the definition of momentum.

There are now four polarization 4-vectors $\epsilon^\lambda(\vec{p})$, instead of the two polarization 3-vectors that we met in the Coulomb gauge. Of these four 4-vectors, we pick ϵ^0 to be timelike, while $\epsilon^{1,2,3}$ are spacelike. We pick the normalization

$$\epsilon^\lambda \cdot \epsilon^{\lambda'} = \eta^{\lambda\lambda'} \tag{6.41}$$

which also means that

$$(\epsilon_\mu)^\lambda (\epsilon_\nu)^{\lambda'} \eta_{\lambda\lambda'} = \eta_{\mu\nu} \tag{6.42}$$

The polarization vectors depend on the photon 4-momentum $p = (|\vec{p}|, \vec{p})$. Of the two spacelike polarizations, we will choose ϵ^1 and ϵ^2 to lie transverse to the momentum:

$$\epsilon^1 \cdot p = \epsilon^2 \cdot p = 0 \tag{6.43}$$

The third vector ϵ^3 is the longitudinal polarization. For example, if the momentum lies along the x^3 direction, so $p \sim (1, 0, 0, 1)$, then

$$\epsilon^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \epsilon^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{6.44}$$

For other 4-momenta, the polarization vectors are the appropriate Lorentz transformations of these vectors, since (6.43) are Lorentz invariant.

We do our usual trick, and translate the field commutation relations (6.39) into those for creation and annihilation operators. We find $[a_{\vec{p}}^\lambda, a_{\vec{q}}^{\lambda'}] = [a_{\vec{p}}^{\lambda\,\dagger}, a_{\vec{q}}^{\lambda'\,\dagger}] = 0$ and

$$[a_{\vec{p}}^\lambda, a_{\vec{q}}^{\lambda'\,\dagger}] = -\eta^{\lambda\lambda'} (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) \tag{6.45}$$

The minus signs here are odd to say the least! For spacelike $\lambda = 1, 2, 3$, everything looks fine,

$$[a^{\lambda}_{\vec{p}}, a^{\lambda'\dagger}_{\vec{q}}] = \delta^{\lambda\lambda'}(2\pi)^3\,\delta^{(3)}(\vec{p} - \vec{q}) \quad \lambda, \lambda' = 1, 2, 3 \tag{6.46}$$

But for the timelike annihilation and creation operators, we have

$$[a^{0}_{\vec{p}}, a^{0\dagger}_{\vec{q}}] = -(2\pi)^3\,\delta^{(3)}(\vec{p} - \vec{q}) \tag{6.47}$$

This is very odd! To see just how strange this is, we take the Lorentz invariant vacuum $|0\rangle$ defined by

$$a^{\lambda}_{\vec{p}}\,|0\rangle = 0 \tag{6.48}$$

Then we can create one-particle states in the usual way,

$$|\vec{p}, \lambda\rangle = a^{\lambda\dagger}_{\vec{p}}|0\rangle \tag{6.49}$$

For spacelike polarization states, $\lambda = 1, 2, 3$, all seems well. But for the timelike polarization $\lambda = 0$, the state $|\vec{p}, 0\rangle$ has negative norm,

$$\langle \vec{p}, 0|\vec{q}, 0\rangle = \langle 0|\, a^{0}_{\vec{p}}\, a^{0\dagger}_{\vec{q}}\,|0\rangle = -(2\pi)^3\,\delta^{(3)}(\vec{p} - \vec{q}) \tag{6.50}$$

Wtf? That's very very strange. A Hilbert space with negative norm means negative probabilities which makes no sense at all. We can trace this negative norm back to the wrong sign of the kinetic term for A_0 in our original Lagrangian: $\mathcal{L} = +\frac{1}{2}\dot{\vec{A}}^2 - \frac{1}{2}\dot{A}^2 + \ldots$

At this point we should remember our constraint equation, $\partial_\mu A^\mu = 0$, which, until now, we've not imposed on our theory. This is going to come to our rescue. We will see that it will remove the timelike, negative norm states, and cut the physical polarizations down to two. We work in the Heisenberg picture, so that

$$\partial_\mu A^\mu = 0 \tag{6.51}$$

makes sense as an operator equation. Then we could try implementing the constraint in the quantum theory in a number of different ways. Let's look at a number of increasingly weak ways to do this

- We could ask that $\partial_\mu A^\mu = 0$ is imposed as an equation on operators. But this can't possibly work because the commutation relations (6.39) won't be obeyed for $\pi^0 = -\partial_\mu A^\mu$. We need some weaker condition.

- We could try to impose the condition on the Hilbert space instead of directly on the operators. After all, that's where the trouble lies! We could imagine that there's some way to split the Hilbert space up into good states $|\Psi\rangle$ and bad states that somehow decouple from the system. With luck, our bad states will include the weird negative norm states that we're so disgusted by. But how can we define the good states? One idea is to impose

$$\partial_\mu A^\mu |\Psi\rangle = 0 \tag{6.52}$$

on all good, physical states $|\Psi\rangle$. But this can't work either! Again, the condition is too strong. For example, suppose we decompose $A_\mu(x) = A^+_\mu(x) + A^-_\mu(x)$ with

$$A^+_\mu(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2|\vec{p}|}}\sum_{\lambda=0}^{3} \epsilon^\lambda_\mu a^\lambda_{\vec{p}}\, e^{-ip\cdot x}$$

$$A^-_\mu(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2|\vec{p}|}}\sum_{\lambda=0}^{3} \epsilon^\lambda_\mu a^{\lambda\,\dagger}_{\vec{p}}\, e^{+ip\cdot x} \tag{6.53}$$

Then, on the vacuum $A^+_\mu|0\rangle = 0$ automatically, but $\partial^\mu A^-_\mu|0\rangle \neq 1$. So not even the vacuum is a physical state if we use (6.52) as our constraint

- Our final attempt will be the correct one. In order to keep the vacuum as a good physical state, we can ask that physical states $|\Psi\rangle$ are defined by

$$\partial^\mu A^+_\mu |\Psi\rangle = 0 \tag{6.54}$$

This ensures that

$$\langle \Psi'|\, \partial_\mu A^\mu |\Psi\rangle = 0 \tag{6.55}$$

so that the operator $\partial_\mu A^\mu$ has vanishing matrix elements between physical states. Equation (6.54) is known as the *Gupta-Bleuler* condition. The linearity of the constraint means that the physical states $|\Psi\rangle$ span a physical Hilbert space $\mathcal{H}_{phys}$.

So what does the physical Hilbert space $\mathcal{H}_{phys}$ look like? And, in particular, have we rid ourselves of those nasty negative norm states so that $\mathcal{H}_{phys}$ has a positive definite inner product defined on it? The answer is actually no, but almost!

Let's consider a basis of states for the Fock space. We can decompose any element of this basis as $|\Psi\rangle = |\psi_T\rangle\,|\varphi\rangle$, where $|\psi_T\rangle$ contains only transverse photons, created by

$a^{1,2\,\dagger}_{\vec{p}}$, while $|\varphi\rangle$ contains the timelike photons created by $a^{0\,\dagger}$ and longitudinal photons created by $a^{3\,\dagger}_{\vec{p}}$. The Gupta-Bleuler condition (6.54) requires

$$(a^3_{\vec{p}} - a^0)\,|\varphi\rangle = 0 \tag{6.56}$$

This means that the physical states must contain combinations of timelike and longitudinal photons. Whenever the state contains a timelike photon of momentum $\vec{p}$, it must also contain a longitudinal photon with the same momentum. In general $|\varphi\rangle$ will be a linear combination of states $|\varphi_n\rangle$ containing n pairs of timelike and longitudinal photons, which we can write as

$$|\varphi\rangle = \sum_{n=0}^{\infty} C_n\,|\varphi_n\rangle \tag{6.57}$$

where $|\varphi_0\rangle = |0\rangle$ is simply the vacuum. It's not hard to show that although the condition (6.56) does indeed decouple the negative norm states, all the remaining states involving timelike and longitudinal photons have zero norm

$$\langle\varphi_m|\,\varphi_n\rangle = \delta_{n0}\delta_{m0} \tag{6.58}$$

This means that the inner product on $\mathcal{H}_{phys}$ is positive semi-definite. Which is an improvement. But we still need to deal with all these zero norm states.

The way we cope with the zero norm states is to treat them as gauge equivalent to the vacuum. Two states that differ only in their timelike and longitudinal photon content, $|\varphi_n\rangle$ with $n \geq 1$ are said to be physically equivalent. We can think of the gauge symmetry of the classical theory as descending to the Hilbert space of the quantum theory. Of course, we can't just stipulate that two states are physically identical unless they give the same expectation value for all physical observables. We can check that this is true for the Hamiltonian, which can be easily computed to be

$$H = \int \frac{d^3p}{(2\pi)^3}\,|\vec{p}|\left(\sum_{i=1}^{3} a^{i\,\dagger}_{\vec{p}} a^i_{\vec{p}} - a^{0\,\dagger}_{\vec{p}} a^0_{\vec{p}}\right) \tag{6.59}$$

But the condition (6.56) ensures that $\langle\Psi|\,a^{3\,\dagger}_{\vec{p}} a^3_{\vec{p}}\,|\Psi\rangle = \langle\Psi|\,a^{0\,\dagger}_{\vec{p}} a^0_{\vec{p}}\,|\Psi\rangle$ so that the contri- butions from the timelike and longitudinal photons cancel amongst themselves in the Hamiltonian. This also renders the Hamiltonian positive definite, leaving us just with the contribution from the transverse photons as we would expect.

In general, one can show that the expectation values of all gauge invariant operators evaluated on physical states are independent of the coefficients C_n in (6.57).

Propagators

Finally, it's a simple matter to compute the propagator in Lorentz gauge. It is given by

$$\langle 0|\, T A_\mu(x) A_\nu(y)\, |0\rangle = \int \frac{d^4p}{(2\pi)^4} \frac{-i\eta_{\mu\nu}}{p^2 + i\epsilon} e^{-ip\cdot(x-y)} \tag{6.60}$$

This is a lot nicer than the propagator we found in Coulomb gauge: in particular, it's Lorentz invariant. We could also return to the Lagrangian (6.37). Had we pushed through the calculation with arbitrary coefficient α, we would find the propagator,

$$\langle 0|\, T A_\mu(x) A_\nu(y)\, |0\rangle = \int \frac{d^4p}{(2\pi)^4} \frac{-i}{p^2 + i\epsilon} \left(\eta_{\mu\nu} + (\alpha - 1)\frac{p_\mu p_\nu}{p^2} \right) e^{-ip\cdot(x-y)} \tag{6.61}$$

6.3 Coupling to Matter

Let's now build an interacting theory of light and matter. We want to write down a Lagrangian which couples A_μ to some matter fields, either scalars or spinors. For example, we could write something like

$$\mathcal{L} = -\tfrac{1}{4} F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu \tag{6.62}$$

where j^μ is some function of the matter fields. The equations of motion read

$$\partial_\mu F^{\mu\nu} = j^\nu \tag{6.63}$$

so, for consistency, we require

$$\partial_\mu j^\mu = 0 \tag{6.64}$$

In other words, j^μ must be a conserved current. But we've got lots of those! Let's look at how we can couple two of them to electromagnetism.

6.3.1 Coupling to Fermions

The Dirac Lagrangian

$$\mathcal{L} = \bar\psi (i \partial\!\!\!/ - m)\psi \tag{6.65}$$

has an internal symmetry $\psi \to e^{-i\alpha}\psi$ and $\bar\psi \to e^{+i\alpha}\bar\psi$, with $\alpha \in \mathbf{R}$. This gives rise to the conserved current $j^\mu_V = \bar\psi \gamma^\mu \psi$. So we could look at the theory of electromagnetism coupled to fermions, with the Lagrangian,

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar\psi (i \partial\!\!\!/ - m)\psi - e\bar\psi \gamma^\mu A_\mu \psi \tag{6.66}$$

where we've introduced a coupling constant e. For the free Maxwell theory, we have seen that the existence of a gauge symmetry was crucial in order to cut down the physical degrees of freedom to the requisite 2. Does our interacting theory above still have a gauge symmetry? The answer is yes. To see this, let's rewrite the Lagrangian as

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi \tag{6.67}$$

where $D_\mu\psi = \partial_\mu\psi + ieA_\mu\psi$ is called the *covariant derivative*. This Lagrangian is invariant under gauge transformations which act as

$$A_\mu \to A_\mu + \partial_\mu\lambda \quad \text{and} \quad \psi \to e^{-ie\lambda}\psi \tag{6.68}$$

for an arbitrary function $\lambda(x)$. The tricky term is the derivative acting on ψ, since this will also hit the $e^{-ie\lambda}$ piece after the transformation. To see that all is well, let's look at how the covariant derivative transforms. We have

$$\begin{aligned} D_\mu\psi &= \partial_\mu\psi + ieA_\mu\psi \\ &\to \partial_\mu(e^{-ie\lambda}\psi) + ie(A_\mu + \partial_\mu\lambda)(e^{-ie\lambda}\psi) \\ &= e^{-ie\lambda}D_\mu\psi \end{aligned} \tag{6.69}$$

so the covariant derivative has the nice property that it merely picks up a phase under the gauge transformation, with the derivative of $e^{-ie\lambda}$ cancelling the transformation of the gauge field. This ensures that the whole Lagrangian is invariant, since $\bar{\psi} \to e^{+ie\lambda(x)}\bar{\psi}$.

Electric Charge

The coupling e has the interpretation of the electric charge of the ψ particle. This follows from the equations of motion of classical electromagnetism $\partial_\mu F^{\mu\nu} = j^\nu$: we know that the j^0 component is the charge density. We therefore have the total charge Q given by

$$Q = e\int d^3x\, \bar{\psi}(\vec{x})\gamma^0\psi(\vec{x}) \tag{6.70}$$

After treating this as a quantum equation, we have

$$Q = e\int \frac{d^3p}{(2\pi)^3}\sum_{s=1}^{2}(b_{\vec{p}}^{s\dagger}b_{\vec{p}}^{s} - c_{\vec{p}}^{s\dagger}c_{\vec{p}}^{s}) \tag{6.71}$$

which is the number of particles, minus the number of antiparticles. Note that the particle and the anti-particle are required by the formalism to have opposite electric charge. For QED, the theory of light interacting with electrons, the electric charge is usually written in terms of the dimensionless ratio α, known as the fine structure constant

$$\alpha = \frac{e^2}{4\pi k c} \approx \frac{1}{137} \tag{6.72}$$

Setting $k = c = 1$, we have $e = \sqrt{4\pi\alpha} \approx 0.3$.

There's a small subtlety here that's worth elaborating on. I stressed that there's a radical difference between the interpretation of a global symmetry and a gauge symmetry. The former takes you from one physical state to another with the same properties and results in a conserved current through Noether's theorem. The latter is a redundancy in our description of the system. Yet in electromagnetism, the gauge symmetry $\psi \to e^{+ie\lambda(x)}\psi$ seems to lead to a conservation law, namely the conservation of electric charge. This is because among the infinite number of gauge symmetries parameterized by a function $\lambda(x)$, there is also a single global symmetry: that with $\lambda(x)$ = constant. This is a true symmetry of the system, meaning that it takes us to another physical state. More generally, the subset of global symmetries from among the gauge symme- tries are those for which $\lambda(x) \to \alpha$ = constant as $x \to \infty$. These take us from one physical state to another.

Finally, let's check that the 4 × 4 matrix C that we introduced in Section 4.5 really deserves the name "charge conjugation matrix". If we take the complex conjugation of the Dirac equation, we have

$$(i\gamma^\mu\partial_\mu - e\gamma^\mu A_\mu - m)\psi = 0 \quad \Rightarrow \quad (-i(\gamma^\mu)^{\wedge}\partial_\mu - e(\gamma^\mu)^{\wedge}A_\mu - m)\psi^{\wedge} = 0$$

Now using the defining equation $C^\dagger\gamma^\mu C = -(\gamma^\mu)^{\wedge}$, and the definition $\psi^{(c)} = C\psi^{\wedge}$, we see that the charge conjugate spinor $\psi^{(c)}$ satisfies

$$(i\gamma^\mu\partial_\mu + e\gamma^\mu A_\mu - m)\psi^{(c)} = 0 \tag{6.73}$$

So we see that the charge conjugate spinor $\psi^{(c)}$ satisfies the Dirac equation, but with charge $-e$ instead of $+e$.

6.3.2 Coupling to Scalars

For a real scalar field, we have no suitable conserved current. This means that we can't couple a real scalar field to a gauge field.

Let's now consider a complex scalar field ϕ. (For this section, I'll depart from our previous notation and call the scalar field ϕ to avoid confusing it with the spinor). We have a symmetry $\phi \to e^{-i\alpha}\phi$. We could try to couple the associated current to the gauge field,

$$\mathcal{L}_{int} = -i((\partial_\mu\phi^\dagger)\phi - \phi^\dagger\partial_\mu\phi)A^\mu \tag{6.74}$$

But this doesn't work because

- The theory is no longer gauge invariant
- The current j^μ that we coupled to A_μ depends on $\partial_\mu\phi$. This means that if we try to compute the current associated to the symmetry, it will now pick up a contribution from the $j^\mu A_\mu$ term. So the whole procedure wasn't consistent.

We solve both of these problems simultaneously by remembering the covariant derivative. In this scalar theory, the combination

$$D_\mu\phi = \partial_\mu\phi + ieA_\mu\phi \tag{6.75}$$

again transforms as $D_\mu\phi \to e^{-ie\lambda}D_\mu\phi$ under a gauge transformation $A_\mu \to A_\mu + \partial_\mu\lambda$ and $\phi \to e^{-ie\lambda}\phi$. This means that we can construct a gauge invariant action for a charged scalar field coupled to a photon simply by promoting all derivatives to covariant derivatives

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + D_\mu\phi^\dagger D^\mu\phi - m^2|\phi|^2 \tag{6.76}$$

In general, this trick works for any theory. If we have a $U(1)$ symmetry that we wish to couple to a gauge field, we may do so by replacing all derivatives by suitable covariant derivatives. This procedure is known as *minimal coupling*.

6.4 QED

Let's now work out the Feynman rules for the full theory of quantum electrodynamics (QED) – the theory of electrons interacting with light. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi \tag{6.77}$$

where $D_\mu = \partial_\mu + ieA_\mu$.

The route we take now depends on the gauge choice. If we worked in Lorentz gauge previously, then we can jump straight to Section 6.5 where the Feynman rules for QED are written down. If, however, we worked in Coulomb gauge, then we still have a bit of work in front of us in order to massage the photon propagator into something Lorentz invariant. We will now do that.

In Coulomb gauge $\nabla\cdot\vec{A} = 0$, the equation of motion arising from varying A_0 is now

$$-\nabla^2 A_0 = e\psi^\dagger\psi \equiv ej_0 \tag{6.78}$$

which has the solution

$$A_0(\vec{x},t) = e\int d^3x' \, \frac{j^0(\vec{x}',t)}{4\pi|\vec{x}-\vec{x}'|} \tag{6.79}$$

In Coulomb gauge we can rewrite the Maxwell part of the Lagrangian as

$$\begin{aligned} L_{\text{Maxwell}} &= \int d^3x \; \tfrac{1}{2}\vec{E}^{\,2} - \tfrac{1}{2}\vec{B}^{\,2} \\ &= \int d^3x \; \tfrac{1}{2}(\dot{\vec{A}} + \nabla A_0)^2 - \tfrac{1}{2}\vec{B}^{\,2} \\ &= \int d^3x \; \tfrac{1}{2}\dot{\vec{A}}^{\,2} + \tfrac{1}{2}(\nabla A_0)^2 - \tfrac{1}{2}\vec{B}^{\,2} \end{aligned} \tag{6.80}$$

where the cross-term has vanished using $\nabla\cdot\vec{A} = 0$. After integrating the second term by parts and inserting the equation for A_0, we have

$$L_{\text{Maxwell}} = \int d^3x \; \tfrac{1}{2}\dot{\vec{A}}^{\,2} - \tfrac{1}{2}\vec{B}^{\,2} + \frac{e^2}{2}\int d^3x' \, \frac{j_0(\vec{x})j_0(\vec{x}')}{4\pi|\vec{x}-\vec{x}'|} \tag{6.81}$$

We find ourselves with a nonlocal term in the action. This is exactly the type of interaction that we boasted in Section 1.1.4 never arises in Nature! It appears here as an artifact of working in Coulomb gauge: it does not mean that the theory of QED is nonlocal. For example, it wouldn't appear if we worked in Lorentz gauge.

We now compute the Hamiltonian. Changing notation slightly from previous chapters, we have the conjugate momenta,

$$\begin{aligned} \vec{\Pi} &= \frac{\partial\mathcal{L}}{\partial\dot{\vec{A}}} = \dot{\vec{A}} \\ \pi_\psi &= \frac{\partial\mathcal{L}}{\partial\dot{\psi}} = i\psi^\dagger \end{aligned} \tag{6.82}$$

which gives us the Hamiltonian

$$H = \int d^3x \; \tfrac{1}{2}\dot{\vec{A}}^{\,2} + \tfrac{1}{2}\vec{B}^{\,2} + \bar{\psi}(-i\gamma^i\partial_i + m)\psi - e\vec{j}\cdot\vec{A} + \frac{e^2}{2}\int d^3x' \, \frac{j^0(\vec{x})j^0(\vec{x}')}{4\pi|\vec{x}-\vec{x}'|}$$

where $\vec{j} = \bar{\psi}\vec{\gamma}\psi$ and $j^0 =$

6.4.1 Naive Feynman Rules

We want to determine the Feynman rules for this theory. For fermions, the rules are the same as those given in Section 5. The new pieces are:

- We denote the photon by a wavy line. Each end of the line comes with an $i, j = 1, 2, 3$ index telling us the component of $\vec{A}$. We calculated the transverse photon propagator in (6.33): it is  and contributes $D^{tr}_{ij} = \frac{i}{p^2 + i\epsilon}\left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2}\right)$

- The vertex  contributes $-ie\gamma^i$. The index on γ^i contracts with the index on the photon line.

- The non-local interaction which, in position space, is given by  contributes a factor of $\frac{i(e\gamma^0)^2\delta(x^0 - y^0)}{4\pi|\vec{x} - \vec{y}|}$

These Feynman rules are rather messy. This is the price we've paid for working in Coulomb gauge. We'll now show that we can massage these expressions into something much more simple and Lorentz invariant. Let's start with the offending instantaneous interaction. Since it comes from the A_0 component of the gauge field, we could try to redefine the propagator to include a D_{00} piece which will capture this term. In fact, it fits quite nicely in this form: if we look in momentum space, we have

$$\frac{\delta(x^0 - y^0)}{4\pi|\vec{x} - \vec{y}|} = \int \frac{d^4p}{(2\pi)^4} \frac{e^{ip\cdot(x-y)}}{|\vec{p}|^2} \tag{6.83}$$

so we can combine the non-local interaction with the transverse photon propagator by defining a new photon propagator

$$D_{\mu\nu}(p) = \begin{cases} +\frac{i}{|\vec{p}|^2} & \mu, \nu = 0 \\ \frac{i}{p^2 + i\epsilon}\left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2}\right) & \mu = i \neq 0, \ \nu = j \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.84}$$

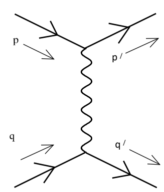
With this propagator, the wavy photon line now carries a $\mu, \nu = 0, 1, 2, 3$ index, with the extra $\mu = 0$ component taking care of the instantaneous interaction. We now need to change our vertex slightly: the $-ie\gamma^i$ above gets replaced by $-ie\gamma^\mu$ which correctly accounts for the $(e\gamma^0)^2$ piece in the instantaneous interaction.

The D_{00} piece of the propagator doesn't look a whole lot different from the transverse photon propagator. But wouldn't it be nice if they were both part of something more symmetric! In fact, they are. We have the following:

Claim: We can replace the propagator $D_{\mu\nu}(p)$ with the simpler, Lorentz invariant propagator

$$D_{\mu\nu}(p) = -\frac{i\eta_{\mu\nu}}{p^2} \tag{6.85}$$

Proof: There is a general proof using current conservation. Here we'll be more pedestrian and show that we can do this for certain Feynman diagrams. In particular, we focus on a particular tree-level diagram that contributes to $e^-e^- \to e^-e^-$ scattering,



$$\sim e^2[\bar{u}(p')\gamma^\mu u(p)]\, D_{\mu\nu}(k)\, [\bar{u}(q')\gamma^\nu u(q)] \tag{6.86}$$

where $k = p - p' = q' - q$. Recall that $u(\vec{p})$ satisfies the equation

$$(\not{p} - m)u(\vec{p}) = 0 \tag{6.87}$$

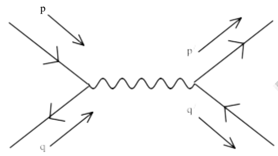
Let's define the spinor contractions $\alpha^\mu = \bar{u}(\vec{p}\,')\gamma^\mu u(\vec{p})$ and $\beta^\nu = \bar{u}(\vec{q}\,')\gamma^\nu u(\vec{q})$. Then since $k = p - p' = q' - q$, we have

$$k_\mu\alpha^\mu = \bar{u}(\vec{p}\,')(\not{p}' - \not{p})u(\vec{p}) = \bar{u}(\vec{p}\,')(m - m)u(\vec{p}) = 0 \tag{6.88}$$

and, similarly, $k_\nu\beta^\nu = 0$. Using this fact, the diagram can be written as

$$\begin{aligned}
\alpha^\mu D_{\mu\nu}\beta^\nu &= i\left(\frac{\vec{\alpha}\cdot\vec{\beta}}{k^2} - \frac{(\vec{\alpha}\cdot\vec{k})(\vec{\beta}\cdot\vec{k})}{k^2|\vec{k}|^2} + \frac{\alpha^0\beta^0}{|\vec{k}|^2}\right) \\
&= i\left(\frac{\vec{\alpha}\cdot\vec{\beta}}{k^2} - \frac{k_0^2\alpha_0\beta_0}{k^2|\vec{k}|^2} + \frac{\alpha_0\beta_0}{|\vec{k}|^2}\right) \\
&= i\left(\frac{\vec{\alpha}\cdot\vec{\beta}}{k^2} - \frac{1}{k^2|\vec{k}|^2}(k_0^2 - k^2)\alpha_0\beta_0\right) \\
&= -\frac{i}{k^2}\alpha\cdot\beta = \alpha^\mu\left(-\frac{i\eta_{\mu\nu}}{k^2}\right)\beta^\nu
\end{aligned} \tag{6.89}$$

which is the claimed result. You can similarly check that the same substitution is legal in the diagram



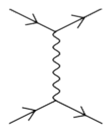
$$\sim e^2[\bar{v}(\vec{q})\gamma^\mu u(\vec{p})]D_{\mu\nu}(k)[\bar{u}(\vec{p}\,')\gamma^\nu v(\vec{q}\,')] \tag{6.90}$$

In fact, although we won't show it here, it's a general fact that in every Feynman diagram we may use the very nice, Lorentz invariant propagator $D_{\mu\nu} = -i\eta_{\mu\nu}/p^2$.

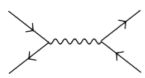
Note: This is the propagator we found when quantizing in Lorentz gauge (using the Feynman gauge parameter). In general, quantizing the Lagrangian (6.37) in Lorentz gauge, we have the propagator

$$D_{\mu\nu} = -\frac{i}{p^2}\left(\eta_{\mu\nu} + (\alpha - 1)\frac{p_\mu p_\nu}{p^2}\right) \tag{6.91}$$

Using similar arguments to those given above, you can show that the $p_\mu p_\nu/p^2$ term cancels in all diagrams. For example, in the following diagrams the $p_\mu p_\nu$ piece of the propagator contributes as



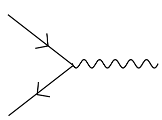
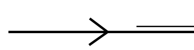
$$\sim \bar{u}(p')\gamma^\mu u(p)\, k_\mu = \bar{u}(p')(\not{p} - \not{p}')u(p) = 0$$



$$\sim \bar{v}(p)\gamma^\mu u(q)\, k_\mu = \bar{v}(p)(\not{p} + \not{q})u(q) = 0 \tag{6.92}$$

6.5 Feynman Rules

Finally, we have the Feynman rules for QED. For vertices and internal lines, we write

- Vertex:  $-ie\gamma^\mu$
- Photon Propagator: $-\dfrac{i\eta_{\mu\nu}}{p^2 + i\epsilon}$
- Fermion Propagator:  $\dfrac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}$

For external lines in the diagram, we attach

- Photons: We add a polarization vector $\epsilon^\mu_{in}/\epsilon^\mu_{out}$ for incoming/outgoing photons. In Coulomb gauge, $\epsilon^0 = 0$ and $\vec{\epsilon}\cdot\vec{p} = 0$.
- Fermions: We add a spinor $u^r(\vec{p})/\bar{u}^r(\vec{p})$ for incoming/outgoing fermions. We add a spinor $\bar{v}^r(\vec{p})/v^r(\vec{p})$ for incoming/outgoing anti-fermions.

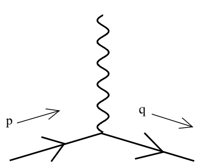
6.5.1 Charged Scalars

> "Pauli asked me to calculate the cross section for pair creation of scalar particles by photons. It was only a short time after Bethe and Heitler had solved the same problem for electrons and positrons. I met Bethe in Copenhagen at a conference and asked him to tell me how he did the calculations. I also inquired how long it would take to perform this task; he answered, "It would take me three days, but you will need about three weeks." He was right, as usual; furthermore, the published cross sections were wrong by a factor of four."
>
> *Viki Weisskopf*

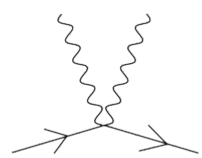
The interaction terms in the Lagrangian for charged scalars come from the covariant derivative terms,

$$\mathcal{L} = D_\mu\psi^\dagger D^\mu\psi = \partial_\mu\psi\,\partial^{\dagger\mu}\psi - ieA_\mu(\psi^\dagger\partial^\mu\psi - \psi\partial^\mu\psi^\dagger) + e^2 A_\mu A^\mu\psi^\dagger\psi \tag{6.93}$$

This gives rise to two interaction vertices. But the cubic vertex is something we haven't seen before: it contains kinetic terms. How do these appear in the Feynman rules? After a Fourier transform, the derivative term means that the interaction is stronger for fermions with higher momentum, so we include a momentum factor in the Feynman rule. There is also a second, "seagull" graph. The two Feynman rules are



$-ie(p+q)_\mu$ and



$+2ie^2\eta_{\mu\nu}$

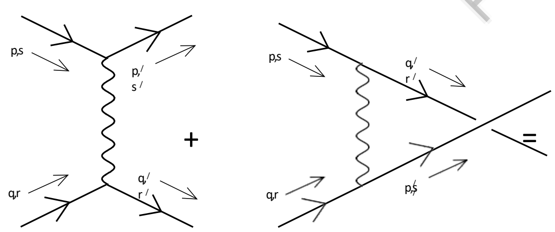
The factor of two in the seagull diagram arises because of the two identical particles appearing in the vertex. (It's the same reason that the 1/4! didn't appear in the Feynman rules for φ^4 theory).

6.6 Scattering in QED

Let's now calculate some amplitudes for various processes in quantum electrodynamics, with a photon coupled to a single fermion. We will consider the analogous set of processes that we saw in Section 3 and Section 5. We have

Electron Scattering

Electron scattering $e^-e^- \to e^-e^-$ is described by the two leading order Feynman diagrams, given by

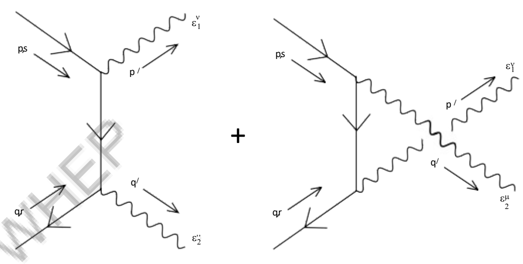


$$= -i(-ie)^2\left(\frac{[\bar u^{s'}(\vec p\,')\gamma^\mu u^s(\vec p)]\,[\bar u^{r'}(\vec q\,')\gamma_\mu u^r(\vec q)]}{(p'-p)^2} - \frac{[\bar u^{s'}(\vec p\,')\gamma^\mu u^r(\vec q)]\,[\bar u^{r'}(\vec q\,')\gamma_\mu u^s(\vec p)]}{(p-q')^2}\right)$$

The overall $-i$ comes from the $-i\eta_{\mu\nu}$ in the propagator, which contract the indices on the γ-matrices (remember that it's really positive for $\mu, \nu = 1, 2, 3$).

Electron Positron Annihilation

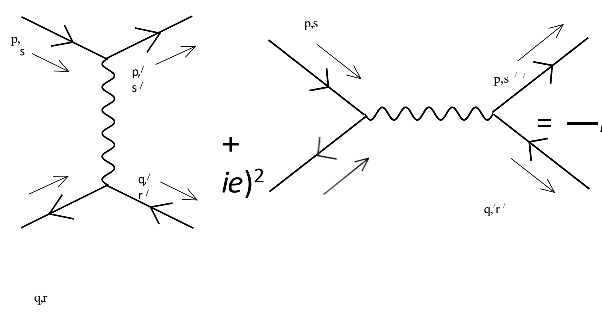
Let's now look at $e^-e^+ \to 2\gamma$, two gamma rays. The two lowest order Feynman diagrams are,



$$= i(-ie)^2\,\bar v^r(\vec q)\left(\frac{\gamma_\mu(\not p-\not p'+m)\gamma_\nu}{(p-p')^2-m^2} + \frac{\gamma_\nu(\not p-\not q'+m)\gamma_\mu}{(p-q')^2-m^2}\right)u^s(\vec p)\,\epsilon_1^\nu(\vec p\,')\,\epsilon_2^\mu(\vec q\,')$$

Electron Positron Scattering

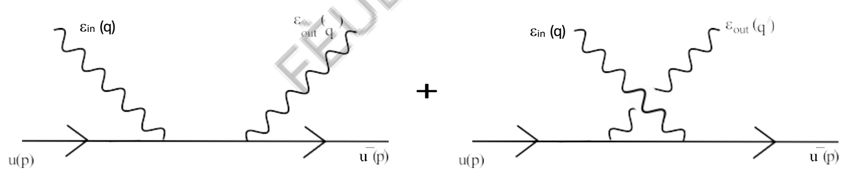
For $e^-e^+ \to e^-e^+$ scattering (sometimes known as Bhabha scattering) the two lowest order Feynman diagrams are



$$= -i(-ie)^2\left(-\frac{[\bar u^{s'}(\vec p\,')\gamma^\mu u^s(\vec p)]\,[\bar v^r(\vec q)\gamma_\mu v^{r'}(\vec q\,')]}{(p-p')^2} + \frac{[\bar v^r(\vec q)\gamma^\mu u^s(\vec p)]\,[\bar u^{s'}(\vec p\,')\gamma_\mu v^{r'}(\vec q\,')]}{(p+q)^2}\right)$$

Compton Scattering

The scattering of photons (in particular x-rays) off electrons $e^-\gamma \to e^-\gamma$ is known as Compton scattering. Historically, the change in wavelength of the photon in the scattering process was one of the conclusive pieces of evidence that light could behave as a particle. The amplitude is given by,



$$= i(-ie)^2\, \bar{u}^{r'}(\vec{p}\,')\left[\frac{\gamma_\mu(\not p+\not q+m)\gamma_\nu}{(p+q)^2-m^2}+\frac{\gamma_\nu(\not p-\not q'+m)\gamma_\mu}{(p-q')^2-m^2}\right]u^s(\vec{p})\,\epsilon^\nu_{in}\,\epsilon^\mu_{out}$$

This amplitude vanishes for longitudinal photons. For example, suppose $\epsilon_{in} \sim q$. Then, using momentum conservation $p + q = p' + q'$, we may write the amplitude as

$$i\mathcal{A} = i(-ie)^2\, \bar{u}^{r'}(\vec{p}\,')\left[\not\epsilon_{out}\frac{(\not p+\not q+m)\not q}{(p+q)^2-m^2}+\frac{\not q(\not p'-\not q+m)\not\epsilon_{out}}{(p'-q)^2-m^2}\right]u^s(\vec{p})$$

$$= i(-ie)^2\, \bar{u}^{r'}(\vec{p}\,')\,\not\epsilon_{out}\,u^s(\vec{p})\left[\frac{2p\cdot q}{(p+q)^2-m^2}+\frac{2p'\cdot q}{(p'-q)^2-m^2}\right] \qquad (6.94)$$

where, in going to the second line, we've performed some γ-matrix manipulations, together with the spinor equations $(\not p - m)u(\vec{p})$ and $\bar{u}(\vec{p}\,')(\not p' - m) = 0$. We now recall the fact that q is a null vector, while $p^2 = (p')^2 = m^2$ since the external legs are on mass- shell. This means that the two denominators in the amplitude read $(p+q)^2 - m^2 = 2p\cdot q$ and $(p' - q)^2 - m^2 = -2p' \cdot q$. This ensures that the combined amplitude vanishes for longitudinal photons as promised. A similar result holds when $\not\epsilon_{out} \sim q'$.

Photon Scattering

In QED, photons no longer pass through each other unimpeded. At one-loop, there is a diagram which leads to photon scattering. Although naively logarithmically divergent, the diagram is actually rendered finite by gauge invariance.

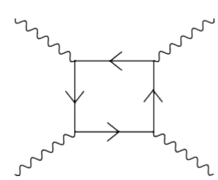
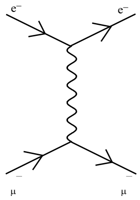


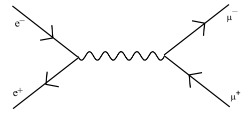
Figure 30:

Adding Muons

Adding a second fermion into the mix, which we could identify as a muon, new processes become possible. For example, we can now have processes such as $e^-\mu^- \to e^-\mu^-$ scattering, and e^+e^- annihilation into a muon anti-muon pair. Using our standard notation of p and q for incoming momenta, and p' and q' for outgoing momenta, we have the amplitudes given by



$$\sim \frac{1}{(p - p')^2} \quad \text{and}$$

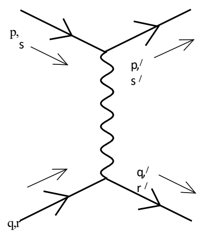


$$\sim \frac{1}{(p + q)^2} \tag{6.95}$$

6.6.1 The Coulomb Potential

We've come a long way. We've understood how to compute quantum amplitudes in a large array of field theories. To end this course, we use our newfound knowledge to rederive a result you learnt in kindergarten: Coulomb's law.

To do this, we repeat our calculation that led us to the Yukawa force in Sections 3.5.2 and 5.7.2. We start by looking at $e^-e^- \to e^-e^-$ scattering. We have



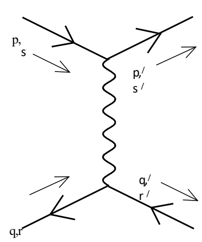
$$= -i(-ie)^2 \frac{[\bar{u}(\vec{p}\,')\gamma^\mu u(\vec{p})][\bar{u}(\vec{q}\,')\gamma_\mu u(\vec{q})]}{(p' - p)^2} \tag{6.96}$$

Following (5.49), the non-relativistic limit of the spinor is $u(p) \to \sqrt{m}\begin{pmatrix} \xi \\ \xi \end{pmatrix}$. This means that the γ^0 piece of the interaction gives a term $\bar{u}^s(\vec{p})\gamma^0 u^r(\vec{q}) \to 2m\delta^{rs}$, while the spatial γ^i, $i = 1, 2, 3$ pieces vanish in the non-relativistic limit: $\bar{u}^s(\vec{p})\gamma^i u^r(\vec{q}) \to 0$. Comparing the scattering amplitude in this limit to that of non-relativistic quantum mechanics, we have the effective potential between two electrons given by,

$$U(\vec{r}) = +e^2 \int \frac{d^3p}{(2\pi)^3} \frac{e^{i\vec{p}\cdot\vec{r}}}{|\vec{p}|^2} = + \frac{e^2}{4\pi r} \tag{6.97}$$

We find the familiar repulsive Coulomb potential. We can trace the minus sign that gives a repulsive potential to the fact that only the A_0 component of the intermediate propagator $\sim -i\eta_{\mu\nu}$ contributes in the non-relativistic limit.

For $e^-e^+ \to e^-e^+$ scattering, the amplitude is



$$= +i(-ie)^2 \frac{[\bar{u}(\vec{p}\,')\gamma^\mu u(\vec{p})][\bar{v}(\vec{q})\gamma_\mu v(\vec{q}\,')]}{(p' - p)^2} \tag{6.98}$$

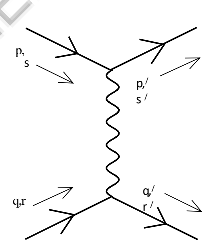
The overall + sign comes from treating the fermions correctly: we saw the same minus sign when studying scattering in Yukawa theory. The difference now comes from looking at the non-relativistic limit. We have $\bar{v}\gamma^0 v \to 2m$, giving us the potential between opposite charges,

$$U(\vec{r}) = -e^2 \int \frac{d^3p}{(2\pi)^3} \frac{e^{i\vec{p}\cdot\vec{r}}}{|\vec{p}|^2} = -\frac{e^2}{4\pi r} \tag{6.99}$$

Reassuringly, we find an attractive force between an electron and positron. The difference from the calculation of the Yukawa force comes again from the zeroth component of the gauge field, this time in the guise of the γ^0 sandwiched between $\bar{v}\gamma^0 v \to 2m$, rather than the $\bar{v}v \to -2m$ that we saw in the Yukawa case.

The Coulomb Potential for Scalars

There are many minus signs in the above calculation which somewhat obscure the crucial one which gives rise to the repulsive force. A careful study reveals the offending sign to be that which sits in front of the A_0 piece of the photon propagator $-i\eta_{\mu\nu}/p^2$. Note that with our signature $(+---)$, the propagating A_i have the correct sign, while A_0 comes with the wrong sign. This is simpler to see in the case of scalar QED, where we don't have to worry about the gamma matrices. From the Feynman rules of Section 6.5.1, we have the non-relativistic limit of scalar e^-e^- scattering,



$$= -i\eta_{\mu\nu}(-ie)^2 \frac{(p+p')^\mu (q+q')^\nu}{(p'-p)^2} \to -i(-ie)^2 \frac{(2m)^2}{-(\vec{p}-\vec{p}\,')^2}$$

where the non-relativistic limit in the numerator involves $(p+p')\cdot(q+q') \approx (p+p')^0(q+q')_0 \approx (2m)^2$ and is responsible for selecting the A_0 part of the photon propagator rather than the A_i piece. This shows that the Coulomb potential for spin 0 particles of the same charge is again repulsive, just as it is for fermions. For e^-e^+ scattering, the amplitude picks up an extra minus sign because the arrows on the legs of the Feynman rules in Section 6.5.1 are correlated with the momentum arrows. Flipping the arrows on one pair of legs in the amplitude introduces the relevant minus sign to ensure that the non-relativistic potential between e^-e^+ is attractive as expected.

6.7 Afterword

In this course, we have laid the foundational framework for quantum field theory. Most of the developments that we've seen were already in place by the middle of the 1930s, pioneered by people such as Jordan, Dirac, Heisenberg, Pauli and Weisskopf [5].

Yet by the end of the 1930s, physicists were ready to give up on quantum field theory. The difficulty lies in the next terms in perturbation theory. These are the terms that correspond to Feynamn diagrams with loops in them, which we have scrupulously avoided computing in this course. The reason we've avoided them is because they typically give infinity! And, after ten years of trying, and failing, to make sense of this, the general feeling was that one should do something else. This from Dirac in 1937,

> *Because of its extreme complexity, most physicists will be glad to see the end of QED*

But the leading figures of the day gave up too soon. It took a new generation of postwar physicists — people like Schwinger, Feynman, Tomonaga and Dyson — to return to quantum field theory and tame the infinities. The story of how they did that will be told in next term's course.

[5] For more details on the history of quantum field theory, see the excellent book "QED and the Men who Made it" by Sam Schweber.

Chapter 2

Basic Set Theory

> **A set is a Many that allows itself to be thought of as a One.**
> - *Georg Cantor*

This chapter introduces set theory, mathematical induction, and formalizes the notion of mathematical functions. The material is mostly elementary. For those of you new to abstract mathematics elementary does not mean *simple* (though much of the material is fairly simple). Rather, elementary means that the material requires very little previous education to understand it. Elementary material can be quite challenging and some of the material in this chapter, if not exactly rocket science, may require that you adjust you point of view to understand it. The single most powerful technique in mathematics is to adjust your point of view until the problem you are trying to solve becomes simple.

Another point at which this material may diverge from your previous experience is that it will require proof. In standard introductory classes in algebra, trigonometry, and calculus there is currently very little emphasis on the discipline of *proof.* Proof is, however, the central tool of mathematics. This text is for a course that is a students formal introduction to tools and methods of proof.

2.1 Set Theory

A *set* is a collection of distinct objects. This means that $\{1, 2, 3\}$ is a set but $\{1, 1, 3\}$ is not because 1 appears twice in the second collection. The second collection is called a *multiset.* Sets are often specified with curly brace notation. The set of even integers can be written:

$$\{2n : \text{n is an integer}\}$$

The opening and closing curly braces denote a set, $2n$ specifies the members of the set, the colon says "such that" or "where" and everything following the colon are conditions that explain or refine the membership. All correct mathematics can be spoken in English. The set definition above is spoken "The set of twice n where n is an integer".

The only problem with this definition is that we do not yet have a formal definition of the integers. The integers are the set of whole numbers, both positive and negative: $\{0, \pm 1, \pm 2, \pm 3, \ldots\}$. We now introduce the operations used to manipulate sets, using the opportunity to practice curly brace notation.

Definition 2.1 *The* **empty set** *is a set containing no objects. It is written as a pair of curly braces with nothing inside* $\{\}$ *or by using the symbol* $\emptyset$*.*

As we shall see, the empty set is a handy object. It is also quite strange. The set of all humans that weigh at least eight tons, for example, is the empty set. Sets whose definition contains a contradiction or impossibility are often empty.

Definition 2.2 *The* **set membership symbol** $\in$ *is used to say that an object is a member of a set. It has a partner symbol* $\notin$ *which is used to say an object is not in a set.*

Definition 2.3 *We say two sets are* **equal** *if they have exactly the same members.*

Example 2.1 *If*

$$S = \{1, 2, 3\}$$

then $3 \in S$ *and* $4 \notin S$. *The set membership symbol is often used in defining operations that manipulate sets. The set*

$$T = \{2, 3, 1\}$$

is equal to S *because they have the same members: 1, 2, and 3. While we usually list the members of a set in a "standard" order (if one is available) there is no requirement to do so and sets are indifferent to the order in which their members are listed.*

Definition 2.4 *The* **cardinality** *of a set is its* size. *For a finite set, the cardinality of a set is the number of members it contains. In symbolic notation the size of a set* S *is written* $|S|$. *We will deal with the idea of the cardinality of an infinite set later.*

Example 2.2 Set cardinality
For the set $S = \{1, 2, 3\}$ *we show cardinality by writing* $|S| = 3$

We now move on to a number of *operations* on sets. You are already familiar with several operations on numbers such as addition, multiplication, and negation.

Definition 2.5 *The* **intersection** *of two sets* S *and* T *is the collection of all objects that are in both sets. It is written* $S \cap T$. *Using curly brace notation*

$$S \cap T = \{x : (x \in S) \text{ and } (x \in T)\}$$

The symbol *and* in the above definition is an example of a Boolean or logical operation. It is only true when both the propositions it joins are also true. It has a symbolic equivalent $\wedge$. This lets us write the formal definition of intersection more compactly:

$$S \cap T = \{x : (x \in S) \wedge (x \in T)\}$$

Example 2.3 Intersections of sets
Suppose $S = \{1, 2, 3, 5\}$,
$T = \{1, 3, 4, 5\}$, *and* $U = \{2, 3, 4, 5\}$.
Then:

$S \cap T = \{1, 3, 5\}$,

$S \cap U = \{2, 3, 5\}$, *and*

$T \cap U = \{3, 4, 5\}$

Definition 2.6 *If* A *and* B *are sets and* $A \cap B = \emptyset$ *then we say that* A *and* B *are* disjoint, *or* disjoint sets.

Definition 2.7 *The* **union** *of two sets* S *and* T *is the collection of all objects that are in either set. It is written* $S \cup T$. *Using curly brace notion*

$$S \cup T = \{x : (x \in S) \text{ or } (x \in T)\}$$

The symbol *or* is another Boolean operation, one that is true if either of the propositions it joins are true. Its symbolic equivalent is $\vee$ which lets us re-write the definition of union as:

$$S \cup T = \{x : (x \in S) \vee (x \in T)\}$$

Example 2.4 Unions of sets.
Suppose $S = \{1, 2, 3\}$, $T = \{1, 3, 5\}$, *and* $U = \{2, 3, 4, 5\}$.
Then:

$S \cup T = \{1, 2, 3, 5\}$,

$S \cup U = \{1, 2, 3, 4, 5\}$, *and*

$T \cup U = \{1, 2, 3, 4, 5\}$

When performing set theoretic computations, you should declare the domain in which you are working. In set theory this is done by declaring a universal set.

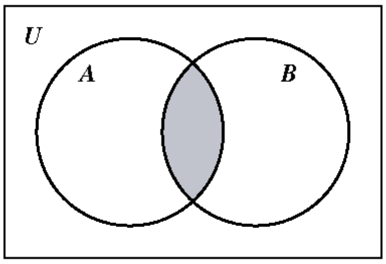
Definition 2.8 *The* **universal set**, *at least for a given collection of set theoretic computations, is the set of all possible objects.*

If we declare our universal set to be the integers then $\{\frac{1}{2}, \frac{2}{3}\}$ is not a well defined set because the objects used to define it are not members of the universal set. The symbols $\{\frac{1}{2}, \frac{2}{3}\}$ do define a set if a universal set that includes $\frac{1}{2}$ and $\frac{2}{3}$ is chosen. The problem arises from the fact that neither of these numbers are integers. The universal set is commonly written $\mathcal{U}$. Now that we have the idea of declaring a universal set we can define another operation on sets.

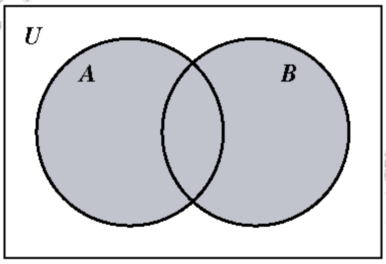
2.1.1 Venn Diagrams

A Venn diagram is a way of depicting the relationship between sets. Each set is shown as a circle and circles overlap if the sets intersect.

Example 2.5 *The following are Venn diagrams for the intersection and union of two sets. The shaded parts of the diagrams are the intersections and unions respectively.*



$A \cap B$



$A \cup B$

Notice that the rectangle containing the diagram is labeled with a U representing the universal set.

Definition 2.9 *The* **compliment** *of a set S is the collection of objects in the universal set that are not in S. The compliment is written S^c. In curly brace notation*

$$S^c = \{x : (x \in \mathcal{U}) \wedge (x \notin S)\}$$

or more compactly as

$$S^c = \{x : x \notin S\}$$

however it should be apparent that the compliment of a set always depends on which universal set is chosen.

There is also a Boolean symbol associated with the complementation operation: the *not* operation. The notation for not is $\neg$. There is not much savings in space as the definition of compliment becomes

$$S^c = \{x : \neg(x \in S)\}$$

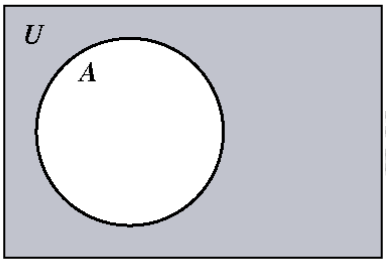
Example 2.6 Set Compliments

(i) *Let the universal set be the integers. Then the compliment of the even integers is the odd integers.*

(ii) *Let the universal set be $\{1, 2, 3, 4, 5\}$, then the compliment of $S = \{1, 2, 3\}$ is $S^c = \{4, 5\}$ while the compliment of $T = \{1, 3, 5\}$ is $T^c = \{2, 4\}$.*

(iii) *Let the universal set be the letters $\{a, e, i, o, u, y\}$. Then $\{y\}^c = \{a, e, i, o, u\}$.*

The Venn diagram for A^c is



A^c

We now have enough set-theory operators to use them to define more operators quickly. We will continue to give English and symbolic definitions.

Definition 2.10 *The* **difference** *of two sets S and T is the collection of objects in S that are not in T. The difference is written $S - T$. In curly brace notation*

$$S - T = \{x : x \in (S \cap (T^c))\},$$

or alternately

$$S - T = \{x : (x \in S) \wedge (x \notin T)\}$$

Notice how intersection and complementation can be used together to create the difference operation and that the definition can be rephrased to use Boolean operations. There is a set of rules that reduces the number of parenthesis required. These are called **operator precedence rules**.

(i) Other things being equal, operations are performed left-to-right.

(ii) Operations between parenthesis are done first, starting with the innermost of nested parenthesis.

(iii) All complimentations are computed next.

(iv) All intersections are done next.

(v) All unions are performed next.

(vi) Tests of set membership and computations, equality or inequality are performed last.

Special operations like the set difference or the symmetric difference, defined below, are not included in the precedence rules and thus always use parenthesis.

Example 2.7 Operator precedence
Since complementation is done before intersection the symbolic definition of the difference of sets can be rewritten:

$$S - T = \{x : x \in S \cap T^c\}$$

If we were to take the set operations

$$A \cup B \cap C^c$$

and put in the parenthesis we would get

$$(A \cup (B \cap (C^c)))$$

Definition 2.11 *The* **symmetric difference** *of two sets S and T is the set of objects that are in one and only one of the sets. The symmetric difference is written $S \Delta T$. In curly brace notation:*

$$S \Delta T = \{(S - T) \cup (T - S)\}$$

Example 2.8 Symmetric differences
Let S be the set of non-negative multiples of two that are no more than twenty four. Let T be the non-negative multiples of three that are no more than twenty four. Then

$$S \Delta T = \{2, 3, 4, 8, 9, 10, 14, 15, 16, 20, 21, 22\}$$

Another way to think about this is that we need numbers that are positive multiples of 2 or 3 (but not both) that are no more than 24.

Another important tool for working with sets is the ability to compare them. We have already defined what it means for two sets to be equal, and so by implication what it means for them to be unequal. We now define another comparator for sets.

Definition 2.12 *For two sets S and T we say that S* **is a subset of** *T if each element of S is also an element of T. In formal notation $S \subseteq T$ if for all $x \in S$ we have $x \in T$.*

If $S \subseteq T$ then we also say T contains S which can be written $T \supseteq S$. If $S \subseteq T$ and $S \neq T$ then we write $S \subset T$ and we say S is a *proper* subset of T.

Example 2.9 Subsets
If $A = \{a, b, c\}$ then A has eight different subsets:

$$\emptyset \quad \{a\} \quad \{b\} \quad \{c\}$$

$$\{a, b\} \quad \{a, c\} \quad \{b, c\} \quad \{a, b, c\}$$

Notice that $A \subseteq A$ and in fact each set is a subset of itself. The empty set $\emptyset$ is a subset of every set.

We are now ready to prove our first proposition. Some new notation is required and we must introduce an important piece of mathematical culture. If we say "A if and only if B" then we mean that either A and B are both true or they are both false in any given circumstance. For example: "an integer x is even if and only if it is a multiple of 2". The phrase "if and only if" is used to establish *logical equivalence.* Mathematically, "A if and only if B" is a way of stating that A and B are simply different ways of saying the same thing. The phrase "if and only if" is abbreviated iff and is represented symbolically as the double arrow $\Leftrightarrow$. Proving an iff statement is done by independently demonstrating that each may be deduced from the other.

Proposition 2.1 *Two sets are equal if and only if each is a subset of the other. In symbolic notation:*

$$(A = B) \Leftrightarrow (A \subseteq B) \wedge (B \subseteq A)$$

Proof:

Let the two sets in question be A and B. Begin by assuming that $A = B$. We know that every set is a subset of itself so $A \subseteq A$. Since $A = B$ we may substitute into this expression on the left and obtain $B \subseteq A$. Similarly we may substitute on the right and obtain $A \subseteq B$. We have thus demonstrated that if $A = B$ then A and B are both subsets of each other, giving us the first half of the iff.

Assume now that $A \subseteq B$ and $B \subseteq A$. Then the definition of subset tells us that any element of A is an element of B. Similarly any element of B is an element of A. This means that A and B have the same elements which satisfies the definition of set equality. We deduce $A = B$ and we have the second half of the iff. □

A note on mathematical grammar: the symbol □ indicates the end of a proof. On a paper turned in by a student it is usually taken to mean "I think the proof ends here". Any proof should have a □ to indicate its end. The student should also note the lack of calculations in the above proof. If a proof cannot be read back in (sometimes overly formal) English then it is probably incorrect. Mathematical symbols should be used for the sake of brevity or clarity, not to obscure meaning.

Proposition 2.2 De Morgan's Laws *Suppose that S and T are sets. DeMorgan's Laws state that*

(i) $(S \cup T)^c = S^c \cap T^c$, *and*

(ii) $(S \cap T)^c = S^c \cup T^c$.

Proof:

Let $x \in (S \cup T)^c$; then x is not a member of S or T. Since x is not a member of S we see that $x \in S^c$. Similarly $x \in T^c$. Since x is a member of both these sets we see that $x \in S^c \cap T^c$ and we see that $(S \cup T)^c \subseteq S^c \cap T^c$. Let $y \in S^c \cap T^c$. Then the definition of intersection tells us that $y \in S^c$ and $y \in T^c$. This in turn lets us deduce that y is not a member of $S \cup T$, since it is not in either set, and so we see that $y \in (S \cup T)^c$. This demonstrates that $S^c \cap T^c \subseteq (S \cup T)^c$. Applying Proposition 2.1 we get that $(S \cup T)^c = S^c \cap T^c$ and we have proven part (i). The proof of part (ii) is left as an exercise. □

In order to prove a mathematical statement you must prove it is always true. In order to disprove a mathematical statement you need only find a single instance where it is false. It is thus possible for a false mathematical statement to be "true most of the time". In the next chapter we will develop the theory of prime numbers. For now we will assume the reader has a modest familiarity with the primes. The statement "Prime numbers are odd" is false once, because 2 is a prime number. All the other prime numbers are odd. The statement is a false one. This very strict definition of what makes a statement true is a convention in mathematics. We call 2 a *counter example.* It is thus necessary to find only one counter-example to demonstrate a statement is (mathematically) false.

Example 2.10 Disproof by counter example
Prove that the statement $A \cup B = A \cap B$ is false.

Let $A = \{1, 2\}$ and $B = \{3, 4\}$. Then $A \cap B = \emptyset$ while $A \cup B = \{1, 2, 3, 4\}$. The sets A and B form a counter-example to the statement.

Problems

Problem 2.1 *Which of the following are sets? Assume that a proper universal set has been chosen and answer by listing the names of the collections of objects that are sets. Warning: at least one of these items has an answer that, while likely, is not 100% certain.*

(i) $A = \{2, 3, 5, 7, 11, 13, 19\}$

(ii) $B = \{A, E, I, O, U\}$

(iii) $C = \{\sqrt{x} : x < 0\}$

(iv) $D = \{1, 2, A, 5, B, Q, 1, V\}$

(v) E is the list of first names of people in the 1972 phone book in Lawrence Kansas in the order they appear in the book. There were more than 35,000 people in Lawrence that year.

(vi) F is a list of the weight, to the nearest kilogram, of all people that were in Canada at any time in 2007.

(vii) G is a list of all weights, to the nearest kilogram, that at least one person in Canada had in 2007.

Problem 2.2 *Suppose that we have the set* $U = \{n : 0 \le n < 100\}$ *of whole numbers as our universal set. Let* P *be the prime numbers in* U*, let* E *be the even numbers in* U*, and let* $F = \{1, 2, 3, 5, 8, 13, 21, 34, 55, 89\}$*. Describe the following sets either by listing them or with a careful English sentence.*

(i) E^c,

(ii) $P \cap F$,

(iii) $P \cap E$,

(iv) $F \cap E \cup F \cap E^c$, and

(v) $F \cup F^c$.

Problem 2.3 *Suppose that we take the universal set* $\mathcal{U}$ *to be the integers. Let* S *be the even integers, let* T *be the integers that can be obtained by tripling any one integer and adding one to it, and let* V *be the set of numbers that are whole multiples of both two and three.*

(i) Write S, T, and V using symbolic notation.

(ii) Compute $S \cap T$, $S \cap V$ and $T \cap V$ and give symbolic representations that do not use the symbols S, T, or V on the right hand side of the equals sign.

Problem 2.4 *Compute the cardinality of the following sets. You may use other texts or the internet.*

(i) Two digit positive odd integers.

(ii) Elements present in a sucrose molecule.

(iii) Isotopes of hydrogen that are not radioactive.

(iv) Planets orbiting the same star as the planet you are standing on that have moons. Assume that Pluto is a minor planet.

(v) Elements with seven electrons in their valence shell. Remember that Ununoctium was discovered in 2002 so be sure to use a relatively recent reference.

(vi) Subsets of $S = \{a, b, c, d\}$ with cardinality 2.

(vii) Prime numbers whose base-ten digits sum to ten. Be careful, some have three digits.

Problem 2.5 *Find an example of an infinite set that has a finite complement, be sure to state the universal set.*

Problem 2.6 *Find an example of an infinite set that has an infinite complement, be sure to state the universal set.*

Problem 2.7 *Add parenthesis to each of the following expressions that enforce the operator precedence rules as in Example 2.7. Notice that the first three describe sets while the last returns a logical value (true of false).*

(i) $A \cup B \cup C \cup D$

(ii) $A \cup B \cap C \cup D$

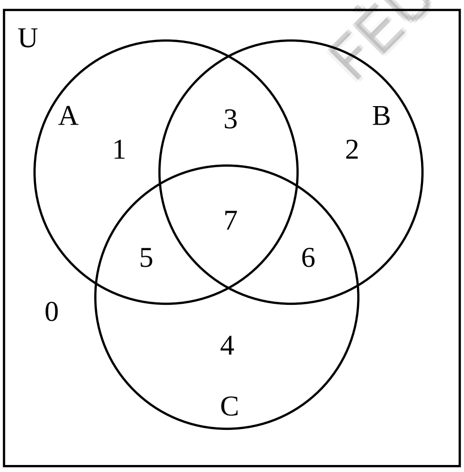
(iii) $A^c \cap B^c \cup C$

(iv) $A \cup B = A \cap C$

Problem 2.8 *Give the Venn diagrams for the following sets.*

(i) $A - B$ (ii) $B - A$ (iii) $A^c \cap B$

(iv) $A \Delta B$ (v) $(A \Delta B)^c$ (vi) $A^c \cup B^c$



Problem 2.9 *Examine the Venn diagram above. Notice that every combination of sets has a unique number in common. Construct a similar collection of four sets.*

Problem 2.10 *Read Problem 2.9. Can a system of sets of this sort be constructed for any number of sets? Explain your reasoning.*

Problem 2.11 *Suppose we take the universal set to be the set of non-negative integers. Let E be the set of even numbers, O be the set of odd numbers and $F = \{0, 1, 2, 3, 5, 8, 13, 21, 34, 89, 144, ...\}$ be the set of Fibonacci numbers. The Fibonacci sequence is $0, 1, 1, 2, 3, 5, 8, \ldots$ in which the next term is obtained by adding the previous two.*

(i) Prove that the intersection of F with E and O are both infinite.

(ii) Make a Venn diagram for the sets E, F, and O, and explain why this is a Mickey-Mouse problem.

Problem 2.12 *A binary operation $\odot$ is* **commutative** *if $x \odot y = y \odot x$. An example of a commutative operation is multiplication. Subtraction is non-commutative. Determine, with proof, if union, intersection, set difference, and symmetric difference are commutative.*

Problem 2.13 *An* **identity** *for an operation $\odot$ is an object i so that, for all objects x, $i \odot x = x \odot i = x$. Find, with proof, identities for the operations* set union *and* set intersection.

Problem 2.14 *Prove part (ii) of Proposition 2.2.*

Problem 2.15 *Prove that*

$$A \cup (B \cup C) = (A \cup B) \cup C$$

Problem 2.16 *Prove that*

$$A \cap (B \cap C) = (A \cap B) \cap C$$

Problem 2.17 *Prove that*

$$A\Delta(B\Delta C) = (A\Delta B)\Delta C$$

Problem 2.18 *Disprove that*

$$A\Delta(B \cup C) = (A\Delta B) \cup C$$

Problem 2.19 *Consider the set $S = \{1, 2, 3, 4\}$. For each $k = 0, 1, \ldots, 4$ how many k element subsets does S have?*

Problem 2.20 *Suppose we have a set S with $n \geq 0$ elements. Find a formula for the number of different subsets of S that have k elements.*

Problem 2.21 *For finite sets S and T, prove*

$$|S \cup T| = |S| + |T| - |S \cap T|$$

2.2 Mathematical Induction

Mathematical induction is a technique used in proving mathematical assertions. The basic idea of induction is that we prove that a statement is true in one case and then also prove that if it is true in a given case it is true in the next case. This then permits the cases for which the statement is true to cascade from the initial true case. We will start with the mathematical foundations of induction.

We assume that the reader is familiar with the symbols $<$, $>$, $\leq$ and $\geq$. From this point on we will denote the set of integers by the symbol $\mathbb{Z}$. The non-negative integers are called the *natural numbers*. The symbol for the set of natural numbers is $\mathbb{N}$. Any mathematical system rests on a foundation of axioms. Axioms are things that we simply assume to be true. We will assume the truth of the following principle, adopting it as an axiom.

The well-ordering principle: Every non-empty set of natural numbers contains a smallest element.

The well ordering principle is an axiom that agrees with the common sense of most people familiar with the natural numbers. An empty set does not contain a smallest member because it contains no members at all. As soon as we have a set of natural numbers with some members then we can order those members in the usual fashion. Having ordered them, one will be smallest. This intuition agreeing with this latter claim depends strongly on the fact the integers are "whole numbers" spaced out in increments of one. To see why this is important consider the smallest positive distance. If such a distance existed, we could cut it in half to obtain a smaller distance - the quantity contradicts its own existence. The well-ordering principle can be used to prove the correctness of induction.

Theorem 2.1 Mathematical Induction I *Suppose that $P(n)$ is a proposition that it either true or false for any given natural numbers n. If*

(i) $P(0)$ is true and,

(ii) when $P(n)$ is true so is $P(n + 1)$

Then we may deduce that $P(n)$ is true for any natural number.

Proof:

Assume that (i) and (ii) are both true statements. Let S be the set of all natural numbers for which $P(n)$ is false. If S is empty then we are done, so assume that S is not empty. Then, by the well ordering principle, S has a least member m. By (i) above $m \neq 0$ and so $m-1$ is a natural number. Since m is the smallest member of S it follows that $P(m-1)$ is true. But this means, by (ii) above, that $P(m)$ is true. We have a contradiction and so our assumption that $S \neq \emptyset$ must be wrong. We deduce S is empty and that as a consequence $P(n)$ is true for all $n \in \mathbb{N}$. □

The technique used in the above proof is called *proof by contradiction*. We start by assuming the logical opposite of what we want to prove, in this case that there is some m for which $P(m)$ is false, and from that assumption we derive an impossibility. If an assumption can be used to demonstrate an impossibility then it is false and its logical opposite is true.

A nice problem on which to demonstrate mathematical induction is counting how many subsets a finite set has.

Proposition 2.3 Subset counting. *A set S with n elements has 2^n subsets.*

Proof:

First we check that the proposition is true when $n = 0$. The empty set has exactly one subset: itself. Since $2^0 = 1$ the proposition is true for $n = 0$. We now assume the proposition is true for some n. Suppose that S is a set with $n+1$ members and that $x \in S$. Then $S - \{x\}$ (the set difference of S and a set $\{x\}$ containing only x) is a set of n elements and so, by the assumption, has 2^n subsets. Now every subset of S either contains x or it fails to. Every subset of S that does not contain x is a subset of $S - \{x\}$ and so there are 2^n such subsets of S. Every subset of S that contains x may be obtained in exactly one way from one that does not by taking the union with $\{x\}$. This means that the number of subsets of S containing or failing to contain x are equal. This means there are 2^n subsets of S containing x. The total number of subsets of S is thus $2^n + 2^n = 2^{n+1}$. So if we assume the proposition is true for n we can demonstrate that it is also true for $n+1$. It follows by mathematical induction that the proposition is true for all natural numbers. □

The set of all subsets of a given set is itself an important object and so has a name.

Definition 2.13 *The set of all subsets of a set S is called the* **powerset** *of S. The notation for the powerset of S is $\mathcal{P}(S)$.*

This definition permits us to rephrase Proposition 2.3 as follows: the power set of a set of n elements has size 2^n.

Theorem 2.1 lets us prove propositions that are true on the natural numbers, starting at zero. A small modification of induction can be used to prove statements that are true only for those $n \geq k$ for any integer k. All that is needed is to use induction on a proposition $Q(n-k)$ where $Q(n-k)$ is logically equivalent to $P(n)$. If $Q(n-k)$ is true for $n - k \geq 0$ then $P(n)$ is true for $n \geq k$ and we have the modified induction. The practical difference is that we start with k instead of zero.

Example 2.11 *Prove that $n^2 \geq 2n$ for all $n \geq 2$.*

Notice that $2^2 = 4 = 2 \times 2$ so the proposition is true when $n = 2$. We next assume that $P(n)$ is true for some n and we compute:

$$\begin{aligned}
n^2 &\geq 2n \\
n^2 + 2n + 1 &\geq 2n + 2n + 1 \\
(n+1)^2 &\geq 2n + 2n + 1 \\
(n+1)^2 &\geq 2n + 1 + 1 \\
(n+1)^2 &\geq 2n + 2 \\
(n+1)^2 &\geq 2(n+1)
\end{aligned}$$

To move from the third step to the fourth step we use the fact that $2n > 1$ when $n \geq 2$. The last step is $P(n+1)$, which means we have deduced $P(n+1)$ from $P(n)$. Using the modified form of induction we have proved that $n^2 \geq 2n$ for all $n \geq 2$.

It is possible to formalize the procedure for using mathematical induction into a three-part process. Once we have a proposition $P(n)$,

(i) First demonstrate a *base case* by directly demonstrating $P(k)$,

(ii) Next make the *induction hypothesis* that $P(n)$ is true for some n,

(iii) Finally, starting with the assumption that $P(n)$ is true, demonstrate $P(n+1)$.

These steps permit us to deduce that $P(n)$ is true for all $n \geq k$.

Example 2.12 *Using induction, prove*

$$1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$$

In this case $P(n)$ is the statement

$$1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$$

Base case: $1 = \frac{1}{2}1(1+1)$, *so $P(1)$ is true.* **Induction hypothesis:** *for some n,*

$$1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$$

Compute:

$$\begin{aligned} 1 + 2 + \cdots + (n+1) &= 1 + 2 + \cdots + n + (n+1) \\ &= \frac{1}{2}n(n+1) + (n+1) \\ &= \frac{1}{2}\left(n(n+1) + 2(n+1)\right) \\ &= \frac{1}{2}\left(n^2 + 3n + 2\right) \\ &= \frac{1}{2}(n+1)(n+2) \\ &= \frac{1}{2}(n+1)((n+1)+1) \end{aligned}$$

and so we have shown that if $P(n)$ is true then so is $P(n+1)$. We have thus proven that $P(n)$ is true for all $n \geq 1$ by mathematical induction.

We now introduce *sigma notation* which makes problems like the one worked in Example 2.12 easier to state and manipulate. The symbol $\sum$ is used to add up lists of numbers. If we wished to sum some formula $f(i)$ over a range from a to b, that is to say $a \leq i \leq b$, then we write :

$$\sum_{i=a}^{b} f(i)$$

On the other hand if S is a set of numbers and we want to add up $f(s)$ for all $s \in S$ we write:

$$\sum_{s \in S} f(s)$$

The result proved in Example 2.12 may be stated in the following form using sigma notation.

$$\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$$

Proposition 2.4 *Suppose that c is a constant and that $f(i)$ and $g(i)$ are formulas. Then*

(i) $\sum_{i=a}^{b} (f(i) + g(i)) = \sum_{i=a}^{b} f(i) + \sum_{i=a}^{b} g(i)$

(ii) $\sum_{i=a}^{b} (f(i) - g(i)) = \sum_{i=a}^{b} f(i) - \sum_{i=a}^{b} g(i)$

(iii) $\sum_{i=a}^{b} c \cdot f(i) = c \cdot \sum_{i=a}^{b} f(i)$.

Proof:

Part (i) and (ii) are both simply the associative law for addition: $a+(b+c) = (a+b)+c$ applied many times. Part (iii) is a similar multiple application of the distributive law $ca + cb = c(a+b)$. □

The sigma notation lets us work with indefinitely long (and even infinite) sums. There are other similar notations. If $A_1, A_2, \ldots, A_n$ are sets then the intersection or union of all these sets can be written:

$$\bigcap_{i=1}^{n} A_i$$

$$\bigcup_{i=1}^{n} A_i$$

Similarly if $f(i)$ is a formula on the integers then

$$\prod_{i=1}^{n} f(i)$$

is the notation for computing the product $f(1) \cdot f(2) \cdot \cdots \cdot f(n)$. This notation is called **Pi** notation.

Definition 2.14 *When we solve an expression involving* $\sum$ *to obtain a formula that does not use* $\sum$ *or "..." as in Example 2.12 then we say we have found a* **closed form** *for the expression. Example 2.12 finds a closed form for* $\sum_{i=1}^{n} i$.

At this point we introduce a famous mathematical sequence in order to create an arena for practicing proofs by induction.

Definition 2.15 *The* **Fibonacci numbers** *are defined as follows.* $f_1 = f_2 = 1$ *and, for* $n \geq 3$, $f_n = f_{n-1} + f_{n-2}$.

Example 2.13 *The Fibonacci numbers with four or fewer digits are:* $f_1 = 1$, $f_2 = 1$, $f_3 = 2$, $f_4 = 3$, $f_5 = 5$, $f_6 = 8$, $f_7 = 13$, $f_8 = 21$, $f_9 = 34$, $f_{10} = 55$, $f_{11} = 89$, $f_{12} = 144$, $f_{13} = 233$, $f_{14} = 377$, $f_{15} = 610$, $f_{16} = 987$, $f_{17} = 1597$, $f_{18} = 2584$, $f_{19} = 4181$, *and* $f_{20} = 6765$.

Example 2.14 *Prove that the Fibonacci number* f_{3n} *is even.*

Solution:

Notice that $f_3 = 2$ *and so the proposition is true when* $n = 1$. *Assume that the proposition is true for some* $n \geq 1$. *Then:*

$$\begin{aligned} f_{3(n+1)} &= f_{3n+3} && (2.1)\\ &= f_{3n+2} + f_{3n+1} && (2.2)\\ &= f_{3n+1} + f_{3n} + f_{3n+1} && (2.3)\\ &= 2 \cdot f_{3n+1} + f_{3n} && (2.4) \end{aligned}$$

but this suffices because f_{3n} *is even by the induction hypothesis while* $2 \cdot f_{3n+1}$ *is also even. The sum is thus even and so* $f_{3(n+1)}$ *is even. If follows by induction that* f_{3n} *is even for all* n. □

Problems

Problem 2.22 *Suppose that* $S = \{a, b, c\}$. *Compute and list explicitly the members of the powerset,* $\mathcal{P}(S)$.

Problem 2.23 *Prove that for a finite set* X *that*

$$|X| \leq |\mathcal{P}(X)|$$

Problem 2.24 *Suppose that* $X \subseteq Y$ *with* $|Y| = n$ *and* $|X| = m$. *Compute the number of subsets of* Y *that contain* X.

Problem 2.25 *Compute the following sums.*

(i) $\sum_{i=1}^{20} i$,

(ii) $\sum_{i=10}^{30} i$, and

(iii) $\sum_{i=-20}^{21} i$.

Problem 2.26 *Using mathematical induction, prove the following formulas.*

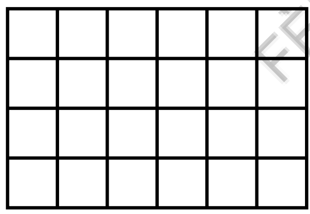
(i) $\sum_{i=1}^{n} 1 = n$,

(ii) $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$, and

(iii) $\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}$.

Problem 2.27 *If* $f(i)$ *and* $g(i)$ *are formulas and* c *and* d *are constants prove that*

$$\sum_{i=a}^{b}(c \cdot f(i) + d \cdot g(i)) = c \cdot \sum_{i=a}^{b} f(i) + d \cdot \sum_{i=a}^{b} g(i)$$



Problem 2.28 *Suppose you want to break an* $n \times m$ *chocolate bar, like the* 6×4 *example shown above, into pieces corresponding to the small squares shown. What is the minimum number of breaks you can make? Prove your answer is correct.*

Problem 2.29 *Prove by induction that the sum of the first* n *odd numbers equals* n^2.

Problem 2.30 *Compute the sum of the first* n *positive even numbers.*

Problem 2.31 *Find a closed form for*

$$\sum_{i=1}^{n} i^2 + 3i + 5$$

Problem 2.32 *Let $f(n,3)$ be the number of subsets of $\{1,2,\ldots,n\}$ of size 3. Using induction, prove that $f(n,3)=\frac{1}{6}n(n-1)(n-2)$.*

Problem 2.33 *Suppose that we have sets $X_1,X_2,\ldots,X_n$ and $Y_1,Y_2,\ldots,Y_n$ such that $X_i \subseteq Y_i$. Prove that the intersection of all the X_i is a subset of the intersection of all the Y_i:*

$$\bigcap_{i=1}^{n} X_i \subseteq \bigcap_{i=1}^{n} Y_i$$

Problem 2.34 *Suppose that $S_1,S_2,\ldots S_n$ are sets. Prove the following generalization of DeMorgan's laws:*

(i) $\left(\bigcap_{i=1}^{n} S_i\right)^c = \bigcup_{i=1}^{n} S_i^c$, *and*

(ii) $\left(\bigcup_{i=1}^{n} S_i\right)^c = \bigcap_{i=1}^{n} S_i^c$.

Problem 2.35 *Prove by induction that the Fibonacci number f_{4n} is a multiple of 3.*

Problem 2.36 *Prove that if r is a real number $r \neq 1$ and $r \neq 0$ then*

$$\sum_{i=0}^{n} r^i = \frac{1-r^{n+1}}{1-r}$$

Problem 2.37 *Prove by induction that the Fibonacci number f_{5n} is a multiple of 5.*

Problem 2.38 *Prove by induction that the Fibonacci number f_n has the value*

$$f_n = \frac{\sqrt{5}}{5}\cdot\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{\sqrt{5}}{5}\cdot\left(\frac{1-\sqrt{5}}{2}\right)^n$$

Problem 2.39 *Prove that for sufficiently large n the Fibonacci number f_n is the integer closest to*

$$\frac{\sqrt{5}}{5}\left(\frac{1+\sqrt{5}}{2}\right)^n$$

and compute the exact value of f_{30}. Show your work (i.e. don't look the result up on the net).

Problem 2.40 *Prove that $\frac{n(n-1)(n-2)(n-3)}{24}$ is a whole number for any whole number n.*

Problem 2.41 *Consider the statement "All cars are the same color." and the following "proof".*

Proof:

We will prove for $n \geq 1$ that for any set of n cars all the cars in the set have the same color.

- *Base Case: n=1 If there is only one car then clearly there is only one color the car can be.*
- *Inductive Hypothesis: Assume that for any set of n cars there is only one color.*
- *Inductive step: Look at any set of n + 1 cars. Number them: $1,2,3,\ldots,n,n+1$. Consider the sets $\{1,2,3,\ldots,n\}$ and $\{2,3,4,...,n+1\}$. Each is a set of only n cars, therefore for each set there is only one color. But the n^{th} car is in both sets so the color of the cars in the first set must be the same as the color of the cars in the second set. Therefore there must be only one color among all n + 1 cars.*
- *The proof follows by induction.* □

What are the problems with this proof?

2.3 Functions

In this section we will define functions and extend much of our ability to work with sets to infinite sets. There are a number of different types of functions and so this section contains a great deal of terminology.

Recall that two finite sets are the same size if they contain the same number of elements. It is possible to make this idea formal by using functions and, once the notion is formally defined, it can be applied to infinite sets.

Definition 2.16 *An* **ordered pair** *is a collection of two elements with the added property that one element comes first and one element comes second. The set containing only x and y (for $x \neq y$) is written $\{x,y\}$. The ordered pair containing x and y with x first is written (x,y). Notice that while $\{x,x\}$ is not a well defined set, (x,x) is a well defined ordered pair because the two copies of x are different by virtue of coming first and second.*

The reason for defining ordered pairs at this point is that it permits us to make an important formal definition that pervades the rest of mathematics.

Definition 2.17 *A* **function** f *with* **domain** S *and* **range** T *is a set of ordered pairs* (s,t) *with first element from* S *and second element from* T *that has the property that every element of* S *appears exactly once as the first element in some ordered pair. We write* $f : S \to T$ *for such a function.*

Example 2.15 *Suppose that* $A = \{a, b, c\}$ *and* $B = \{0, 1\}$ *then*

$$f = \{(a, 0), (b, 1), (c, 0)\}$$

is a function from A *to* B*. The function* $f : A \to B$ *can also be specified by saying* $f(a) = 0$, $f(b) = 1$ *and* $f(c) = 0$.

The set of ordered pairs $\{(a, 0), (b, 1)\}$ *is not a function from* A *to* B *because* c *is not the first coordinate of any ordered pair. The set of ordered pairs* $\{(a, 0), (a, 1), (b, 0), (c, 0)\}$ *is not a function from* A *to* B *because* a *appears as the first coordinate of two different ordered pairs.*

In calculus you may have learned the *vertical line rule* that states that the graph of a function may not intersect a vertical line at more than one point. This corresponds to requiring that each point in the domain of the function appear in only one ordered pair. In set theory, all functions are required to state their domain and range when they are defined. In calculus functions had a domain that was a subset of the real numbers and you were sometimes required to identify the subset.

Example 2.16 *This example contrasts the way functions were treated in a typical calculus course with the way we treat them in set theory.*

Calculus: *find the domain of the function*

$$f(x) = \sqrt{x}$$

Since we know that the square root function exists only for non-negative real numbers the domain is $\{x : x \geq 0\}$.

Set theory: *the function* $f = \sqrt{x}$ *from the non-negative real numbers to the real numbers is the set of ordered pairs* $\{(r^2, r) : r \geq 0\}$*. This function is well defined because each non-negative real number is the square of some positive real number.*

The major contrasts between functions in calculus and functions in set theory are:

(i) The domain of functions in calculus are often specified only by implication (you have to know how all the functions used work) and are almost always a subset of the real numbers. The domain in set theory must be explicitly specified and may be any set at all.

(ii) Functions in calculus typically had graphs that you could draw and look at. Geometric intuition driven by the graphs plays a major role in our understanding of functions. Functions in set theory are seldom graphed and often don't have a graph.

A point of similarity between calculus and set theory is that the range of the function is not explicitly specified. When we have a function $f : S \to T$ then the range of f is a subset of T.

Definition 2.18 *If* f *is a function then we denote the domain of* f *by* $dom(f)$ *and the range of* f *by* $rng(f)$

Example 2.17 *Suppose that* $f(n) : \mathbb{N} \to \mathbb{N}$ *is defined by* $f(n) = 2n$*. Then the domain and range of* f *are the integers:* $dom(f) = rng(f) = \mathbb{N}$*. If we specify the ordered pairs of* f *we get*

$$f = \{(n, 2n) : n \in \mathbb{N}\}$$

There are actually two definitions of range that are used in mathematics. The definition we are using, the set from which second coordinates of ordered pairs in a function are drawn, is also the definition typically using in computer science. The other definition is the set of second coordinates that actually appear in ordered pairs. This set, which we will define formally later, is the *image* of the function. To make matters even worse the set we are calling the range of a function is also called the *co-domain*. We include these confusing terminological notes for students that may try and look up supplemental material.

Definition 2.19 *Let X, Y, and Z be sets. The* **composition** *of two functions $f : X \to Y$ and $g : Y \to Z$ is a function $h : X \to Z$ for which $h(x) = g(f(x))$ for all $x \in X$. We write $g \circ f$ for the composition of g with f.*

The definition of the composition of two functions requires a little checking to make sure it makes sense. Since *every* point must appear as a first coordinate of an ordered pair in a function, every result of applying f to an element of X is an element of Y to which g can be applied. This means that h is a well-defined set of ordered pairs. Notice that the order of composition is important - if the sets X, Y, and Z are distinct there is only one order in which composition even makes sense.

Example 2.18 *Suppose that $f : \mathbb{N} \to \mathbb{N}$ is given by $f(n) = 2n$ while $g : \mathbb{N} \to \mathbb{N}$ is given by $g(n) = n + 4$. Then*

$$(g \circ f)(n) = 2n + 4$$

while

$$(f \circ g)(n) = 2(n + 4) = 2n + 8$$

We now start a series of definitions that divide functions into a number of classes. We will arrive at a point where we can determine if the mapping of a function is reversible, if there is a function that exactly reverses the action of a given function.

Definition 2.20 *A function $f : S \to T$ is* **injective** *or* **one-to-one** *if no element of T (no second coordinate) appears in more than one ordered pair. Such a function is called an* **injection**.

Example 2.19 *The function $f : \mathbb{N} \to \mathbb{N}$ given by $f(n) = 2n$ is an injection. The ordered pairs of f are $(n, 2n)$ and so any number that appears as a second coordinate does so once.*

The function $g : \mathbb{Z} \to \mathbb{Z}$ given by $g(n) = n^2$ is not an injection. To see this notice that g contains the ordered pairs $(1, 1)$ and $(-1, 1)$ so that 1 appears twice as the second coordinate of an ordered pair.

Definition 2.21 *A function $f : S \to T$ is* **surjective** *or* **onto** *if every element of T appears in an ordered pair. Surjective functions are called* **surjections**.

We use the symbol $\mathbb{R}$ for the real numbers. We also assume familiarity with interval notation for contiguous subsets of the reals. For real numbers $a \leq b$

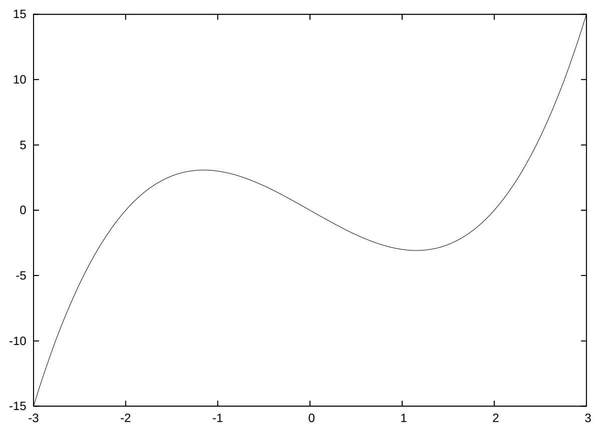
$$\begin{array}{lcl} (a, b) & \text{is} & \{x : a < x < b\} \\ (a, b] & \text{is} & \{x : a < x \leq b\} \\ [a, b) & \text{is} & \{x : a \leq x < b\} \\ [a, b] & \text{is} & \{x : a \leq x \leq b\} \end{array}$$

Example 2.20 *The function $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(n) = 5 - n$ is a surjection. If we set $m = 5 - n$ then $n = 5 - m$. This means that if we want to find some n so that $f(n)$ is, for example, 8, then $5 - 8 = -3$ and we see that $f(-3) = 8$. This demonstrates that all m have some n so that $f(n) = m$, showing that all m appear as the second coordinate of an ordered pair in f.*

The function $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = \frac{x^2}{1+x^2}$ is not a surjection because $-1 < g(x) < 1$ for all $x \in \mathbb{R}$.

Definition 2.22 *A function that is both surjective and injective is said to be* **bijective**. *Bijective functions are called* **bijections**.

Example 2.21 *The function $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(n) = n$ is a bijection. All of its ordered pairs have the same first and second coordinate. This function is called the* identity function.



The function $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = x^3 - 4x$ is not a bijection. It is not too hard to show that it is a surjection, but it fails to be an injection. The portion of the graph shown above demonstrates that $g(x)$ takes on the same value more than once. This means that *some numbers appear twice as second coordinates of ordered pairs in g. We can use the graph because g is a function from the real numbers to the real numbers.*

For a function $f : S \to T$ to be a bijection every element of S appears in an ordered pair as the first member of an ordered pair and every element of T appears in an ordered pair as the second member of an ordered pair. Another way to view a bijection is as a matching of the elements of S and T so that every element of S is paired with an element of T. For finite sets this is clearly only possible if the sets are the same size and, in fact, this is the formal definition of "same size" for sets.

Definition 2.23 *Two sets S and T are defined to be* **the same size** *or to have* **equal cardinality** *if there is a bijection $f : S \to T$.*

Example 2.22 *The sets $A = \{a, b, c\}$ and $Z = \{1, 2, 3\}$ are the same size. This is obvious because they have the same number of elements, $|A| = |Z| = 3$ but we can construct an explicit bijection*

$$f = \{(a, 3), (b, 1), (c, 2)\}$$

with each member of A appearing once as a first coordinate and each member of B appearing once as a second coordinate. This bijection is a witness *that A and B are the same size.*

Let E be the set of even integers. Then the function

$$g : \mathbb{Z} \to E$$

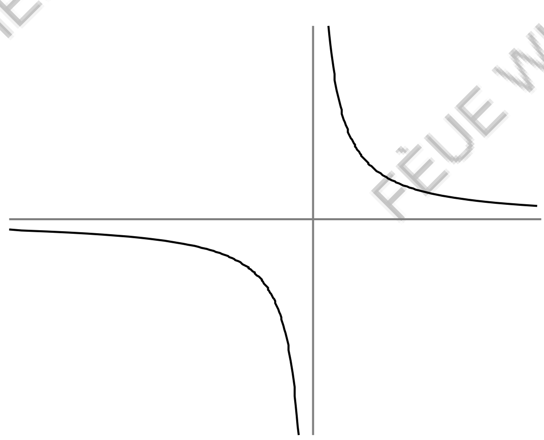
in which $g(n) = 2n$ is a bijection. Notice that each integers can be put into g and that each even integer has exactly one integer that can be doubled to make it. The existence of g is a witness that the set of integers and the set of even integers are the same size. This may seem a bit bizarre because the set $\mathbb{Z} - E$ is the infinite set of odd integers. In fact one hallmark of an infinite set is that it can be the same size as a proper subset. This also means we now have an equality set for sizes of infinite sets. We will do a good deal more with this in Chapter 3.

Bijections have another nice property: they can be unambiguously reversed.

Definition 2.24 *The* **inverse** *of a function $f : S \to T$ is a function $g : T \to S$ so that for all $x \in S$, $g(f(x)) = x$ and for all $y \in T$, $f(g(y)) = y$.*

If a function f has an inverse we use the notation f^{-1} for that inverse. Since an exponent of -1 also means reciprocal in some circumstances this can be a bit confusing. The notational confusion is resolved by considering context. So long as we keep firmly in mind that functions are sets of ordered pairs it is easy to prove the proposition/definition that follows after the next example.

Example 2.23 *If E is the set of even integers then the bijection $f(n) = 2n$ from $\mathbb{Z}$ to E has the inverse $f^{-1} : E \to \mathbb{Z}$ given by $g(2n) = n$. Notice that defining the rule for g as depending on the argument $2n$ seamlessly incorporates the fact that the domain of g is the even integers.*



If $g(x) = \frac{x}{x-1}$, shown above with its asymptotes $x = 1$ and $y = 1$ then f is a function from the set $H = \mathbb{R} - \{1\}$ to itself. The function was chosen to have asymptotes at equal x and y values; this is a bit unusual. The function g is a bijection. Notice that the graph intersects any horizontal or vertical line in at most one point. Every value except $x = 1$ may be put into g meaning that g is a function on H. Since the vertical asymptote goes off to ∞ in both directions, all values in H come out of g. This demonstrates g is a bijection. This means that it has an inverse which we now compute using a standard *technique from calculus classes.*

$$\begin{aligned} y &= \frac{x}{x-1} \\ y(x-1) &= x \\ xy - y &= x \\ xy - x &= y \\ x(y-1) &= y \\ x &= \frac{y}{y-1} \end{aligned}$$

which tells us that $g^{-1}(x) = \frac{x}{x-1}$ *so* $g = g^{-1}$*: the function is its own inverse.*

Proposition 2.5 *A function has an inverse if and only if it is a bijection.*

Proof:
Suppose that $f : S \to T$ is a bijection. Then if $g : T \to S$ has ordered pairs that are the exact reverse of those given by f it is obvious that for all $x \in S$, $g(f(x)) = x$, likewise that for all $y \in T$, $f(g(y)) = y$. We have that bijections posses inverses. It remains to show that non-bijections do not have inverses.

If $f : S \to T$ is not a bijection then either it is not a surjection or it is not an injection. If f is not a surjection then there is some $t \in T$ that appears in no ordered pair of f. This means that no matter what $g(t)$ is, $f(g(t)) \neq t$ and we fail to have an inverse. If, on the other hand, $f : S \to T$ is a surjection but fails to be an injection then for some distinct $a, b \in S$ we have that $f(a) = t = f(b)$. For $g : T \to S$ to be an inverse of f we would need $g(t) = a$ and $g(t) = b$, forcing t to appear as the first coordinate of two ordered pairs in g and so rendering g a non-function. We thus have that non-bijections do not have inverses. □

The type of inverse we are discussing above is a *two-sided inverse*. The functions f and f^{-1} are mutually inverses of one another. It is possible to find a function that is a one-way inverse of a function so that $f(g(x)) = x$ but $g(f(x))$ is not even defined. These are called *one-sided inverses*.

Note on mathematical grammar: Recall that when two notions, such as "bijection" and "has an inverse" are equivalent we use the phrase "if and only if" (abbreviated iff) to phrase a proposition declaring that the notions are equivalent. A proposition that A iff B is proven by first assuming A and deducing B and then separately assuming B and deducing A. The formal symbol for A iff B is $A \Leftrightarrow B$. Likewise we have symbols for the ability to deduce B given A, $A \Rightarrow B$ and vice-versa $B \Rightarrow A$. These symbols are spoken "A implies B" and "B implies A" respectively.

Proposition 2.6 *Suppose that* X*,* Y*, and* Z *are sets. If* $f : X \to Y$ *and* $g : Y \to Z$ *are bijections then so is* $g \circ f : X \to Z$*.*

Proof: *this proof is left as an exercise.*

Definition 2.25 *Suppose that* $f : A \to B$ *is a function. The* **image of** A **in** B *is the subset of* B *made of elements that appear as the second element of ordered pairs in* f*. Colloquially the image of* f *is the set of elements of* B *hit by* f*. We use the notation* $Im(f)$ *for images. In other words* $Im(f) = \{f(a) : a \in A\}$*.*

Example 2.24 *If* $f : \mathbb{N} \to \mathbb{N}$ *is given by the rule* $f(n) = 3n$ *then the set* $T = \{0, 3, 6, \ldots\}$ *of natural numbers that are multiples of three is the image of* f*. Notation:* $Im(f) = T$*.*
If $g : \mathbb{R} \to \mathbb{R}$ *given by* $g(x) = x^2$ *then*

$$Im(g) = \{y : y \geq 0, y \in \mathbb{R}\}$$

There is a name for the set of all ordered pairs drawn from two sets.

Definition 2.26 *If* A *and* B *are sets then the set of all ordered pairs with the first element from* A *and the second from* B *is called the* **Cartesian Product** *of* A *and* B*.*

The notation for the Cartesian product of A and B is $A \times B$. using curly brace notation:

$$A \times B = \{(a, b) : a \in A, b \in B\}$$

Example 2.25 *If* $A = \{1, 2\}$ *and* $B = \{x, y\}$ *then*

$$A \times B = \{(1, x), (1, y), (2, x), (2, y)\}$$

The **Cartesian plane** *is an example of a Cartesian product of the real numbers with themselves:* $\mathbb{R} \times \mathbb{R}$*.*

2.3.1 Permutations

In this section we will look at a very useful sort of function, bijections of finite sets.

Definition 2.27 *A* **permutation** *is a bijection of a finite set with itself. Likewise a bijection of a finite set X with itself is called a* **permutation of X**.

Example 2.26 *Let $A = \{a, b, c\}$ then the possible permutations of A consist of the following six functions:*

{(a,a)(b,b)(c,c)}	{(a,a)(b,c)(c,b)}
{(a,b)(b,a)(c,c)}	{(a,b)(b,c)(c,a)}
{(a,c)(b,a)(c,b)}	{(a,c)(b,b)(c,a)}

Notice that the number of permutations of three objects does not depend on the identity of those objects. In fact there are always six permutations of any set of three objects. We now define a handy function that uses a rather odd notation. The method of showing permutations in Example 2.26, explicit listing of ordered pairs, is a bit cumbersome.

Definition 2.28 *Assume that we have agreed on an order, e.g. a,b,c, for the members of a set $X = \{a, b, c\}$. Then* **one-line** *notation for a permutation f consists of listing the first coordinate of the ordered pairs in the agreed on order. The table in Example 2.26 would become:*

abc	**acb**
bac	**bca**
cab	**cba**

in one line notation. Notice the saving of space.

Definition 2.29 *The* **factorial** *of a natural number n is the product*

$$n(n-1)(n-2)\cdots 3\cdot 2\cdot 1 = \prod_{i=1}^{n} i$$

with the convention that the factorial of 0 is 1. We denote the factorial of n as $n!$, spoken "n factorial".

Example 2.27 *Here are the first few factorials:*

n	0	1	2	3	4	5	6	7
$n!$	1	1	2	6	24	120	720	5040

Proposition 2.7 *The number of permutations of a finite set with n elements is $n!$.*

Proof: *this proof is left as an exercise.*

Notice that one implication of Proposition 2.6 is that the composition of two permutations is a permutation. This means that the set of permutations of a set is *closed* under functional composition.

Definition 2.30 *A* **fixed point** *of a function $f : S \to S$ is any $x \in S$ such that $f(x) = x$. We say that* **f fixes x**.

Problems

Problem 2.42 *Suppose for finite sets A and B that $f : A \to B$ is an injective function. Prove that*

$$|B| \geq |A|$$

Problem 2.43 *Suppose that for finite sets A and B that $f : A \to B$ is a surjective function. Prove that $|A| \geq |B|$.*

Problem 2.44 *Using functions from the integers to the integers give an example of*

(i) A function that is an injection but not a surjection.

(ii) A function that is a surjection but not an injection.

(iii) A function that is neither an injection nor a surjection.

(iv) A bijection that is not the identity function.

Problem 2.45 *For each of the following functions from the real numbers to the real numbers say if the function is surjective or injective. It may be neither.*

(i) $f(x) = x^2$ (ii) $g(x) = x^3$

(iii) $h(x) = \begin{cases} \sqrt{x} & x \geq 0 \\ -\sqrt{-x} & x < 0 \end{cases}$

Interlude

The Collatz Conjecture

One of the most interesting features of mathematics is that it is possible to phrase problems in a few lines that turn out to be incredibly hard. The Collatz conjecture was first posed in 1937 by Lothar Collatz. Define the function f from the natural numbers to the natural numbers with the rule

$$f(n) = \begin{cases} 3n+1 & \text{n odd} \\ \frac{n}{2} & \text{n even} \end{cases}$$

Collatz' conjecture is that if you apply f repeatedly to a positive integer then the resulting sequence of numbers eventually arrives at one. If we start with 17, for example, the result of repeatedly applying f is:

$$f(17) = 52, f(52) = 26, f(26) = 13, f(13) = 40, f(40) = 20, f(20) = 10,$$
$$f(10) = 5, f(5) = 16, f(16) = 8, f(8) = 4, f(4) = 2, f(2) = 1$$

The sequences of numbers generated by repeatedly applying f to a natural number are called *hailstone sequences* with the collapse of the value when a large power of 2 appears being analogous to the impact of a hailstone. If we start with the number 27 then 111 steps are required to reach one and the largest intermediate number is 9232. This quite irregular behavior of the sequence is not at all apparent in the original phrasing of the problem.

The Collatz conjecture has been checked for numbers up to 5×2^{61} (about 5.764×10^{18}) by using a variety of computational tricks. It has not, however, been proven or disproven. The very simple statement of the problem causes mathematicians to underestimate the difficulty of the problem. At one point a mathematician suggested that the problem might have been developed by the Russians as a way to slow American mathematical research. This was after several of his colleagues spent months working on the problem without obtaining results.

A simple (but incorrect) argument suggests that hailstone sequences ought to grow indefinitely. Half of all numbers are odd, half are even. The function f slightly more than triples odd numbers and divides even numbers in half. Thus, on average, f increases the value of numbers. The problem is this: half of all even numbers are multiples of four and so are divided in half twice. One-quarter of all even numbers are multiples of eight and so get divided in half three times, and so on. The net effect of factors that are powers of two is to defeat the simple argument that f grows "on average".

Problem 2.46 *True or false (and explain): The function $f(x) = \frac{x-1}{x+1}$ is a bijection from the real numbers to the real numbers.*

Problem 2.47 *Find a function that is an injection of the integers into the even integers that does not appear in any of the examples in this chapter.*

Problem 2.48 *Suppose that $B \subset A$ and that there exists a bijection $f : A \to B$. What may be reasonably deduced about the set A?*

Problem 2.49 *Suppose that A and B are finite sets. Prove that $|A \times B| = |A| \cdot |B|$.*

Problem 2.50 *Suppose that we define $h : \mathbb{N} \to \mathbb{N}$ as follows. If n is even then $h(n) = n/2$ but if n is odd then $h(n) = 3n+1$. Determine if h is a (i) surjection or (ii) injection.*

Problem 2.51 *Prove proposition 2.6.*

Problem 2.52 *Prove or disprove: the composition of injections is an injection.*

Problem 2.53 *Prove or disprove: the composition of surjections is a surjection.*

Problem 2.54 *Prove proposition 2.7.*

Problem 2.55 *List all permutations of*

$$X = \{1, 2, 3, 4\}$$

using one-line notation.

Problem 2.56 *Suppose that X is a set and that f, g, and h are permutations of X. Prove that the equation $f \circ g = h$ has a solution g for any given permutations f and h.*

Problem 2.57 *Examine the permutation f of $Q = \{a, b, c, d, e\}$ which is* **bcaed** *in one line notation. If we create the series $f, f \circ f, f \circ (f \circ f), \ldots$ does the identity function,* **abcde**, *ever appear in the series? If so, what is its first appearance? If not, why not?*

Problem 2.58 *If f is a permutation of a finite set, prove that the sequence $f, f \circ f, f \circ (f \circ f), \ldots$ must contain repeated elements.*

Problem 2.59 *Suppose that X and Y are finite sets and that $|X| = |Y| = n$. Prove that there are $n!$ bijections of X with Y.*

Problem 2.60 *Suppose that X and Y are sets with $|X| = n$, $|Y| = m$. Count the number of functions from X to Y.*

Problem 2.61 *Suppose that X and Y are sets with $|X| = n$, $|Y| = m$ for $m > n$. Count the number of injections of X into Y.*

Problem 2.62 *For a finite set S with a subset T prove that the permutations of S that have all members of T as fixed points form a set that is closed under functional composition.*

Problem 2.63 *Compute the number of permutations of a set S with n members that fix at least $m < n$ points.*

Problem 2.64 *Using any technique at all, estimate the fraction of permutations of an n-element set that have no fixed points. This problem is intended as an exploration.*

Problem 2.65 *Let X be a finite set with $|X| = n$. Let $C = X \times X$. How many subsets of C have the property that every element of X appears once as a first coordinate of some ordered pair and once as a second coordinate of some ordered pair?*

Problem 2.66 *An alternate version of Sigma ($\sum$) and Pi ($\prod$) notation works by using a set as an index. So if $S = \{1, 3, 5, 7\}$ then*

$$\sum_{s \in S} s = 16 \quad \textit{and} \quad \prod_{s \in S} s = 105$$

Given all the material so far, give and defend reasonable values for the sum and product of an empty set.

Problem 2.67 *Suppose that $f_\alpha : [0, 1] \to [0, 1]$ for $-1 < \alpha < \infty$ is given by*

$$f_\alpha(x) = \frac{(\alpha + 1)x}{\alpha x + 1},$$

prove that f_α is a bijection.

Problem 2.68 *Find, to five decimals accuracy:*

$$Ln(200!)$$

Explain how you obtained the answer.

2.4 $\infty + 1$

We conclude the chapter with a brief section that demonstrates a strange thing that can be accomplished with set notation. We choose to represent the natural numbers $0, 1, 2, \ldots$ by sets that contain the number of elements counted by the corresponding natural number. We also choose to do so as simply as possible, using only curly braces and commas. Given this the numbers and their corresponding sets are:

0 : {}
1 : {{}} = {0}
2 : {{},{{}}}={0,1}
3 : {{},{{}},{{},{{}}}}={0,1,2}
4 : {{},{{}},{{},{{}}},{{},{{}},{{},{{}}}}}
={0,1,2,3}

The trick for the above representation is this. Zero is represented by the empty set. One is represented by the set of the only thing we have constructed - zero, represented as the empty set. Similarly the representation of two is the set of the representation of zero and one (the empty set and the set of the empty set). This representation is incredibly inefficient but it uses a very small number of symbols. This representation also has a useful property. As always, we will start with a definition.

Definition 2.31 *The* **minimal set representation** *of the natural numbers is constructed as follows:*

(i) Let 0 be represented by the empty set.

(ii) For $n > 0$ let n be represented by the set $\{0, 1, \ldots, n-1\}$.

The shorthand $\{0, 1\}$ for $\{\{\}, \{\{\}\}\}$ is called the *simplified notation* for the minimal set representation. We now give the useful property of the minimal set representation.

Proposition 2.8 $n + 1 = n \cup \{n\}$

Proof:
This follows directly from Definition 2.31 by considering the set difference of the representations of n and $n-1$. □
The definition says that any set of the representations of consecutive natural numbers, starting at zero, is the representation of the next natural number. This permits us to conclude that the set of all natural numbers

$$\{0, 1, 2, \ldots\}$$

fits the definition of a natural number. Which natural number is it? It is easy to see, in the minimal set representation, that for natural numbers m and n, $m < n$ implies that the representation of m is a subset of the representation of n. Every finite natural number is a subset of the set of all natural numbers and so we conclude that $\{0, 1, 2, \ldots\}$ is an infinite natural number. The set notation thus permits us to construct an infinite number.

The set consisting of the representations of all finite natural numbers is an infinite natural number. The number has been given the name ω, the lower-case omega. In addition to being a letter omega traditionally also means "the last". The number ω comes after all the finite natural numbers. If we now apply Proposition 2.8 we see that

$$\omega \cup \{\omega\} = \omega + 1$$

This means that we can add one to an infinite number. Is the resulting number $\omega + 1$ a different number from ω? It turns out the answer is "'yes", because the representations of these numbers are different as sets. The representation of ω contains no infinite sets while the representation of $\omega + 1$ contains one.

Problems

Problem 2.69 *Find the representation for 5 using the curly-brace-and-comma notation.*

Problem 2.70 *Give the minimal set representation of $\omega + 2$ using the simplified notation.*

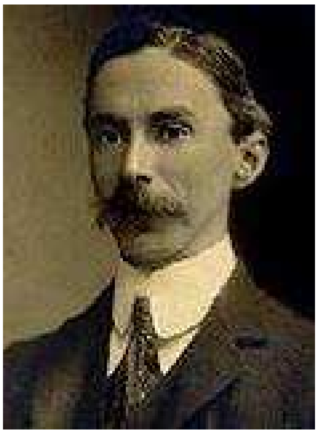
Problem 2.71 *Suppose that $n > m$ are natural numbers and that S is the minimal set representation of n while T is the minimal set representation of m. Is the representation of $n - m$ a member of the set difference $S - T$?*

Problem 2.72 *Give a formula, as a function of n, for the number of times that the symbol* **{** *appears in the representation of n.*

Problem 2.73 *Prove or disprove: there are an infinite number of distinct infinite numbers.*

Interlude

Russell's Paradox



Bertrand Arthur William Russell, 3rd Earl Russell, OM, FRS (18 May 1872-2 February 1970), commonly known as simply Bertrand Russell, was a British philosopher, logician, mathematician, historian, religious skeptic, social reformer, socialist and pacifist. Although he spent the majority of his life in England, he was born in Wales, where he also died.

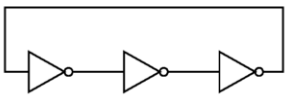
Let Q be the set of all sets that do not contain themselves as a member. Consider the question: "Does Q contain itself?" If the answer to this question is no then Q, by definition must contain itself. If, however, Q contains itself then it is by definition unable to contain itself. This rather annoying contradiction, constructed by Russell, had a rather amusing side effect.

Friedrich Frege had just finished the second of a three volume set of works called the *Basic Laws of Arithmetic* that was supposed to remove all intuition from mathematics and place it on a purely logical basis. Russell wrote Frege, explaining his paradox. Frege added an appendix to his second volume that attempted to avoid Russell's paradox. The third volume was never published.

It is possible to resolve Russell's paradox by being much more careful about what objects may be defined to be sets; the *category* of all sets that do not contain themselves gives rise to no contradiction (it does give rise to an entire field of mathematics, category theory). The key to resolving the paradox from a set theoretic perspective is that one cannot assume that, for every property, there is a set of all things satisfying that property. This is a reason why it is important that a set is properly defined. Another consequence of Russell's paradox is a warning that self-referential statements are both potentially interesting and fairly dangerous, at least on the intellectual plane.

The original phrasing of Russell's paradox was in terms of normal and abnormal sets. A set is *normal* if it fails to contain itself and abnormal otherwise. Consider the set of all normal sets. If this set is abnormal, it contains itself but by definition the set contains only normal sets and hence it is itself normal. The normality of this set forces the set to contain itself, which makes it abnormal. This is simply a rephrasing of the original contradiction.

Puzzle: what does the circuit below have to do with Russell's paradox and what use is it?



Preprint typeset in JHEP style - HYPER VERSION

January 2009

String Theory

University of Cambridge Part III Mathematical Tripos

Dr David Tong

Department of Applied Mathematics and Theoretical Physics,
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge, CB3 OWA, UK

http://www.damtp.cam.ac.uk/user/tong/string.html
d.tong@damtp.cam.ac.uk

Recommended Books and Resources

- J. Polchinski, *String Theory*

This two volume work is the standard introduction to the subject. Our lectures will more or less follow the path laid down in volume one covering the bosonic string. The book contains explanations and descriptions of many details that have been deliberately (and, I suspect, at times inadvertently) swept under a very large rug in these lectures. Volume two covers the superstring.

- M. Green, J. Schwarz and E. Witten, *Superstring Theory*

Another two volume set. It is now over 20 years old and takes a slightly old-fashioned route through the subject, with no explicit mention of conformal field theory. However, it does contain much good material and the explanations are uniformly excellent. Volume one is most relevant for these lectures.

- B. Zwiebach, *A First Course in String Theory*

This book grew out of a course given to undergraduates who had no previous exposure to general relativity or quantum field theory. It has wonderful pedagogical discussions of the basics of lightcone quantization. More surprisingly, it also has some very clear descriptions of several advanced topics, even though it misses out all the bits in between.

- P. Di Francesco, P. Mathieu and D. Sénéchal, *Conformal Field Theory*

This big yellow book is affectionately known as the yellow pages. It's a great way to learn conformal field theory. At first glance, it comes across as slightly daunting because it's big. (And yellow). But you soon realise that it's big because it starts at the beginning and provides detailed explanations at every step. The material necessary for this course can be found in chapters 5 and 6.

Further References: "*String Theory and M-Theory*" by Becker, Becker and Schwarz and "*String Theory in a Nutshell*" (it's a big nutshell) by Kiritsis both deal with the bosonic string fairly quickly, but include more advanced topics that may be of interest. The book "*D-Branes*" by Johnson has lively and clear discussions about the many joys of D-branes. Links to several excellent online resources, including video lectures by Shiraz Minwalla, are listed on the course webpage.

Contents

Acknowledgements

These lectures are aimed at beginning graduate students. They assume a working knowledge of quantum field theory and general relativity. The lectures were given over one semester and are based broadly on Volume one of the book by Joe Polchinski. I inherited the course from Michael Green whose notes were extremely useful. I also benefited enormously from the insightful and entertaining video lectures by Shiraz Minwalla.

I'm grateful to Anirban Basu, Niklas Beisert, Joe Bhaseen, Diego Correa, Nick Dorey, Michael Green, Anshuman Maharana, Malcolm Perry and Martin Schnabl for discussions and help with various aspects of these notes. I'm also grateful to the students, especially Carlos Guedes, for their excellent questions and superhuman typo-spotting abilities. Finally, my thanks to Alex Considine for infinite patience and understanding over the weeks these notes were written. I am supported by the Royal Society.

0. Introduction

String theory is an ambitious project. It purports to be an all-encompassing theory of the universe, unifying the forces of nature, including gravity, in a single quantum mechanical framework.

The premise of string theory is that, at the fundamental level, matter does not consist of point-particles but rather of tiny loops of string. From this slightly absurd beginning, the laws of physics emerge. General relativity, electromagnetism and Yang-Mills gauge theories all appear in a surprising fashion. However, they come with baggage. String theory gives rise to a host of other ingredients, most strikingly extra spatial dimensions of the universe beyond the three that we have observed. The purpose of this course is to understand these statements in detail.

These lectures differ from most other courses that you will take in a physics degree. String theory is speculative science. There is no experimental evidence that string theory is the correct description of our world and scant hope that hard evidence will arise in the near future. Moreover, string theory is very much a work in progress and certain aspects of the theory are far from understood. Unresolved issues abound and it seems likely that the final formulation has yet to be written. For these reasons, I'll begin this introduction by suggesting some answers to the question: Why study string theory?

Reason 1. String theory is a theory of quantum gravity

String theory unifies Einstein's theory of general relativity with quantum mechanics. Moreover, it does so in a manner that retains the explicit connection with both quantum theory and the low-energy description of spacetime.

But quantum gravity contains many puzzles, both technical and conceptual. What does spacetime look like at the shortest distance scales? How can we understand physics if the causal structure fluctuates quantum mechanically? Is the big bang truely the beginning of time? Do singularities that arise in black holes really signify the end of time? What is the microscopic origin of black hole entropy and what is it telling us? What is the resolution to the information paradox? Some of these issues will be reviewed later in this introduction.

Whether or not string theory is the true description of reality, it offers a framework in which one can begin to explore these issues. For some questions, string theory has given very impressive and compelling answers. For others, string theory has been almost silent.

Reason 2. String theory may be *the* theory of quantum gravity

With broad brush, string theory looks like an extremely good candidate to describe the real world. At low-energies it naturally gives rise to general relativity, gauge theories, scalar fields and chiral fermions. In other words, it contains all the ingredients that make up our universe. It also gives the only presently credible explanation for the value of the cosmological constant although, in fairness, I should add that the explanation is so distasteful to some that the community is rather amusingly split between whether this is a good thing or a bad thing. Moreover, string theory incorporates several ideas which do not yet have experimental evidence but which are considered to be likely candidates for physics beyond the standard model. Prime examples are supersymmetry and axions.

However, while the broad brush picture looks good, the finer details have yet to be painted. String theory does not provide unique predictions for low-energy physics but instead offers a bewildering array of possibilities, mostly dependent on what is hidden in those extra dimensions. Partly, this problem is inherent to any theory of quantum gravity: as we'll review shortly, it's a long way down from the Planck scale to the domestic energy scales explored at the LHC. Using quantum gravity to extract predictions for particle physics is akin to using QCD to extract predictions for how coffee makers work. But the mere fact that it's hard is little comfort if we're looking for convincing evidence that string theory describes the world in which we live.

While string theory cannot at present offer falsifiable predictions, it has nonetheless inspired new and imaginative proposals for solving outstanding problems in particle physics and cosmology. There are scenarios in which string theory might reveal itself in forthcoming experiments. Perhaps we'll find extra dimensions at the LHC, perhaps we'll see a network of fundamental strings stretched across the sky, or perhaps we'll detect some feature of non-Gaussianity in the CMB that is characteristic of D-branes at work during inflation. My personal feeling however is that each of these is a long shot and we may not know whether string theory is right or wrong within our lifetimes. Of course, the history of physics is littered with naysayers, wrongly suggesting that various theories will never be testable. With luck, I'll be one of them.

Reason 3. String theory provides new perspectives on gauge theories

String theory was born from attempts to understand the strong force. Almost forty years later, this remains one of the prime motivations for the subject. String theory provides tools with which to analyze down-to-earth aspects of quantum field theory that are far removed from high-falutin' ideas about gravity and black holes.

Of immediate relevance to this course are the pedagogical reasons to invest time in string theory. At heart, it is the study of conformal field theory and gauge symmetry. The techniques that we'll learn are not isolated to string theory, but apply to countless systems which have direct application to real world physics.

On a deeper level, string theory provides new and very surprising methods to understand aspects of quantum gauge theories. Of these, the most startling is the *AdS/CFT correspondence*, first conjectured by Juan Maldacena, which gives a relationship between strongly coupled quantum field theories and gravity in higher dimensions. These ideas have been applied in areas ranging from nuclear physics to condensed matter physics and have provided qualitative (and arguably quantitative) insights into strongly coupled phenomena.

Reason 4. String theory provides new results in mathematics

For the past 250 years, the close relationship between mathematics and physics has been almost a one-way street: physicists borrowed many things from mathematicians but, with a few noticeable exceptions, gave little back. In recent times, that has changed. Ideas and techniques from string theory and quantum field theory have been employed to give new "proofs" and, perhaps more importantly, suggest new directions and insights in mathematics. The most well known of these is *mirror symmetry*, a relationship between topologically different Calabi-Yau manifolds.

The four reasons described above also crudely characterize the string theory community: there are "relativists" and "phenomenologists" and "field theorists" and "mathematicians". Of course, the lines between these different sub-disciplines are not fixed and one of the great attractions of string theory is its ability to bring together people working in different areas — from cosmology to condensed matter to pure mathematics — and provide a framework in which they can profitably communicate. In my opinion, it is this cross-fertilization between fields which is the greatest strength of string theory.

0.1 Quantum Gravity

This is a starter course in string theory. Our focus will be on the perturbative approach to the bosonic string and, in particular, why this gives a consistent theory of quantum gravity. Before we leap into this, it is probably best to say a few words about quantum gravity itself. Like why it's hard. And why it's important. (And why it's not).

The Einstein Hilbert action is given by

$$S_{EH} = \frac{1}{16\pi G_N}\int d^4x\sqrt{-g}\mathcal{R}$$

Newton's constant G_N can be written as

$$8\pi G_N = \frac{\hbar c}{M_{pl}^2}$$

Throughout these lectures we work in units with $\hbar = c = 1$. The Planck mass M_{pl} defines an energy scale

$$M_{pl} \approx 2 \times 10^{18} \text{ GeV} .$$

(This is sometimes referred to as the reduced Planck mass, to distinguish it from the scale without the factor of 8π, namely $\sqrt{1/G_N} \approx 1 \times 10^{19}$ GeV).

There are a couple of simple lessons that we can already take from this. The first is that the relevant coupling in the quantum theory is $1/M_{pl}$. To see that this is indeed the case from the perspective of the action, we consider small perturbations around flat Minkowski space,

$$g_{\mu\nu} = \eta_{\mu\nu} + \frac{1}{M_{pl}} h_{\mu\nu}$$

The factor of $1/M_{pl}$ is there to ensure that when we expand out the Einstein-Hilbert action, the kinetic term for h is canonically normalized, meaning that it comes with no powers of M_{pl}. This then gives the kind of theory that you met in your first course on quantum field theory, albeit with an infinite series of interaction terms,

$$S_{EH} = \int d^4x \; (\partial h)^2 + \frac{1}{M_{pl}} \, h \, (\partial h)^2 + \frac{1}{M_{pl}^2} \, h^2 \, (\partial h)^2 + \ldots$$

Each of these terms is schematic: if you were to do this explicitly, you would find a mess of indices contracted in different ways. We see that the interactions are suppressed by powers of M_{pl}. This means that quantum perturbation theory is an expansion in the dimensionless ratio E^2/M_{pl}^2, where E is the energy associated to the process of interest. We learn that gravity is weak, and therefore under control, at low-energies. But gravitational interactions become strong as the energy involved approaches the Planck scale. In the language of the renormalization group, couplings of this type are known as *irrelevant*.

The second lesson to take away is that the Planck scale M_{pl} is very very large. The LHC will probe the electroweak scale, $M_{EW} \sim 10^3$ GeV. The ratio is $M_{EW}/M_{pl} \sim 10^{-15}$. For this reason, quantum gravity will not affect your daily life, even if your daily life involves the study of the most extreme observable conditions in the universe.

Gravity is Non-Renormalizable

Quantum field theories with irrelevant couplings are typically ill-behaved at high-energies, rendering the theory ill-defined. Gravity is no exception. Theories of this type are called *non-renormalizable*, which means that the divergences that appear in the Feynman diagram expansion cannot be absorbed by a finite number of counterterms. In pure Einstein gravity, the symmetries of the theory are enough to ensure that the one-loop S-matrix is finite. The first divergence occurs at two-loops and requires the introduction of a counterterm of the form,

$$\Gamma \sim \frac{1}{\epsilon}\frac{1}{M_{pl}^4}\int d^4x\sqrt{-g}\,\mathcal{R}^{\mu\nu}{}_{\rho\sigma}\mathcal{R}^{\rho\sigma}{}_{\lambda\kappa}\mathcal{R}^{\lambda\kappa}{}_{\mu\nu}$$

with $\epsilon = 4 - D$. All indications point towards the fact that this is the first in an infinite number of necessary counterterms.

Coupling gravity to matter requires an interaction term of the form,

$$S_{int} = \int d^4x\,\frac{1}{M_{pl}}\,h_{\mu\nu}T^{\mu\nu} + \mathcal{O}(h^2)$$

This makes the situation marginally worse, with the first divergence now appearing at one-loop. The Feynman diagram in the figure shows particle scattering through the exchange of two gravitons. When the momentum k running in the loop is large, the diagram is badly divergent: it scales as

$$\frac{1}{M_{pl}^4}\int^{\infty} d^4k$$

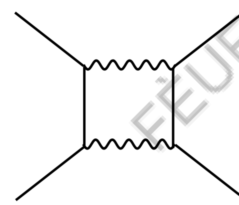


Figure 1:

Non-renormalizable theories are commonplace in the history of physics, the most commonly cited example being Fermi's theory of the weak interaction. The first thing to say about them is that they are far from useless! Non-renormalizable theories are typically viewed as *effective* field theories, valid only up to some energy scale Λ. One deals with the divergences by simply admitting ignorance beyond this scale and treating Λ as a UV cut-off on any momentum integral. In this way, we get results which are valid to an accuracy of E/Λ (perhaps raised to some power). In the case of the weak interaction, Fermi's theory accurately predicts physics up to an energy scale of $\sqrt{1/G_F} \sim 100$ GeV. In the case of quantum gravity, Einstein's theory works to an accuracy of $(E/M_{pl})^2$.

However, non-renormalizable theories are typically unable to describe physics at their cut-off scale Λ or beyond. This is because they are missing the true ultra-violet degrees of freedom which tame the high-energy behaviour. In the case of the weak force, these new degrees of freedom are the W and Z bosons. We would like to know what missing degrees of freedom are needed to complete gravity.

Singularities

Only a particle physicist would phrase all questions about the universe in terms of scattering amplitudes. In general relativity we typically think about the geometry as a whole, rather than bastardizing the Einstein-Hilbert action and discussing perturbations around flat space. In this language, the question of high-energy physics turns into one of short distance physics. Classical general relativity is not to be trusted in regions where the curvature of spacetime approaches the Planck scale and ultimately becomes singular. A quantum theory of gravity should resolve these singularities.

The question of spacetime singularities is morally equivalent to that of high-energy scattering. Both probe the ultra-violet nature of gravity. A spacetime geometry is made of a coherent collection of gravitons, just as the electric and magnetic fields in a laser are made from a collection of photons. The short distance structure of spacetime is governed – after Fourier transform – by high momentum gravitons. Understanding spacetime singularities and high-energy scattering are different sides of the same coin.

There are two situations in general relativity where singularity theorems tell us that the curvature of spacetime gets large: at the big bang and in the center of a black hole. These provide two of the biggest challenges to any putative theory of quantum gravity.

Gravity is Subtle

It is often said that general relativity contains the seeds of its own destruction. The theory is unable to predict physics at the Planck scale and freely admits to it. Problems such as non-renormalizability and singularities are, in a Rumsfeldian sense, known unknowns. However, the full story is more complicated and subtle. On the one hand, the issue of non-renormalizability may not quite be the crisis that it first appears. On the other hand, some aspects of quantum gravity suggest that general relativity isn't as honest about its own failings as is usually advertised. The theory hosts a number of unknown unknowns, things that we didn't even know that we didn't know. We won't have a whole lot to say about these issues in this course, but you should be aware of them. Here I mention only a few salient points.

Firstly, there is a key difference between Fermi's theory of the weak interaction and gravity. Fermi's theory was unable to provide predictions for any scattering process at energies above $\sqrt{1/G_F}$. In contrast, if we scatter two objects at extremely high-energies in gravity — say, at energies $E \gg M_{pl}$ — then we know exactly what will happen: we form a big black hole. We don't need quantum gravity to tell us this. Classical general relativity is sufficient. If we restrict attention to scattering, the crisis of non-renormalizability is not problematic at ultra-high energies. It's troublesome only within a window of energies around the Planck scale.

Similar caveats hold for singularities. If you are foolish enough to jump into a black hole, then you're on your own: without a theory of quantum gravity, no one can tell you what fate lies in store at the singularity. Yet, if you are smart and stay outside of the black hole, you'll be hard pushed to see any effects of quantum gravity. This is because Nature has conspired to hide Planck scale curvatures from our inquisitive eyes. In the case of black holes this is achieved through cosmic censorship which is a conjecture in classical general relativity that says singularities are hidden behind horizons. In the case of the big bang, it is achieved through inflation, washing away any traces from the very early universe. Nature appears to shield us from the effects of quantum gravity, whether in high-energy scattering or in singularities. I think it's fair to say that no one knows if this conspiracy is pointing at something deep, or is merely inconvenient for scientists trying to probe the Planck scale.

While horizons may protect us from the worst excesses of singularities, they come with problems of their own. These are the unknown unknowns: difficulties that arise when curvatures are small and general relativity says "trust me". The entropy of black holes and the associated paradox of information loss strongly suggest that local quantum field theory breaks down at macroscopic distance scales. Attempts to formulate quantum gravity in de Sitter space, or in the presence of eternal inflation, hint at similar difficulties. Ideas of holography, black hole complimentarity and the AdS/CFT correspondence all point towards non-local effects and the emergence of spacetime. These are the deep puzzles of quantum gravity and their relationship to the ultra-violet properties of gravity is unclear.

As a final thought, let me mention the one observation that has an outside chance of being related to quantum gravity: the cosmological constant. With an energy scale of $\Lambda \sim 10^{-3}$ eV it appears to have little to do with ultra-violet physics. If it does have its origins in a theory of quantum gravity, it must either be due to some subtle "unknown unknown", or because it is explained away as an environmental quantity as in string theory.

Is the Time Ripe?

Our current understanding of physics, embodied in the standard model, is valid up to energy scales of 10^3 GeV. This is 15 orders of magnitude away from the Planck scale. Why do we think the time is now ripe to tackle quantum gravity? Surely we are like the ancient Greeks arguing about atomism. Why on earth do we believe that we've developed the right tools to even address the question?

The honest answer, I think, is hubris.

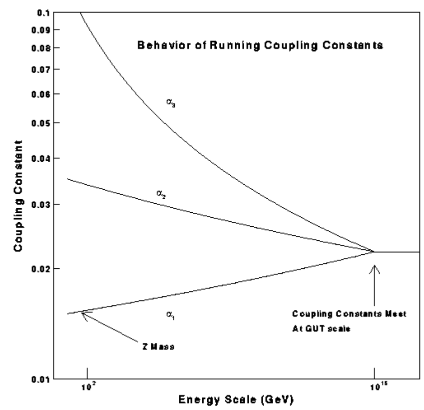


Figure 2:

However, there is mild circumstantial evidence that the framework of quantum field theory might hold all the way to the Planck scale without anything very dramatic happening in between. The main argument is unification. The three coupling constants of Nature run logarithmically, meeting miraculously at the GUT energy scale of 10^{15} GeV. Just slightly later, the fourth force of Nature, gravity, joins them. While not overwhelming, this does provide a hint that perhaps quantum field theory can be taken seriously at these ridiculous scales. Historically I suspect this was what convinced large parts of the community that it was ok to speak about processes at 10^{18} GeV.

Finally, perhaps the most compelling argument for studying physics at the Planck scale is that string theory *does* provide a consistent unified quantum theory of gravity and the other forces. Given that we have this theory sitting in our laps, it would be foolish not to explore its consequences. The purpose of these lecture notes is to begin this journey.

1. The Relativistic String

All lecture courses on string theory start with a discussion of the point particle. Ours is no exception. We'll take a flying tour through the physics of the relativistic point particle and extract a couple of important lessons that we'll take with us as we move onto string theory.

1.1 The Relativistic Point Particle

We want to write down the Lagrangian describing a relativistic particle of mass m. In anticipation of string theory, we'll consider D-dimensional Minkowski space $\mathbf{R}^{1,D-1}$. Throughout these notes, we work with signature

$$\eta_{\mu\nu} = \mathrm{diag}(-1, +1, +1, \ldots, +1)$$

Note that this is the opposite signature to my quantum field theory notes.

If we fix a frame with coordinates $X^\mu = (t, \vec{x})$ the action is simple:

$$S = -m \int dt \sqrt{1 - \dot{\vec{x}} \cdot \dot{\vec{x}}} \, . \tag{1.1}$$

To see that this is correct we can compute the momentum $\vec{p}$, conjugate to $\vec{x}$, and the energy E which is equal to the Hamiltonian,

$$\vec{p} = \frac{m\dot{\vec{x}}}{\sqrt{1 - \dot{\vec{x}} \cdot \dot{\vec{x}}}} \quad , \quad E = \sqrt{m^2 + \vec{p}^{\,2}} \, ,$$

both of which should be familiar from courses on special relativity.

Although the Lagrangian (1.1) is correct, it's not fully satisfactory. The reason is that time t and space $\vec{x}$ play very different roles in this Lagrangian. The position $\vec{x}$ is a dynamical degree of freedom. In contrast, time t is merely a parameter providing a label for the position. Yet Lorentz transformations are supposed to mix up t and $\vec{x}$ and such symmetries are not completely obvious in (1.1). Can we find a new Lagrangian in which time and space are on equal footing?

One possibility is to treat both time and space as labels. This leads us to the concept of field theory. However, in this course we will be more interested in the other possibility: we will promote time to a dynamical degree of freedom. At first glance, this may appear odd: the number of degrees of freedom is one of the crudest ways we have to characterize a system. We shouldn't be able to add more degrees of freedom at will without fundamentally changing the system that we're talking about. Another way of saying this is that the particle has the option to move in space, but it doesn't have the option to move in time. It *has* to move in time. So we somehow need a way to promote time to a degree of freedom without it really being a true dynamical degree of freedom! How do we do this? The answer, as we will now show, is gauge symmetry.

Consider the action,

$$S = -m \int d\tau \, \sqrt{-\dot{X}^\mu \dot{X}^\nu \eta_{\mu\nu}} \; , \tag{1.2}$$

where $\mu = 0, \ldots, D-1$ and $\dot{X}^\mu = dX^\mu/d\tau$. We've introduced a new parameter τ which labels the position along the worldline of the particle as shown by the dashed lines in the figure. This action has a simple interpretation: it is just the proper time $\int ds$ along the worldline.

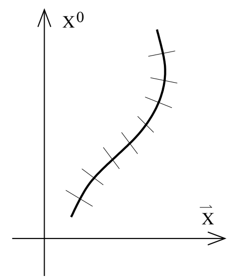


Figure 3:

Naively it looks as if we now have D physical degrees of freedom rather than $D-1$ because, as promised, the time direction $X^0 \equiv t$ is among our dynamical variables: $X^0 = X^0(\tau)$. However, this is an illusion. To see why, we need to note that the action (1.2) has a very important property: reparameterization invariance. This means that we can pick a different parameter $\tilde{\tau}$ on the worldline, related to τ by any monotonic function

$$\tilde{\tau} = \tilde{\tau}(\tau) \; .$$

Let's check that the action is invariant under transformations of this type. The integration measure in the action changes as $d\tau = d\tilde{\tau}\, |d\tau/d\tilde{\tau}|$. Meanwhile, the velocities change as $dX^\mu/d\tau = (dX^\mu/d\tilde{\tau})\,(d\tilde{\tau}/d\tau)$. Putting this together, we see that the action can just as well be written in the $\tilde{\tau}$ reparameterization,

$$S = -m \int d\tilde{\tau} \, \sqrt{-\frac{dX^\mu}{d\tilde{\tau}} \frac{dX^\nu}{d\tilde{\tau}} \eta_{\mu\nu}} \; .$$

The upshot of this is that not all D degrees of freedom X^μ are physical. For example, suppose you find a solution to this system, so that you know how X^0 changes with τ and how X^1 changes with τ and so on. Not all of that information is meaningful because τ itself is not meaningful. In particular, we could use our reparameterization invariance to simply set

$$\tau = X^0(\tau) \equiv t \tag{1.3}$$

If we plug this choice into the action (1.2) then we recover our initial action (1.1). The reparameterization invariance is a *gauge symmetry* of the system. Like all gauge symmetries, it's not really a symmetry at all. Rather, it is a redundancy in our description. In the present case, it means that although we seem to have D degrees of freedom X^μ, one of them is fake.

The fact that one of the degrees of freedom is a fake also shows up if we look at the momenta,

$$p_\mu = \frac{\partial L}{\partial \dot{X}^\mu} = \frac{m\dot{X}^\nu \eta_{\mu\nu}}{\sqrt{-\dot{X}^\lambda \dot{X}^\rho\, \eta_{\lambda\rho}}} \tag{1.4}$$

These momenta aren't all independent. They satisfy

$$p_\mu p^\mu + m^2 = 0 \tag{1.5}$$

This is a constraint on the system. It is, of course, the mass-shell constraint for a relativistic particle of mass m. From the worldline perspective, it tells us that the particle isn't allowed to sit still in Minkowski space: at the very least, it had better keep moving in a timelike direction with $(p^0)^2 \geq m^2$.

One advantage of the action (1.2) is that the Poincaré symmetry of the particle is now manifest, appearing as a global symmetry on the worldline

$$X^\mu \rightarrow \Lambda^\mu{}_\nu X^\nu + c^\mu \tag{1.6}$$

where Λ is a Lorentz transformation satisfying $\Lambda^\mu{}_\nu \eta^{\nu\rho} \Lambda^\sigma{}_\rho = \eta^{\mu\sigma}$, while c^μ corresponds to a constant translation. We have made all the symmetries manifest at the price of introducing a gauge symmetry into our system. A similar gauge symmetry will arise in the relativistic string and much of this course will be devoted to understanding its consequences.

1.1.1 Quantization

It's a trivial matter to quantize this action. We introduce a wavefunction $\Psi(X)$. This satisfies the usual Schrödinger equation,

$$i\,\frac{\partial \Psi}{\partial \tau} = H\Psi\ .$$

But, computing the Hamiltonian $H = \dot{X}^\mu p_\mu - L$, we find that it vanishes: $H = 0$. This shouldn't be surprising. It is simply telling us that the wavefunction doesn't depend on

τ. Since the wavefunction is something physical while, as we have seen, τ is not, this is to be expected. Note that this doesn't mean that time has dropped out of the problem. On the contrary, in this relativistic context, time X^0 is an operator, just like the spatial coordinates $\vec{x}$. This means that the wavefunction Ψ is immediately a function of space and time. It is not like a static state in quantum mechanics, but more akin to the fully integrated solution to the non-relativistic Schrödinger equation.

The classical system has a constraint given by (1.5). In the quantum theory, we impose this constraint as an operator equation on the wavefunction, namely $(p^\mu p_\mu + m^2)\Psi = 0$. Using the usual representation of the momentum operator $p_\mu = -i\partial/\partial X^\mu$, we recognize this constraint as the Klein-Gordon equation

$$\left(-\frac{\partial}{\partial X^\mu}\frac{\partial}{\partial X^\nu}\eta^{\mu\nu} + m^2\right)\Psi(X) = 0 \tag{1.7}$$

Although this equation is familiar from field theory, it's important to realize that the interpretation is somewhat different. In relativistic field theory, the Klein-Gordon equation is the equation of motion obeyed by a scalar field. In relativistic quantum mechanics, it is the equation obeyed by the wavefunction. In the early days of field theory, the fact that these two equations are the same led people to think one should view the wavefunction as a classical field and quantize it a second time. This isn't correct, but nonetheless the language has stuck and it is common to talk about the point particle perspective as "first quantization" and the field theory perspective as "second quantization".

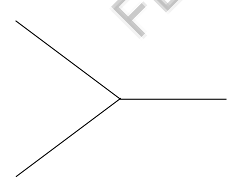


Figure 4:

So far we've considered only a free point particle. How can we introduce interactions into this framework? We would have to first decide which interactions are allowed: perhaps the particle can split into two; perhaps it can fuse with other particles? Obviously, there is a huge range of options for us to choose from. We would then assign amplitudes for these processes to happen. There would be certain restrictions coming from the requirement of unitarity which, among other things, would lead to the necessity of anti-particles. We could draw diagrams associated to the different interactions — an example is given in the figure — and in this manner we would slowly build up the Feynman diagram expansion that is familiar from field theory. In fact, this was pretty much the way Feynman himself approached the topic of QED. However, in practice we rarely construct particle interactions in this way because the field theory framework provides a much better way of looking at things. In contrast, this way of building up interactions is exactly what we will later do for strings.

1.1.2 Ein Einbein

There is another action that describes the relativistic point particle. We introduce yet another field on the worldline, $e(\tau)$, and write

$$S = \frac{1}{2}\int d\tau \, \left(e^{-1}\dot{X}^2 - em^2\right) \;, \tag{1.8}$$

where we've used the notation $\dot{X}^2 = \dot{X}^\mu \dot{X}^\nu \eta_{\mu\nu}$. For the rest of these lectures, terms like X^2 will always mean an implicit contraction with the spacetime Minkowski metric.

This form of the action makes it look as if we have coupled the worldline theory to 1d gravity, with the field $e(\tau)$ acting as an einbein (in the sense of vierbeins that are introduced in general relativity). To see this, note that we could change notation and write this action in the more suggestive form

$$S = -\frac{1}{2}\int d\tau \, \sqrt{-g_{\tau\tau}}\left(g^{\tau\tau}\dot{X}^2 + m^2\right) \;. \tag{1.9}$$

where $g_{\tau\tau} = (g^{\tau\tau})^{-1}$ is the metric on the worldline and $e = \sqrt{-g_{\tau\tau}}$

Although our action appears to have one more degree of freedom, e, it can be easily checked that it has the same equations of motion as (1.2). The reason for this is that e is completely fixed by its equation of motion, $\dot{X}^2 + e^2 m^2 = 0$. Substituting this into the action (1.8) recovers (1.2)

The action (1.8) has a couple of advantages over (1.2). Firstly, it works for massless particles with $m = 0$. Secondly, the absence of the annoying square root means that it's easier to quantize in a path integral framework.

The action (1.8) retains invariance under reparameterizations which are now written in a form that looks more like general relativity. For transformations parameterized by an infinitesimal η, we have

$$\tau \to \tilde{\tau} = \tau - \eta(\tau) \quad , \quad \delta e = \frac{d}{d\tau}(\eta(\tau)e) \quad , \quad \delta X^\mu = \frac{dX^\mu}{d\tau}\eta(\tau) \tag{1.10}$$

The einbein e transforms as a density on the worldline, while each of the coordinates X^μ transforms as a worldline scalar.

1.2 The Nambu-Goto Action

A particle sweeps out a worldline in Minkowski space. A string sweeps out a *worldsheet*. We'll parameterize this worldsheet by one timelike coordinate τ, and one spacelike coordinate σ. In this section we'll focus on closed strings and take σ to be periodic, with range

$$\sigma \in [0, 2\pi) \; . \tag{1.11}$$

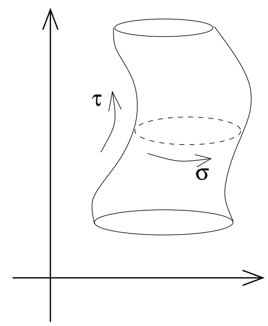


Figure 5:

We will sometimes package the two worldsheet coordinates together as $\sigma^\alpha = (\tau, \sigma)$, $\alpha = 0, 1$. Then the string sweeps out a surface in spacetime which defines a map from the worldsheet to Minkowski space, $X^\mu(\sigma, \tau)$ with $\mu = 0, \ldots, D-1$. For closed strings, we require

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

In this context, spacetime is sometimes referred to as the *target space* to distinguish it from the worldsheet.

We need an action that describes the dynamics of this string. The key property that we will ask for is that nothing depends on the coordinates σ^α that we choose on the worldsheet. In other words, the string action should be reparameterization invariant. What kind of action does the trick? Well, for the point particle the action was proportional to the length of the worldline. The obvious generalization is that the action for the string should be proportional to the area, A, of the worldsheet. This is certainly a property that is characteristic of the worldsheet itself, rather than any choice of parameterization.

How do we find the area A in terms of the coordinates $X^\mu(\sigma, \tau)$? The worldsheet is a curved surface embedded in spacetime. The induced metric, $\gamma_{\alpha\beta}$, on this surface is the pull-back of the flat metric on Minkowski space,

$$\gamma_{\alpha\beta} = \frac{\partial X^\mu}{\partial \sigma^\alpha} \frac{\partial X^\nu}{\partial \sigma^\beta} \eta_{\mu\nu} \; . \tag{1.12}$$

Then the action which is proportional to the area of the worldsheet is given by,

$$S = -T \int d^2\sigma \; \sqrt{-\det \gamma} \; . \tag{1.13}$$

Here T is a constant of proportionality. We will see shortly that it is the *tension* of the string, meaning the mass per unit length.

We can write this action a little more explicitly. The pull-back of the metric is given by,

$$\gamma_{\alpha\beta} = \begin{pmatrix} \dot{X}^2 & \dot{X}\cdot X' \\ \dot{X}\cdot X' & X'^2 \end{pmatrix} .$$

where $\dot{X}^\mu = \partial X^\mu/\partial\tau$ and $X^{\mu\prime} = \partial X^\mu/\partial\sigma$. The action then takes the form,

$$S = -T\int d^2\sigma\sqrt{-(\dot{X})^2\,(X')^2 + (\dot{X}\cdot X')^2} \; . \tag{1.14}$$

This is the *Nambu-Goto* action for a relativistic string.

Action = Area: A Check

If you're unfamiliar with differential geometry, the argument about the pull-back of the metric may be a bit slick. Thankfully, there's a more pedestrian way to see that the action (1.14) is equal to the area swept out by the worldsheet. It's slightly simpler to make this argument for a surface embedded in Euclidean space rather than Minkowski space. We choose some parameterization of the sheet in terms of τ and σ, as drawn in the figure, and we write the coordinates of Euclidean space as $\vec{X}(\sigma,\tau)$. We'll compute the area of the infinitesimal shaded region. The vectors tangent to the boundary are,

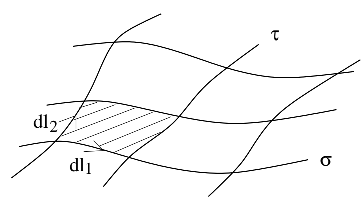


Figure 6:

$$\vec{dl}_1 = \frac{\partial\vec{X}}{\partial\sigma} \quad , \quad \vec{dl}_2 = \frac{\partial\vec{X}}{\partial\tau} \; .$$

If the angle between these two vectors is θ, then the area is then given by

$$ds^2 = |\vec{dl}_1||\vec{dl}_2|\sin\theta = \sqrt{dl_1^2\,dl_2^2(1-\cos^2\theta)} = \sqrt{dl_1^2\,dl_2^2 - (\vec{dl}_1\cdot\vec{dl}_2)^2} \tag{1.15}$$

which indeed takes the form of the integrand of (1.14).

Tension and Dimension

Let's now see that T has the physical interpretation of tension. We write Minkowski coordinates as $X^\mu = (t,\vec{x})$. We work in a gauge with $X^0 \equiv t = R\tau$, where R is a constant that is needed to balance up dimensions (see below) and will drop out at the end of the argument. Consider a snapshot of a string configuration at a time when

$d\vec{x}/d\tau = 0$ so that the instantaneous kinetic energy vanishes. Evaluating the action for a time dt gives

$$S = -T \int d\tau d\sigma R \sqrt{(d\vec{x}/d\sigma)^2} = -T \int dt \,(\text{spatial length of string}) \ . \tag{1.16}$$

But, when the kinetic energy vanishes, the action is proportional to the time integral of the potential energy,

$$\text{potential energy} = T \times (\text{spatial length of string}) \ .$$

So T is indeed the energy per unit length as claimed. We learn that the string acts rather like an elastic band and its energy increases linearly with length. (This is different from the elastic bands you're used to which obey Hooke's law where energy increased quadratically with length). To minimize its potential energy, the string will want to shrink to zero size. We'll see that when we include quantum effects this can't happen because of the usual zero point energies.

There is a slightly annoying way of writing the tension that has its origin in ancient history, but is commonly used today

$$T = \frac{1}{2\pi\alpha'} \tag{1.17}$$

where α' is pronounced "alpha-prime". In the language of our ancestors, α' is referred to as the "universal Regge slope". We'll explain why later in this course.

At this point, it's worth pointing out some conventions that we have, until now, left implicit. The spacetime coordinates have dimension $[X] = -1$. In contrast, the worldsheet coordinates are taken to be dimensionless, $[\sigma] = 0$. (This can be seen in our identification $\sigma \equiv \sigma + 2\pi$). The tension is equal to the mass per unit length and has dimension $[T] = 2$. Obviously this means that $[\alpha'] = -2$. We can therefore associate a length scale, l_s, by

$$\alpha' = l_s^2 \tag{1.18}$$

The *string scale* l_s is the natural length that appears in string theory. In fact, in a certain sense (that we will make more precise later in the course) this length scale is the only parameter of the theory.

Actual Strings vs. Fundamental Strings

There are several situations in Nature where string-like objects arise. Prime examples include magnetic flux tubes in superconductors and chromo-electric flux tubes in QCD. Cosmic strings, a popular speculation in cosmology, are similar objects, stretched across the sky. In each of these situations, there are typically two length scales associated to the string: the tension, T and the width of the string, L. For all these objects, the dynamics is governed by the Nambu-Goto action as long as the curvature of the string is much greater than L. (In the case of superconductors, one should work with a suitable non-relativistic version of the Nambu-Goto action).

However, in each of these other cases, the Nambu-Goto action is not the end of the story. There will typically be additional terms in the action that depend on the width of the string. The form of these terms is not universal, but often includes a *rigidity* piece of form $L\int K^2$, where K is the extrinsic curvature of the worldsheet. Other terms could be added to describe fluctuations in the width of the string.

The string scale, l_s, or equivalently the tension, T, depends on the kind of string that we're considering. For example, if we're interested in QCD flux tubes then we would take

$$T \sim (1 \text{ GeV})^2 \tag{1.19}$$

In this course we will consider *fundamental strings* which have zero width. What this means in practice is that we take the Nambu-Goto action as the complete description for all configurations of the string. These strings will have relevance to quantum gravity and the tension of the string is taken to be much larger, typically an order of magnitude or so below the Planck scale.

$$T \lesssim M_{pl}^2 = (10^{18} \text{ GeV})^2 \tag{1.20}$$

However, I should point out that when we try to view string theory as a fundamental theory of quantum gravity, we don't really know what value T should take. As we will see later in this course, it depends on many other aspects, most notably the string coupling and the volume of the extra dimensions.

1.2.1 Symmetries of the Nambu-Goto Action

The Nambu-Goto action has two types of symmetry, each of a different nature.

- Poincaré invariance of the spacetime (1.6). This is a global symmetry from the perspective of the worldsheet, meaning that the parameters $\Lambda^\mu{}_\nu$ and c^μ which label the symmetry transformation are constants and do not depend on worldsheet coordinates σ^α.

- Reparameterization invariance, $\sigma^\alpha \to \tilde{\sigma}^\alpha(\sigma)$. As for the point particle, this is a gauge symmetry. It reflects the fact that we have a redundancy in our description because the worldsheet coordinates σ^α have no physical meaning.

1.2.2 Equations of Motion

To derive the equations of motion for the Nambu-Goto string, we first introduce the momenta which we call Π because there will be countless other quantities that we want to call p later,

$$\Pi^\tau_\mu = \frac{\partial \mathcal{L}}{\partial \dot{X}^\mu} = -T \frac{(\dot{X}\cdot X')X'_\mu - (X'^2)\dot{X}_\mu}{\sqrt{(\dot{X}\cdot X')^2 - \dot{X}^2\, X'^2}} \ .$$

$$\Pi^\sigma_\mu = \frac{\partial \mathcal{L}}{\partial X'^\mu} = -T \frac{(\dot{X}\cdot X')\dot{X}_\mu - (\dot{X}^2)X'_\mu}{\sqrt{(\dot{X}\cdot X')^2 - \dot{X}^2\, X'^2}} \ .$$

The equations of motion are then given by,

$$\frac{\partial \Pi^\tau_\mu}{\partial \tau} + \frac{\partial \Pi^\sigma_\mu}{\partial \sigma} = 0$$

These look like nasty, non-linear equations. In fact, there's a slightly nicer way to write these equations, starting from the earlier action (1.13). Recall that the variation of a determinant is $\delta\sqrt{-\gamma} = \frac{1}{2}\sqrt{-\gamma}\,\gamma^{\alpha\beta}\delta\gamma_{\alpha\beta}$. Using the definition of the pull-back metric $\gamma_{\alpha\beta}$, this gives rise to the equations of motion

$$\partial_\alpha(\sqrt{-\det\gamma}\,\gamma^{\alpha\beta}\partial_\beta X^\mu) = 0 \ , \tag{1.21}$$

Although this notation makes the equations look a little nicer, we're kidding ourselves. Written in terms of X^μ, they are still the same equations. Still nasty.

1.3 The Polyakov Action

The square-root in the Nambu-Goto action means that it's rather difficult to quantize using path integral techniques. However, there is another form of the string action which is classically equivalent to the Nambu-Goto action. It eliminates the square root at the expense of introducing another field,

$$S = -\frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{-g}g^{\alpha\beta}\,\partial_\alpha X^\mu \partial_\beta X^\nu\,\eta_{\mu\nu} \tag{1.22}$$

where $g \equiv \det g$. This is the *Polyakov* action. (Polyakov didn't discover the action, but he understood how to work with it in the path integral and for this reason it carries his name. The path integral treatment of this action will be the subject of Chapter 5).

The new field is $g_{\alpha\beta}$. It is a dynamical metric on the worldsheet. From the perspective of the worldsheet, the Polyakov action is a bunch of scalar fields X coupled to 2d gravity.

The equation of motion for X^μ is

$$\partial_\alpha(\sqrt{-g}g^{\alpha\beta}\partial_\beta X^\mu) = 0 \ , \tag{1.23}$$

which coincides with the equation of motion (1.21) from the Nambu-Goto action, except that $g_{\alpha\beta}$ is now an independent variable which is fixed by its own equation of motion. To determine this, we vary the action (remembering again that $\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}g_{\alpha\beta}\delta g^{\alpha\beta} = +\frac{1}{2}\sqrt{-g}g^{\alpha\beta}\delta g_{\alpha\beta}$),

$$\delta S = -\frac{T}{2}\int d^2\sigma \ \delta g^{\alpha\beta}\left(\sqrt{-g}\,\partial_\alpha X^\mu\partial_\beta X^\nu - \tfrac{1}{2}\sqrt{-g}\,g_{\alpha\beta}g^{\rho\sigma}\partial_\rho X^\mu\partial_\sigma X^\nu\right)\eta_{\mu\nu} = 0 \ . \tag{1.24}$$

The worldsheet metric is therefore given by,

$$g_{\alpha\beta} = 2f(\sigma)\,\partial_\alpha X\cdot\partial_\beta X \ , \tag{1.25}$$

where the function $f(\sigma)$ is given by,

$$f^{-1} = g^{\rho\sigma}\,\partial_\rho X\cdot\partial_\sigma X$$

A comment on the potentially ambiguous notation: here, and below, any function $f(\sigma)$ is always short-hand for $f(\sigma,\tau)$: it in no way implies that f depends only on the spatial worldsheet coordinate.

We see that $g_{\alpha\beta}$ isn't quite the same as the pull-back metric $\gamma_{\alpha\beta}$ defined in equation (1.12); the two differ by the conformal factor f. However, this doesn't matter because, rather remarkably, f drops out of the equation of motion (1.23). This is because the $\sqrt{-g}$ term scales as f, while the inverse metric $g^{\alpha\beta}$ scales as f^{-1} and the two pieces cancel. We therefore see that Nambu-Goto and the Polyakov actions result in the same equation of motion for X.

In fact, we can see more directly that the Nambu-Goto and Polyakov actions coincide. We may replace $g_{\alpha\beta}$ in the Polyakov action (1.22) with its equation of motion $g_{\alpha\beta} = 2f\,\gamma_{\alpha\beta}$. The factor of f also drops out of the action for the same reason that it dropped out of the equation of motion. In this manner, we recover the Nambu-Goto action (1.13).

1.3.1 Symmetries of the Polyakov Action

The fact that the presence of the factor $f(\sigma,\tau)$ in (1.25) didn't actually affect the equations of motion for X^μ reflects the existence of an extra symmetry which the Polyakov action enjoys. Let's look more closely at this. Firstly, the Polyakov action still has the two symmetries of the Nambu-Goto action,

- Poincaré invariance. This is a global symmetry on the worldsheet.

$$X^\mu \to \Lambda^\mu{}_\nu X^\nu + c^\mu \ .$$

- Reparameterization invariance, also known as diffeomorphisms. This is a gauge symmetry on the worldsheet. We may redefine the worldsheet coordinates as $\sigma^\alpha \to \tilde{\sigma}^\alpha(\sigma)$. The fields X^μ transform as worldsheet scalars, while $g_{\alpha\beta}$ transforms in the manner appropriate for a 2d metric.

$$\begin{aligned} X^\mu(\sigma) &\to \tilde{X}^\mu(\tilde{\sigma}) = X^\mu(\sigma) \\ g_{\alpha\beta}(\sigma) &\to \tilde{g}_{\alpha\beta}(\tilde{\sigma}) = \frac{\partial\sigma^\gamma}{\partial\tilde{\sigma}^\alpha}\frac{\partial\sigma^\delta}{\partial\tilde{\sigma}^\beta}\, g_{\gamma\delta}(\sigma) \end{aligned}$$

It will sometimes be useful to work infinitesimally. If we make the coordinate change $\sigma^\alpha \to \tilde{\sigma}^\alpha = \sigma^\alpha - \eta^\alpha(\sigma)$, for some small η. The transformations of the fields then become,

$$\begin{aligned} \delta X^\mu(\sigma) &= \eta^\alpha \partial_\alpha X^\mu \\ \delta g_{\alpha\beta}(\sigma) &= \nabla_\alpha \eta_\beta + \nabla_\beta \eta_\alpha \end{aligned}$$

where the covariant derivative is defined by $\nabla_\alpha \eta_\beta = \partial_\alpha \eta_\beta - \Gamma^\sigma_{\alpha\beta}\eta_\sigma$ with the Levi-Civita connection associated to the worldsheet metric given by the usual expression,

$$\Gamma^\sigma_{\alpha\beta} = \tfrac{1}{2}\, g^{\sigma\rho}(\partial_\alpha g_{\beta\rho} + \partial_\beta g_{\rho\alpha} - \partial_\rho g_{\alpha\beta})$$

Together with these familiar symmetries, there is also a new symmetry which is novel to the Polyakov action. It is called *Weyl invariance*.

- Weyl Invariance. Under this symmetry, $X^\mu(\sigma) \to X^\mu(\sigma)$, while the metric changes as

$$g_{\alpha\beta}(\sigma) \to \Omega^2(\sigma)\, g_{\alpha\beta}(\sigma) \ . \tag{1.26}$$

Or, infinitesimally, we can write $\Omega^2(\sigma) = e^{2\phi(\sigma)}$ for small ϕ so that

$$\delta g_{\alpha\beta}(\sigma) = 2\phi(\sigma)\, g_{\alpha\beta}(\sigma) \ .$$

It is simple to see that the Polyakov action is invariant under this transformation: the factor of Ω^2 drops out just as the factor of f did in equation (1.25), canceling between $\sqrt{-g}$ and the inverse metric $g^{\alpha\beta}$. This is a gauge symmetry of the string, as seen by the fact that the parameter Ω depends on the worldsheet coordinates σ. This means that two metrics which are related by a Weyl transformation (1.26) are to be considered as the same physical state.

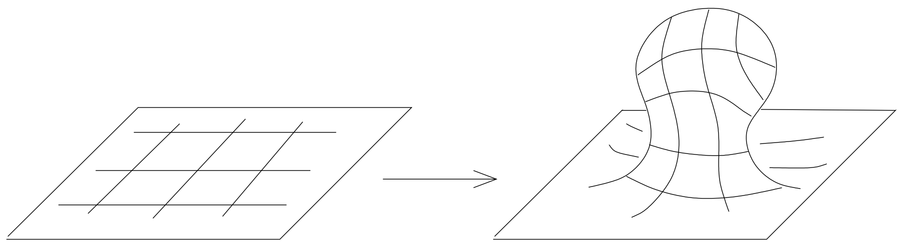


Figure 7: An example of a Weyl transformation

How should we think of Weyl invariance? It is not a coordinate change. Instead it is the invariance of the theory under a local change of scale which preserves the angles between all lines. For example the two worldsheet metrics shown in the figure are viewed by the Polyakov string as equivalent. This is rather surprising! And, as you might imagine, theories with this property are extremely rare. It should be clear from the discussion above that the property of Weyl invariance is special to two dimensions, for only there does the scaling factor coming from the determinant $\sqrt{-g}$ cancel that coming from the inverse metric. But even in two dimensions, if we wish to keep Weyl invariance then we are strictly limited in the kind of interactions that can be added to the action. For example, we would not be allowed a potential term for the worldsheet scalars of the form,

$$\int d^2\sigma \sqrt{-g}\, V(X) \ .$$

These break Weyl invariance. Nor can we add a worldsheet cosmological constant term,

$$\mu \int d^2\sigma \sqrt{-g} \ .$$

This too breaks Weyl invariance. We will see later in this course that the requirement of Weyl invariance becomes even more stringent in the quantum theory. We will also see what kind of interactions terms can be added to the worldsheet. Indeed, much of this course can be thought of as the study of theories with Weyl invariance.

1.3.2 Fixing a Gauge

As we have seen, the equation of motion (1.23) looks pretty nasty. However, we can use the redundancy inherent in the gauge symmetry to choose coordinates in which they simplify. Let's think about what we can do with the gauge symmetry.

Firstly, we have two reparameterizations to play with. The worldsheet metric has three independent components. This means that we expect to be able to set any two of the metric components to a value of our choosing. We will choose to make the metric locally conformally flat, meaning

$$g_{\alpha\beta} = e^{2\phi}\eta_{\alpha\beta} \ , \tag{1.27}$$

where $\phi(\sigma, \tau)$ is some function on the worldsheet. You can check that this is possible by writing down the change of the metric under a coordinate transformation and seeing that the differential equations which result from the condition (1.27) have solutions, at least locally. Choosing a metric of the form (1.27) is known as *conformal gauge*.

We have only used reparameterization invariance to get to the metric (1.27). We still have Weyl transformations to play with. Clearly, we can use these to remove the last independent component of the metric and set $\phi = 0$ such that,

$$g_{\alpha\beta} = \eta_{\alpha\beta} \ . \tag{1.28}$$

We end up with the flat metric on the worldsheet in Minkowski coordinates.

A Diversion: How to make a metric flat

The fact that we can use Weyl invariance to make any two-dimensional metric flat is an important result. Let's take a quick diversion from our main discussion to see a different proof that isn't tied to the choice of Minkowski coordinates on the worldsheet. We'll work in 2d Euclidean space to avoid annoying minus signs. Consider two metrics related by a Weyl transformation, $g'_{\alpha\beta} = e^{2\phi}g_{\alpha\beta}$. One can check that the Ricci scalars of the two metrics are related by,

$$\sqrt{g'}R' = \sqrt{g}(R - 2\nabla^2\phi) \ . \tag{1.29}$$

We can therefore pick a ϕ such that the new metric has vanishing Ricci scalar, $R' = 0$, simply by solving this differential equation for ϕ. However, in two dimensions (but not in higher dimensions) a vanishing Ricci scalar implies a flat metric. The reason is simply that there aren't too many indices to play with. In particular, symmetry of the Riemann tensor in two dimensions means that it must take the form,

$$R_{\alpha\beta\gamma\delta} = \frac{R}{2}(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}) \ .$$

So $R' = 0$ is enough to ensure that $R'_{\alpha\beta\gamma\delta} = 0$, which means that the manifold is flat. In equation (1.28), we've further used reparameterization invariance to pick coordinates in which the flat metric is the Minkowski metric.

The equations of motion and the stress-energy tensor

With the choice of the flat metric (1.28), the Polyakov action simplifies tremendously and becomes the theory of D free scalar fields. (In fact, this simplification happens in any conformal gauge).

$$S = -\frac{1}{4\pi\alpha'} \int d^2\sigma \; \partial_\alpha X \cdot \partial^\alpha X \;, \tag{1.30}$$

and the equations of motion for X^μ reduce to the free wave equation,

$$\partial_\alpha \partial^\alpha X^\mu = 0 \;. \tag{1.31}$$

Now that looks too good to be true! Are the horrible equations (1.23) really equivalent to a free wave equation? Well, not quite. There is something that we've forgotten: we picked a choice of gauge for the metric $g_{\alpha\beta}$. But we must still make sure that the equation of motion for $g_{\alpha\beta}$ is satisfied. In fact, the variation of the action with respect to the metric gives rise to a rather special quantity: it is the stress-energy tensor, $T_{\alpha\beta}$. With a particular choice of normalization convention, we define the stress-energy tensor to be

$$T_{\alpha\beta} = -\frac{2}{T}\frac{1}{\sqrt{-g}}\frac{\partial S}{\partial g^{\alpha\beta}} \;.$$

We varied the Polyakov action with respect to $g_{\alpha\beta}$ in (1.24). When we set $g_{\alpha\beta} = \eta_{\alpha\beta}$ we get

$$T_{\alpha\beta} = \partial_\alpha X \cdot \partial_\beta X - \tfrac{1}{2}\eta_{\alpha\beta}\,\eta^{\rho\sigma}\partial_\rho X \cdot \partial_\sigma X \;. \tag{1.32}$$

The equation of motion associated to the metric $g_{\alpha\beta}$ is simply $T_{\alpha\beta} = 0$. Or, more explicitly,

$$\begin{aligned} T_{01} &= \dot{X} \cdot X' = 0 \\ T_{00} = T_{11} &= \tfrac{1}{2}(\dot{X}^2 + X'^2) = 0 \;. \end{aligned} \tag{1.33}$$

We therefore learn that the equations of motion of the string are the free wave equations (1.31) subject to the two constraints (1.33) arising from the equation of motion $T_{\alpha\beta} = 0$.

Getting a feel for the constraints

Let's try to get some intuition for these constraints. There is a simple meaning of the first constraint in (1.33): we must choose our parameterization such that lines of constant σ are perpendicular to the lines of constant τ, as shown in the figure.

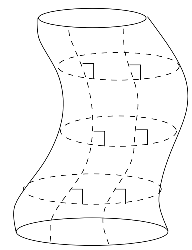


Figure 8:

But we can do better. To gain more physical insight, we need to make use of the fact that we haven't quite exhausted our gauge symmetry. We will discuss this more in Section 2.2, but for now one can check that there is enough remnant gauge symmetry to allow us to go to static gauge,

$$X^0 \equiv t = R\tau \ ,$$

so that $(X^0)' = 0$ and $\dot{X}^0 = R$, where R is a constant that is needed on dimensional grounds. The interpretation of this constant will become clear shortly. Then, writing $X^\mu = (t, \vec{x})$, the equation of motion for spatial components is the free wave equation,

$$\ddot{\vec{x}} - \vec{x}'' = 0$$

while the constraints become

$$\begin{aligned} \dot{\vec{x}} \cdot \vec{x}' &= 0 \\ \dot{\vec{x}}^{\,2} + \vec{x}'^{\,2} &= R^2 \end{aligned} \tag{1.34}$$

The first constraint tells us that the motion of the string must be perpendicular to the string itself. In other words, the physical modes of the string are transverse oscillations. There is no longitudinal mode. We'll also see this again in Section 2.2.

From the second constraint, we can understand the meaning of the constant R: it is related to the length of the string when $\dot{\vec{x}} = 0$,

$$\int d\sigma \, \sqrt{(d\vec{x}/d\sigma)^2} = 2\pi R \ .$$

Of course, if we have a stretched string with $\dot{\vec{x}} = 0$ at one moment of time, then it won't stay like that for long. It will contract under its own tension. As this happens, the second constraint equation relates the length of the string to the instantaneous velocity of the string.

1.4 Mode Expansions

Let's look at the equations of motion and constraints more closely. The equations of motion (1.31) are easily solved. We introduce lightcone coordinates on the worldsheet,

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

in terms of which the equations of motion simply read

$$\partial_+ \partial_- X^\mu = 0$$

The most general solution is,

$$X^\mu(\sigma, \tau) = X^\mu_L(\sigma^+) + X^\mu_R(\sigma^-)$$

for arbitrary functions X^μ_L and X^μ_R. These describe left-moving and right-moving waves respectively. Of course the solution must still obey both the constraints (1.33) as well as the periodicity condition,

$$X^\mu(\sigma, \tau) = X^\mu(\sigma + 2\pi, \tau) \ . \tag{1.35}$$

The most general, periodic solution can be expanded in Fourier modes,

$$\begin{aligned} X^\mu_L(\sigma^+) &= \tfrac{1}{2}x^\mu + \tfrac{1}{2}\alpha' p^\mu\, \sigma^+ + i\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{1}{n}\, \tilde{\alpha}^\mu_n\, e^{-in\sigma^+} \ , \\ X^\mu_R(\sigma^-) &= \tfrac{1}{2}x^\mu + \tfrac{1}{2}\alpha' p^\mu\, \sigma^- + i\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{1}{n}\, \alpha^\mu_n\, e^{-in\sigma^-} \ . \end{aligned} \tag{1.36}$$

This mode expansion will be very important when we come to the quantum theory. Let's make a few simple comments here.

- Various normalizations in this expression, such as the α' and factor of $1/n$ have been chosen for later convenience.

- X_L and X_R do not individually satisfy the periodicity condition (1.35) due to the terms linear in $\sigma^\pm$. However, the sum of them is invariant under $\sigma \to \sigma + 2\pi$ as required.

- The variables x^μ and p^μ are the position and momentum of the center of mass of the string. This can be checked, for example, by studying the Noether currents arising from the spacetime translation symmetry $X^\mu \to X^\mu + c^\mu$. One finds that the conserved charge is indeed p^μ.

- Reality of X^μ requires that the coefficients of the Fourier modes, α^μ_n and $\tilde{\alpha}^\mu_n$, obey

$$\alpha^\mu_n = (\alpha^\mu_{-n})^\star \quad , \quad \tilde{\alpha}^\mu_n = (\tilde{\alpha}^\mu_{-n})^\star \ . \tag{1.37}$$

1.4.1 The Constraints Revisited

We still have to impose the two constraints (1.33). In the worldsheet lightcone coordinates $\sigma^\pm$, these become,

$$(\partial_+ X)^2 = (\partial_- X)^2 = 0 \ . \tag{1.38}$$

These equations give constraints on the momenta p^μ and the Fourier modes α_n^μ and $\tilde{\alpha}_n^\mu$. To see what these are, let's look at

$$\begin{aligned}\partial_- X^\mu = \partial_- X_R^\mu &= \frac{\alpha'}{2} p^\mu + \sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \alpha_n^\mu \, e^{-in\sigma^-} \\ &= \sqrt{\frac{\alpha'}{2}} \sum_n \alpha_n^\mu e^{-in\sigma^-}\end{aligned}$$

where in the second line the sum is over all $n \in \mathbf{Z}$ and we have defined α_0^μ to be

$$\alpha_0^\mu \equiv \sqrt{\frac{\alpha'}{2}} p^\mu \ .$$

The constraint (1.38) can then be written as

$$\begin{aligned}(\partial_- X)^2 &= \frac{\alpha'}{2} \sum_{m,p} \alpha_m \cdot \alpha_p \, e^{-i(m+p)\sigma^-} \\ &= \frac{\alpha'}{2} \sum_{m,n} \alpha_m \cdot \alpha_{n-m} \, e^{-in\sigma^-} \\ &\equiv \alpha' \sum_n L_n \, e^{-in\sigma^-} = 0 \ .\end{aligned}$$

where we have defined the sum of oscillator modes,

$$L_n = \frac{1}{2} \sum_m \alpha_{n-m} \cdot \alpha_m \ . \tag{1.39}$$

We can also do the same for the left-moving modes, where we again define an analogous sum of operator modes,

$$\tilde{L}_n = \frac{1}{2} \sum_m \tilde{\alpha}_{n-m} \cdot \tilde{\alpha}_m \ . \tag{1.40}$$

with the zero mode defined to be,

$$\tilde{\alpha}_0^\mu \equiv \sqrt{\frac{\alpha'}{2}} p^\mu \ .$$

The fact that $\tilde{\alpha}_0^\mu = \alpha_0^\mu$ looks innocuous but is a key point to remember when we come to quantize the string. The L_n and $\tilde{L}_n$ are the Fourier modes of the constraints. Any classical solution of the string of the form (1.36) must further obey the infinite number of constraints,

$$L_n = \tilde{L}_n = 0 \qquad n \in \mathbf{Z} \; .$$

We'll meet these objects L_n and $\tilde{L}_n$ again in a more general context when we come to discuss conformal field theory.

The constraints arising from L_0 and $\tilde{L}_0$ have a rather special interpretation. This is because they include the square of the spacetime momentum p^μ. But, the square of the spacetime momentum is an important quantity in Minkowski space: it is the square of the rest mass of a particle,

$$p_\mu p^\mu = -M^2 \; .$$

So the L_0 and $\tilde{L}_0$ constraints tell us the effective mass of a string in terms of the excited oscillator modes, namely

$$M^2 = \frac{4}{\alpha'} \sum_{n>0} \alpha_n \cdot \alpha_{-n} = \frac{4}{\alpha'} \sum_{n>0} \tilde{\alpha}_n \cdot \tilde{\alpha}_{-n} \tag{1.41}$$

Because both α_0^μ and $\tilde{\alpha}_0^\mu$ are equal to $\sqrt{\alpha'/2}\, p^\mu$, we have two expressions for the invariant mass: one in terms of right-moving oscillators α_n^μ and one in terms of left-moving oscillators $\tilde{\alpha}_n^\mu$. And these two terms must be equal to each other. This is known as *level matching*. It will play an important role in the next section where we turn to the quantum theory.

2. The Quantum String

Our goal in this section is to quantize the string. We have seen that the string action involves a gauge symmetry and whenever we wish to quantize a gauge theory we're presented with a number of different ways in which we can proceed. If we're working in the canonical formalism, this usually boils down to one of two choices:

- We could first quantize the system and then subsequently impose the constraints that arise from gauge fixing as operator equations on the physical states of the system. For example, in QED this is the Gupta-Bleuler method of quantization that we use in Lorentz gauge. In string theory it consists of treating all fields X^μ, including time X^0, as operators and imposing the constraint equations (1.33) on the states. This is usually called covariant quantization.

- The alternative method is to first solve all of the constraints of the system to determine the space of physically distinct classical solutions. We then quantize these physical solutions. For example, in QED, this is the way we proceed in Coulomb gauge. Later in this chapter, we will see a simple way to solve the constraints of the free string.

Of course, if we do everything correctly, the two methods should agree. Usually, each presents a slightly different challenge and offers a different viewpoint.

In these lectures, we'll take a brief look at the first method of covariant quantization. However, at the slightest sign of difficulties, we'll bail! It will be useful enough to see where the problems lie. We'll then push forward with the second method described above which is known as lightcone quantization in string theory. Although we'll succeed in pushing quantization through to the end, our derivations will be a little cheap and unsatisfactory in places. In Section 5 we'll return to all these issues, armed with more sophisticated techniques from conformal field theory.

2.1 A Lightning Look at Covariant Quantization

We wish to quantize D free scalar fields X^μ whose dynamics is governed by the action (1.30). We subsequently wish to impose the constraints

$$\dot{X} \cdot X' = \dot{X}^2 + X'^2 = 0 \ . \tag{2.1}$$

The first step is easy. We promote X^μ and their conjugate momenta $\Pi_\mu = (1/2\pi\alpha')\dot{X}_\mu$ to operator valued fields obeying the canonical equal-time commutation relations,

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

We translate these into commutation relations for the Fourier modes x^μ, p^μ, α_n^μ and $\tilde{\alpha}_n^\mu$. Using the mode expansion (1.36) we find

$$[x^\mu, p_\nu] = i\delta^\mu_\nu \quad \text{and} \quad [\alpha_n^\mu, \alpha_m^\nu] = [\tilde{\alpha}_n^\mu, \tilde{\alpha}_m^\nu] = n\,\eta^{\mu\nu}\delta_{n+m,0}\ , \tag{2.2}$$

with all others zero. The commutation relations for x^μ and p^μ are expected for operators governing the position and momentum of the center of mass of the string. The commutation relations of α_n^μ and $\tilde{\alpha}_n^\mu$ are those of harmonic oscillator creation and annihilation operators in disguise. And the disguise isn't that good. We just need to define (ignoring the μ index for now)

$$a_n = \frac{\alpha_n}{\sqrt{n}} \quad , \quad a_n^\dagger = \frac{\alpha_{-n}}{\sqrt{n}} \qquad\qquad n > 0 \tag{2.3}$$

Then (2.2) gives the familiar $[a_n, a_m^\dagger] = \delta_{mn}$. So each scalar field gives rise to two infinite towers of creation and annihilation operators, with α_n acting as a rescaled annihilation operator for $n > 0$ and as a creation operator for $n < 0$. There are two towers because we have right-moving modes α_n and left-moving modes $\tilde{\alpha}_n$.

With these commutation relations in hand we can now start building the Fock space of our theory. We introduce a vacuum state of the string $|0\rangle$, defined to obey

$$\alpha_n^\mu\,|0\rangle = \tilde{\alpha}_n^\mu\,|0\rangle = 0 \qquad \text{for} \quad n > 0 \tag{2.4}$$

The vacuum state of string theory has a different interpretation from the analogous object in field theory. This is not the vacuum state of spacetime. It is instead the vacuum state of a single string. This is reflected in the fact that the operators x^μ and p^μ give extra structure to the vacuum. The true ground state of the string is $|0\rangle$, tensored with a spatial wavefunction $\Psi(x)$. Alternatively, if we work in momentum space, the vacuum carries another quantum number, p^μ, which is the eigenvalue of the momentum operator. We should therefore write the vacuum as $|0;p\rangle$, which still obeys (2.4), but now also

$$\hat{p}^\mu\,|0;p\rangle = p^\mu|0;p\rangle \tag{2.5}$$

where (for the only time in these lecture notes) we've put a hat on the momentum operator $\hat{p}^\mu$ on the left-hand side of this equation to distinguish it from the eigenvalue p^μ on the right-hand side.

We can now start to build up the Fock space by acting with creation operators α_n^μ and $\tilde{\alpha}_n^\mu$ with $n < 0$. A generic state comes from acting with any number of these creation operators on the vacuum,

$$(\alpha_{-1}^{\mu_1})^{n_{\mu_1}}(\alpha_{-2}^{\mu_2})^{n_{\mu_2}}\ldots(\tilde{\alpha}_{-1}^{\nu_1})^{n_{\nu_1}}(\tilde{\alpha}_{-2}^{\nu_2})^{n_{\nu_2}}\ldots|0;p\rangle$$

Each state in the Fock space is a different excited state of the string. Each has the interpretation of a different species of particle in spacetime. We'll see exactly what particles they are shortly. But for now, notice that because there's an infinite number of ways to excite a string there are an infinite number of different species of particles in this theory.

2.1.1 Ghosts

There's a problem with the Fock space that we've constructed: it doesn't have positive norm. The reason for this is that one of the scalar fields, X^0, comes with the wrong sign kinetic term in the action (1.30). From the perspective of the commutation relations, this issue raises its head in presence of the spacetime Minkowski metric in the expression

$$[\alpha_n^\mu, \alpha_m^{\nu\,\dagger}] = n\,\eta^{\mu\nu}\,\delta_{n,m}\ .$$

This gives rise to the offending negative norm states, which come with an odd number of timelike oscillators excited, for example

$$\langle p';0|\alpha_1^0\alpha_{-1}^0|0;p\rangle \sim -\delta^D(p-p')$$

This is the first problem that arises in the covariant approach to quantization. States with negative norm are referred to as *ghosts*. To make sense of the theory, we have to make sure that they can't be produced in any physical processes. Of course, this problem is familiar from attempts to quantize QED in Lorentz gauge. In that case, gauge symmetry rides to the rescue since the ghosts are removed by imposing the gauge fixing constraint. We must hope that the same happens in string theory.

2.1.2 Constraints

Although we won't push through with this programme at the present time, let us briefly look at what kind of constraints we have in string theory. In terms of Fourier modes, the classical constraints can be written as $L_n = \tilde{L}_n = 0$, where

$$L_n = \frac{1}{2}\sum_m \alpha_{n-m}\cdot\alpha_m$$

and similar for $\tilde{L}_n$. As in the Gupta-Bleuler quantization of QED, we don't impose all of these as operator equations on the Hilbert space. Instead we only require that the operators L_n and $\tilde{L}_n$ have vanishing matrix elements when sandwiched between two physical states $|\text{phys}\rangle$ and $|\text{phys}'\rangle$,

$$\langle \text{phys}'|L_n|\text{phys}\rangle = \langle \text{phys}'|\tilde{L}_n|\text{phys}\rangle = 0$$

Because $L_n^\dagger = L_{-n}$, it is therefore sufficient to require

$$L_n|\text{phys}\rangle = \tilde{L}_n|\text{phys}\rangle = 0 \quad \text{for} \;\; n > 0 \tag{2.6}$$

However, we still haven't explained how to impose the constraints L_0 and $\tilde{L}_0$. And these present a problem that doesn't arise in the case of QED. The problem is that, unlike for L_n with $n \neq 0$, the operator L_0 is not uniquely defined when we pass to the quantum theory. There is an operator ordering ambiguity arising from the commutation relations (2.2). Commuting the α_n^μ operators past each other in L_0 gives rise to extra constant terms.

Question: How do we know what order to put the α_n^μ operators in the quantum operator L_0? Or the $\tilde{\alpha}_n^\mu$ operators in $\tilde{L}_0$?

Answer: We don't! Yet. Naively it looks as if each different choice will define a different theory when we impose the constraints. To make this ambiguity manifest, for now let's just pick a choice of ordering. We define the quantum operators to be normal ordered, with the annihilation operators α_n^i, $n > 0$, moved to the right,

$$L_0 = \sum_{m=1}^{\infty} \alpha_{-m} \cdot \alpha_m + \frac{1}{2}\,\alpha_0^2 \quad , \quad \tilde{L}_0 = \sum_{m=1}^{\infty} \tilde{\alpha}_{-m} \cdot \tilde{\alpha}_m + \frac{1}{2}\,\tilde{\alpha}_0^2$$

Then the ambiguity rears its head in the different constraint equations that we could impose, namely

$$(L_0 - a)|\text{phys}\rangle = (\tilde{L}_0 - a)|\text{phys}\rangle = 0 \tag{2.7}$$

for some constant a.

As we saw classically, the operators L_0 and $\tilde{L}_0$ play an important role in determining the spectrum of the string because they include a term quadratic in the momentum $\alpha_0^\mu = \tilde{\alpha}_0^\mu = \sqrt{\alpha'/2}\,p^\mu$. Combining the expression (1.41) with our constraint equation for L_0 and $\tilde{L}_0$, we find the spectrum of the string is given by,

$$M^2 = \frac{4}{\alpha'}\left(-a + \sum_{m=1}^{\infty} \alpha_{-m} \cdot \alpha_m\right) = \frac{4}{\alpha'}\left(-a + \sum_{m=1}^{\infty} \tilde{\alpha}_{-m} \cdot \tilde{\alpha}_m\right)$$

We learn therefore that the undetermined constant a has a direct physical effect: it changes the mass spectrum of the string. In the quantum theory, the sums over α_n^μ modes are related to the number operators for the harmonic oscillator: they count the number of excited modes of the string. The level matching in the quantum theory tells us that the number of left-moving modes must equal the number of right-moving modes.

Ultimately, we will find that the need to decouple the ghosts forces us to make a unique choice for the constant a. (Spoiler alert: it turns out to be $a = 1$). In fact, the requirement that there are no ghosts is much stronger than this. It also restricts the number of scalar fields that we have in the theory. (Another spoiler: $D = 26$). If you're interested in how this works in covariant formulation then you can read about it in the book by Green, Schwarz and Witten. Instead, we'll show how to quantize the string and derive these values for a and D in lightcone gauge. However, after a trip through the world of conformal field theory, we'll come back to these ideas in a context which is closer to the covariant approach.

2.2 Lightcone Quantization

We will now take the second path described at the beginning of this section. We will try to find a parameterization of all classical solutions of the string. This is equivalent to finding the classical phase space of the theory. We do this by solving the constraints (2.1) in the classical theory, leaving behind only the physical degrees of freedom.

Recall that we fixed the gauge to set the worldsheet metric to

$$g_{\alpha\beta} = \eta_{\alpha\beta} .$$

However, this isn't the end of our gauge freedom. There still remain gauge transformations which preserve this choice of metric. In particular, any coordinate transformation $\sigma \to \tilde{\sigma}(\sigma)$ which changes the metric by

$$\eta_{\alpha\beta} \to \Omega^2(\sigma)\eta_{\alpha\beta} \ , \tag{2.8}$$

can be undone by a Weyl transformation. What are these coordinate transformations? It's simplest to answer this using lightcone coordinates on the worldsheet,

$$\sigma^\pm = \tau \pm \sigma \ , \tag{2.9}$$

where the flat metric on the worldsheet takes the form,

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

In these coordinates, it's clear that any transformation of the form

$$\sigma^+ \to \tilde{\sigma}^+(\sigma^+) \quad , \quad \sigma^- \to \tilde{\sigma}^-(\sigma^-) \ , \tag{2.10}$$

simply multiplies the flat metric by an overall factor (2.8) and so can be undone by a compensating Weyl transformation. Some quick comments on this surviving gauge symmetry:

- Recall that in Section 1.3.2 we used the argument that 3 gauge invariances (2 reparameterizations + 1 Weyl) could be used to fix 3 components of the worldsheet metric $g_{\alpha\beta}$. What happened to this argument? Why do we still have some gauge symmetry left? The reason is that $\tilde{\sigma}^{\pm}$ are functions of just a single variable, not two. So we did fix nearly all our gauge symmetries. What is left is a set of measure zero amongst the full gauge symmetry that we started with.

- The remaining reparameterization invariance (2.10) has an important physical implication. Recall that the solutions to the equations of motion are of the form $X^\mu_L(\sigma^+) + X^\mu_R(\sigma^-)$ which looks like $2D$ functions worth of solutions. Of course, we still have the constraints which, in terms of $\sigma^\pm$, read

$$(\partial_+ X)^2 = (\partial_- X)^2 = 0 \ , \tag{2.11}$$

which seems to bring the number down to $2(D-1)$ functions. But the reparameterization invariance (2.10) tells us that even some of these are fake since we can always change what we mean by $\sigma^\pm$. The physical solutions of the string are therefore actually described by $2(D-2)$ functions. But this counting has a nice interpretation: the degrees of freedom describe the *transverse* fluctuations of the string.

- The above comment reaches the same conclusion as the discussion in Section 1.3.2. There, in an attempt to get some feel for the constraints, we claimed that we could go to static gauge $X^0 = R\tau$ for some dimensionful parameter R. It is easy to check that this is simple to do using reparameterizations of the form (2.10). However, to solve the string constraints in full, it turns out that static gauge is not that useful. Rather we will use something called "lightcone gauge".

2.2.1 Lightcone Gauge

We would like to gauge fix the remaining reparameterization invariance (2.10). The best way to do this is called lightcone gauge. In counterpoint to the worldsheet lightcone coordinates (2.9), we introduce the spacetime lightcone coordinates,

$$X^\pm = \sqrt{\frac{1}{2}}(X^0 \pm X^{D-1}) \ . \tag{2.12}$$

Note that this choice picks out a particular time direction and a particular spatial direction. It means that any calculations that we do involving $X^\pm$ will not be manifestly Lorentz invariant. You might think that we needn't really worry about this. We could try to make the following argument: "The equations may not *look* Lorentz invariant but, since we started from a Lorentz invariant theory, at the end of the day any physical process is guaranteed to obey this symmetry". Right?! Well, unfortunately not. One of the more interesting and subtle aspects of quantum field theory is the possibility of anomalies: these are symmetries of the classical theory that do not survive the journey of quantization. When we come to the quantum theory, if our equations don't look Lorentz invariant then there's a real possibility that it's because the underlying physics actually isn't Lorentz invariant. Later we will need to spend some time figuring out under what circumstances our quantum theory keeps the classical Lorentz symmetry.

In lightcone coordinates, the spacetime Minkowski metric reads

$$ds^2 = -2dX^+dX^- + \sum_{i=1}^{D-2} dX^i dX^i$$

This means that indices are raised and lowered with $A_+ = -A^-$ and $A_- = -A^+$ and $A_i = A^i$. The product of spacetime vectors reads $A \cdot B = -A^+B^- - A^-B^+ + A^iB^i$.

Let's look at the solution to the equation of motion for X^+. It reads,

$$X^+ = X_L^+(\sigma^+) + X_R^+(\sigma^-) \ .$$

We now gauge fix. We use our freedom of reparameterization invariance to choose coordinates such that

$$X_L^+ = \tfrac{1}{2}\,x^+ + \tfrac{1}{2}\,\alpha' p^+ \sigma^+ \quad , \quad X_R^+ = \tfrac{1}{2}\,x^+ + \tfrac{1}{2}\,\alpha' p^+ \sigma^- \ .$$

You might think that we could go further and eliminate p^+ and x^+ but this isn't possible because we don't quite have the full freedom of reparameterization invariance since all functions should remain periodic in σ. The upshot of this choice of gauge is that

$$X^+ = x^+ + \alpha' p^+ \tau \ . \tag{2.13}$$

This is *lightcone gauge*. Notice that, as long as $p^+ \neq 0$, we can always shift x^+ by a shift in τ.

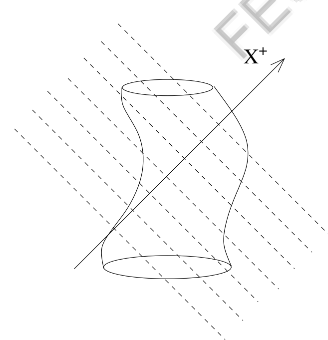


Figure 9:

There's something a little disconcerting about the choice (2.13). We've identified a timelike worldsheet coordinate with a null spacetime coordinate. Nonetheless, as you can see from the figure, it seems to be a good parameterization of the worldsheet. One could imagine that the parameterization might break if the string is actually massless and travels in the X^- direction, with $p^+ = 0$. But otherwise, all should be fine.

Solving for X^-

The choice (2.13) does the job of fixing the reparameterization invariance (2.10). As we will now see, it also renders the constraint equations trivial. The first thing that we have to worry about is the possibility of extra constraints arising from this new choice of gauge fixing. This can be checked by looking at the equation of motion for X^+,

$$\partial_+\partial_- X^- = 0$$

But we can solve this by the usual ansatz,

$$X^- = X^-_L(\sigma^+) + X^-_R(\sigma^-)\ .$$

We're still left with all the other constraints (2.11). Here we see the real benefit of working in lightcone gauge (which is actually what makes quantization possible at all): X^- is completely determined by these constraints. For example, the first of these reads

$$2\partial_+ X^- \partial_+ X^+ = \sum_{i=1}^{D-2} \partial_+ X^i \partial_+ X^i \tag{2.14}$$

which, using (2.13), simply becomes

$$\partial_+ X^-_L = \frac{1}{\alpha' p^+} \sum_{i=1}^{D-2} \partial_+ X^i \partial_+ X^i\ . \tag{2.15}$$

Similarly,

$$\partial_- X^-_R = \frac{1}{\alpha' p^+} \sum_{i=1}^{D-2} \partial_- X^i \partial_- X^i\ . \tag{2.16}$$

So, up to an integration constant, the function $X^-(\sigma^+, \sigma^-)$ is completely determined in terms of the other fields. If we write the usual mode expansion for $X^-_{L/R}$

$$X^-_L(\sigma^+) = \tfrac{1}{2}x^- + \tfrac{1}{2}\,\alpha' p^-\,\sigma^+ + i\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{1}{n}\,\tilde{\alpha}^-_n\, e^{-in\sigma^+}\ ,$$

$$X^-_R(\sigma^-) = \tfrac{1}{2}x^- + \tfrac{1}{2}\,\alpha' p^-\,\sigma^- + i\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{1}{n}\,\alpha^-_n\, e^{-in\sigma^-}\ .$$

then x^- is the undetermined integration constant, while p^-, α^-_n and $\tilde{\alpha}^-_n$ are all fixed by the constraints (2.15) and (2.16). For example, the oscillator modes α^-_n are given by,

$$\alpha^-_n = \sqrt{\frac{1}{2\alpha'}} \frac{1}{p^+} \sum_{m=-\infty}^{+\infty} \sum_{i=1}^{D-2} \alpha^i_{n-m}\alpha^i_m\ , \tag{2.17}$$

A special case of this is the $\alpha_0^- = \sqrt{\alpha'/2}\, p^-$ equation, which reads

$$\frac{\alpha' p^-}{2} = \frac{1}{2p^+} \sum_{i=1}^{D-2} \left(\tfrac{1}{2}\,\alpha' p^i p^i + \sum_{n\neq 0} \alpha_n^i \alpha_{-n}^i \right) . \tag{2.18}$$

We also get another equation for p^- from the $\tilde{\alpha}_0^-$ equation arising from (2.15)

$$\frac{\alpha' p^-}{2} = \frac{1}{2p^+} \sum_{i=1}^{D-2} \left(\tfrac{1}{2}\,\alpha' p^i p^i + \sum_{n\neq 0} \tilde{\alpha}_n^i \tilde{\alpha}_{-n}^i \right) . \tag{2.19}$$

From these two equations, we can reconstruct the old, classical, level matching conditions (1.41). But now with a difference:

$$M^2 = 2p^+p^- - \sum_{i=1}^{D-2} p^i p^i = \frac{4}{\alpha'} \sum_{i=1}^{D-2} \sum_{n>0} \alpha_{-n}^i \alpha_n^i = \frac{4}{\alpha'} \sum_{i=1}^{D-2} \sum_{n>0} \tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i . \tag{2.20}$$

The difference is that now the sum is over oscillators α^i and $\tilde{\alpha}^i$ only, with $i = 1, \ldots, D-2$. We'll refer to these as *transverse* oscillators. Note that the string isn't necessarily living in the X^0-X^{D-1} plane, so these aren't literally the transverse excitations of the string. Nonetheless, if we specify the α^i then all other oscillator modes are determined. In this sense, they are the physical excitation of the string.

Let's summarize the state of play so far. The most general classical solution is described in terms of $2(D-2)$ transverse oscillator modes α_n^i and $\tilde{\alpha}_n^i$, together with a number of zero modes describing the center of mass and momentum of the string: x^i, p^i, p^+ and x^-. But x^+ can be absorbed by a shift of τ in (2.13) and p^- is constrained to obey (2.18) and (2.19). In fact, p^- can be thought of as (proportional to) the lightcone Hamiltonian. Indeed, we know that p^- generates translations in x^+, but this is equivalent to shifts in τ.

2.2.2 Quantization

Having identified the physical degrees of freedom, let's now quantize. We want to impose commutation relations. Some of these are easy:

$$\begin{gathered} [x^i, p^j] = i\delta^{ij} \quad , \quad [x^-, p^+] = -i \\ [\alpha_n^i, \alpha_m^j] = [\tilde{\alpha}_n^i, \tilde{\alpha}_m^j] = n\delta^{ij}\delta_{n+m,0} \ . \end{gathered} \tag{2.21}$$

all of which follow from the commutation relations (2.2) that we saw in covariant quantization[1].

What to do with x^+ and p^-? We could implement p^- as the Hamiltonian acting on states. In fact, it will prove slightly more elegant (but equivalent) if we promote both x^+ and p^- to operators with the expected commutation relation,

$$[x^+, p^-] = -i \ . \tag{2.22}$$

This is morally equivalent to writing $[t, H] = -i$ in non-relativistic quantum mechanics, which is true on a formal level. In the present context, it means that we can once again choose states to be eigenstates of p^μ, with $\mu = 0, \ldots, D$, but the constraints (2.18) and (2.19) must still be imposed as operator equations on the physical states. We'll come to this shortly.

The Hilbert space of states is very similar to that described in covariant quantization: we define a vacuum state, $|0;p\rangle$ such that

$$\hat{p}^\mu |0;p\rangle = p^\mu |0;p\rangle \quad , \quad \alpha_n^i |0;p\rangle = \tilde{\alpha}_n^i |0;p\rangle = 0 \quad \text{for} \ \ n > 0 \tag{2.23}$$

and we build a Fock space by acting with the creation operators α_{-n}^i and $\tilde{\alpha}_{-n}^i$ with $n > 0$. The difference with the covariant quantization is that we only act with transverse oscillators which carry a spatial index $i = 1, \ldots, D-2$. For this reason, the Hilbert space is, by construction, positive definite. We don't have to worry about ghosts.

[1]**Mea Culpa:** We're not really supposed to do this. The whole point of the approach that we're taking is to quantize just the physical degrees of freedom. The resulting commutation relations are not, in general, inherited from the larger theory that we started with simply by closing our eyes and forgetting about all the other fields that we've gauge fixed. We can see the problem by looking at (2.17), where α_n^- is determined in terms of α_n^i. This means that the commutation relations for α_n^i might be infected by those of α_n^- which could potentially give rise to extra terms. The correct procedure to deal with this is to figure out the Poisson bracket structure of the physical degrees of freedom in the classical theory. Or, in fancier language, the symplectic form on the phase space which schematically looks like

$$\omega \sim \int d\sigma \ \ - d\dot{X}^+ \wedge dX^- - d\dot{X}^- \wedge dX^+ + 2 d\dot{X}^i \wedge dX^i \ ,$$

The reason that the commutation relations (2.21) do not get infected is because the α^- terms in the symplectic form come multiplying X^+. Yet X^+ is given in (2.13). It has no oscillator modes. That means that the symplectic form doesn't pick up the Fourier modes of X^- and so doesn't receive any corrections from α_n^-. The upshot of this is that the naive commutation relations (2.21) are actually right.

The Constraints

Because p^- is not an independent variable in our theory, we must impose the constraints (2.18) and (2.19) by hand as operator equations which define the physical states. In the classical theory, we saw that these constraints are equivalent to mass-shell conditions (2.20).

But there's a problem when we go to the quantum theory. It's the same problem that we saw in covariant quantization: there's an ordering ambiguity in the sum over oscillator modes on the right-hand side of (2.20). If we choose all operators to be normal ordered then this ambiguity reveals itself in an overall constant, a, which we have not yet determined. The final result for the mass of states in lightcone gauge is:

$$M^2 = \frac{4}{\alpha'}\left(\sum_{i=1}^{D-2}\sum_{n>0}\alpha^i_{-n}\alpha^i_n - a\right) = \frac{4}{\alpha'}\left(\sum_{i=1}^{D-2}\sum_{n>0}\tilde{\alpha}^i_{-n}\tilde{\alpha}^i_n - a\right)$$

Since we'll use this formula quite a lot in what follows, it's useful to introduce quantities related to the number operators of the harmonic oscillator,

$$N = \sum_{i=1}^{D-2}\sum_{n>0}\alpha^i_{-n}\alpha^i_n \quad , \quad \tilde{N} = \sum_{i=1}^{D-2}\sum_{n>0}\tilde{\alpha}^i_{-n}\tilde{\alpha}^i_n \ . \tag{2.24}$$

These are not quite number operators because of the factor of $1/\sqrt{n}$ in (2.3). The value of N and $\tilde{N}$ is often called the level. Which, if nothing else, means that the name "level matching" makes sense. We now have

$$M^2 = \frac{4}{\alpha'}(N-a) = \frac{4}{\alpha'}(\tilde{N}-a) \ . \tag{2.25}$$

How are we going to fix a? Later in the course we'll see the correct way to do it. For now, I'm just going to give you a quick and dirty derivation.

The Casimir Energy

> "I told him that the sum of an infinite no. of terms of the series: $1+2+3+4+\ldots = -\frac{1}{12}$ under my theory. If I tell you this you will at once point out to me the lunatic asylum as my goal. "
>
> *Ramanujan, in a letter to G.H.Hardy.*

What follows is a heuristic derivation of the normal ordering constant a. Suppose that we didn't notice that there was any ordering ambiguity and instead took the naive classical result directly over to the quantum theory, that is

$$\frac{1}{2}\sum_{n\neq 0}\alpha^i_{-n}\alpha^i_n = \frac{1}{2}\sum_{n<0}\alpha^i_{-n}\alpha^i_n + \frac{1}{2}\sum_{n>0}\alpha^i_{-n}\alpha^i_n \ .$$

where we've left the sum over $i = 1, \ldots, D-2$ implicit. We'll now try to put this in normal ordered form, with the annihilation operators α_n^i with $n > 0$ on the right-hand side. It's the first term that needs changing. We get

$$\frac{1}{2}\sum_{n<0}\left[\alpha_n^i\alpha_{-n}^i - n(D-2)\right] + \frac{1}{2}\sum_{n>0}\alpha_{-n}^i\alpha_n^i = \sum_{n>0}\alpha_{-n}^i\alpha_n^i + \frac{D-2}{2}\sum_{n>0} n \quad .$$

The final term clearly diverges. But it at least seems to have a physical interpretation: it is the sum of zero point energies of an infinite number of harmonic oscillators. In fact, we came across exactly the same type of term in the course on quantum field theory where we learnt that, despite the divergence, one can still extract interesting physics from this. This is the physics of the Casimir force.

Let's recall the steps that we took to derive the Casimir force. Firstly, we introduced an ultra-violet cut-off $\epsilon \ll 1$, probably muttering some words about no physical plates being able to withstand very high energy quanta. Unfortunately, those words are no longer available to us in string theory, but let's proceed regardless. We replace the divergent sum over integers by the expression,

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

Obviously the $1/\epsilon^2$ piece diverges as $\epsilon \to 0$. This term should be renormalized away. In fact, this is necessary to preserve the Weyl invariance of the Polyakov action since it contributes to a cosmological constant on the worldsheet. After this renormalization, we're left with the wonderful answer, first intuited by Ramanujan

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

While heuristic, this argument does predict the correct physical Casimir energy measured in one-dimensional systems. For example, this effect is seen in simulations of quantum spin chains.

What does this mean for our string? It means that we should take the unknown constant a in the mass formula (2.25) to be,

$$M^2 = \frac{4}{\alpha'}\left(N - \frac{D-2}{24}\right) = \frac{4}{\alpha'}\left(\tilde{N} - \frac{D-2}{24}\right) \; . \tag{2.26}$$

This is the formula that we will use to determine the spectrum of the string.

Zeta Function Regularization

I appreciate that the preceding argument is not totally convincing. We could spend some time making it more robust at this stage, but it's best if we wait until later in the course when we will have the tools of conformal field theory at our disposal. We will eventually revisit this issue and provide a respectable derivation of the Casimir energy in Section 4.4.1. For now I merely offer an even less convincing argument, known as zeta-function regularization.

The zeta-function is defined, for $\mathrm{Re}(s) > 1$, by the sum

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

But $\zeta(s)$ has a unique analytic continuation to all values of s. In particular,

$$\zeta(-1) = -\frac{1}{12} .$$

Good? Good. This argument is famously unconvincing the first time you meet it! But it's actually a very useful trick for getting the right answer.

2.3 The String Spectrum

Finally, we're in a position to analyze the spectrum of a single, free string.

2.3.1 The Tachyon

Let's start with the ground state $|0;p\rangle$ defined in (2.23). With no oscillators excited, the mass formula (2.26) gives

$$M^2 = -\frac{1}{\alpha'}\frac{D-2}{6} . \tag{2.27}$$

But that's a little odd. It's a negative mass-squared. Such particles are called *tachyons*.

In fact, tachyons aren't quite as pathological as you might think. If you've heard of these objects before, it's probably in the context of special relativity where they're strange beasts which always travel faster than the speed of light. But that's not the right interpretation. Rather we should think more in the language of quantum field theory. Suppose that we have a field in spacetime — let's call it $T(X)$ — whose quanta will give rise to this particle. The mass-squared of the particle is simply the quadratic term in the action, or

$$M^2 = \left.\frac{\partial^2 V(T)}{\partial T^2}\right|_{T=0}$$

So the negative mass-squared in (2.27) is telling us that we're expanding around a maximum of the potential for the tachyon field as shown in the figure. Note that from this perspective, the Higgs field in the standard model at $H = 0$ is also a tachyon.

The fact that string theory turns out to sit at an unstable point in the tachyon field is unfortunate. The natural question is whether the potential has a good minimum elsewhere, as shown in the figure to the right. No one knows the answer to this! Naive attempts to understand this don't work. We know that around $T = 0$, the leading order contribution to the potential is negative and quadratic. But there are further terms that we can compute using techniques that we'll describe in Section 6. An expansion of the tachyon potential around $T = 0$ looks like

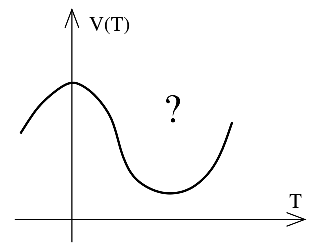


Figure 11:

$$V(T) = \frac{1}{2}M^2T^2 + c_3T^3 + c_4T^4 + \ldots$$

It turns out that the T^3 term in the potential does give rise to a minimum. But the T^4 term destabilizes it again. Moreover, the T field starts to mix with other scalar fields in the theory that we will come across soon. The ultimate fate of the tachyon in the bosonic string is not yet understood.

The tachyon is a problem for the bosonic string. It may well be that this theory makes no sense — or, at the very least, has no time-independent stable solutions. Or perhaps we just haven't worked out how to correctly deal with the tachyon. Either way, the problem does not arise when we introduce fermions on the worldsheet and study the superstring. This will involve several further technicalities which we won't get into in this course. Instead, our time will be put to better use if we continue to study the bosonic string since all the lessons that we learn will carry over directly to the superstring. However, one should be aware that the problem of the unstable vacuum will continue to haunt us throughout this course.

Although we won't describe it in detail, at several times along our journey we'll make an aside about how calculations work out for the superstring.

2.3.2 The First Excited States

We now look at the first excited states. If we act with a creation operator α^j_{-1}, then the level matching condition (2.25) tells us that we also need to act with a $\tilde{\alpha}^i_{-1}$ operator. This gives us $(D-2)^2$ particle states,

$$\tilde{\alpha}^i_{-1}\alpha^j_{-1}\,|0;p\rangle \ , \qquad (2.28)$$

each of which has mass

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

But now we seem to have a problem. Our states have space indices $i, j = 1, \ldots, D-2$. The operators α^i and $\tilde{\alpha}^i$ each transform in the vector representation of $SO(D-2) \subset SO(1, D-1)$ which is manifest in lightcone gauge. But ultimately we want these states to fit into some representation of the full Lorentz $SO(1, D-1)$ group. That looks as if it's going to be hard to arrange. This is the first manifestation of the comment that we made after equation (2.12): it's tricky to see Lorentz invariance in lightcone gauge.

To proceed, let's recall Wigner's classification of representations of the Poincaré group. We start by looking at massive particles in $\mathbf{R}^{1,D-1}$. After going to the rest frame of the particle by setting $p^\mu = (p, 0, \ldots, 0)$, we can watch how any internal indices transform under the little group $SO(D-1)$ of spatial rotations. The upshot of this is that any massive particle must form a representation of $SO(D-1)$. But the particles described by (2.28) have $(D-2)^2$ states. There's no way to package these states into a representation of $SO(D-1)$ and this means that there's no way that the first excited states of the string can form a massive representation of the D-dimensional Poincaré group.

It looks like we're in trouble. Thankfully, there's a way out. If the states are massless, then we can't go to the rest frame. The best that we can do is choose a spacetime momentum for the particle of the form $p^\mu = (p, 0, \ldots, 0, p)$. In this case, the particles fill out a representation of the little group $SO(D-2)$. This means that massless particles get away with having fewer internal states than massive particles. For example, in four dimensions the photon has two polarization states, but a massive spin-1 particle must have three.

The first excited states (2.28) happily sit in a representation of $SO(D-2)$. We learn that if we want the quantum theory to preserve the $SO(1, D-1)$ Lorentz symmetry that we started with, then these states will have to be massless. And this is only the case if the dimension of spacetime is

$$D = 26 \ .$$

This is our first derivation of the critical dimension of the bosonic string.

Moreover, we've found that our theory contains a bunch of massless particles. And massless particles are interesting because they give rise to long range forces. Let's look more closely at what massless particles the string has given us. The states (2.28) transform in the $\mathbf{24} \otimes \mathbf{24}$ representation of $SO(24)$. These decompose into three irreducible representations:

$$\text{traceless symmetric} \;\oplus\; \text{anti-symmetric} \;\oplus\; \text{singlet (=trace)}$$

To each of these modes, we associate a massless field in spacetime such that the string oscillation can be identified with a quantum of these fields. The fields are:

$$G_{\mu\nu}(X) \quad , \quad B_{\mu\nu}(X) \quad , \quad \Phi(X) \tag{2.29}$$

Of these, the first is the most interesting and we shall have more to say momentarily. The second is an anti-symmetric tensor field which is usually called the anti-symmetric tensor field. It also goes by the names of the "Kalb-Ramond field" or, in the language of differential geometry, the "2-form". The scalar field is called the *dilaton*. These three massless fields are common to all string theories. We'll learn more about the role these fields play later in the course.

The particle in the symmetric traceless representation of $SO(24)$ is particularly interesting. This is a massless spin 2 particle. However, there are general arguments, due originally to Feynman and Weinberg, that *any* theory of interacting massless spin two particles must be equivalent to general relativity[2]. We should therefore identify the field $G_{\mu\nu}(X)$ with the metric of spacetime. Let's pause briefly to review the thrust of these arguments.

Why Massless Spin 2 = General Relativity

Let's call the spacetime metric $G_{\mu\nu}(X)$. We can expand around flat space by writing

$$G_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}(X) \ .$$

Then the Einstein-Hilbert action has an expansion in powers of h. If we truncate to quadratic order, we simply have a free theory which we may merrily quantize in the usual canonical fashion: we promote $h_{\mu\nu}$ to an operator and introduce the associated creation and annihilation operators $a_{\mu\nu}$ and $a^\dagger_{\mu\nu}$. This way of looking at gravity is anathema to those raised in the geometrical world of general relativity. But from a particle physics language it is very standard: it is simply the quantization of a massless spin 2 field, $h_{\mu\nu}$.

[2] A very readable description of this can be found in the first few chapters of the Feynman Lectures on Gravitation.

However, even on this simple level, there is a problem due to the indefinite signature of the spacetime Minkowski metric. The canonical quantization relations of the creation and annihilation operators are schematically of the form,

$$[a_{\mu\nu}, a^\dagger_{\rho\sigma}] \sim \eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho}$$

But this will lead to a Hilbert space with negative norm states coming from acting with time-like creation operators. For example, the one-graviton state of the form,

$$a^\dagger_{0i}|0\rangle \tag{2.30}$$

suffers from a negative norm. This should be becoming familiar by now: it is the usual problem that we run into if we try to covariantly quantize a gauge theory. And, indeed, general relativity is a gauge theory. The gauge transformations are diffeomorphisms. We would hope that this saves the theory of quantum gravity from these negative norm states.

Let's look a little more closely at what the gauge symmetry looks like for small fluctuations $h_{\mu\nu}$. We've butchered the Einstein-Hilbert action and left only terms quadratic in h. Including all the index contractions, we find

$$S_{EH} = \frac{M_{pl}^2}{2}\int d^4x\ \left[\partial_\mu h^\rho{}_\rho \partial_\nu h^{\mu\nu} - \partial^\rho h^{\mu\nu}\partial_\mu h_{\rho\nu} + \frac{1}{2}\partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu} - \frac{1}{2}\partial_\mu h^\nu{}_\nu \partial^\mu h^\rho{}_\rho\right] + \ldots$$

One can check that this truncated action is invariant under the gauge symmetry,

$$h_{\mu\nu} \longrightarrow h_{\mu\nu} + \partial_\mu\xi_\nu + \partial_\nu\xi_\mu \tag{2.31}$$

for any function $\xi_\mu(X)$. The gauge symmetry is the remnant of diffeomorphism invariance, restricted to small deviations away from flat space. With this gauge invariance in hand one can show that, just like QED, the negative norm states decouple from all physical processes.

To summarize, theories of massless spin 2 fields only make sense if there is a gauge symmetry to remove the negative norm states. In general relativity, this gauge symmetry descends from diffeomorphism invariance. The argument of Feynman and Weinberg now runs this logic in reverse. It goes as follows: suppose that we have a massless, spin 2 particle. Then, at the linearized level, it must be invariant under the gauge symmetry (2.31) in order to eliminate the negative norm states. Moreover, this symmetry must survive when interaction terms are introduced. But the only way to do this is to ensure that the resulting theory obeys diffeomorpism invariance. That means the theory of any interacting, massless spin 2 particle is Einstein gravity, perhaps supplemented by higher derivative terms.

We haven't yet shown that string theory includes interactions for $h_{\mu\nu}$ but we will come to this later in the course. More importantly, we will also explicitly see how Einstein's field equations arise directly in string theory.

A Comment on Spacetime Gauge Invariance

We've surreptitiously put $\mu, \nu = 0, \ldots, 25$ indices on the spacetime fields, rather than $i, j = 1, \ldots, 24$. The reason we're allowed to do this is because both $G_{\mu\nu}$ and $B_{\mu\nu}$ enjoy a spacetime gauge symmetry which allows us to eliminate appropriate modes. Indeed, this is exactly the gauge symmetry (2.31) that entered the discussion above. It isn't possible to see these spacetime gauge symmetries from the lightcone formalism of the string since, by construction, we find only the physical states (although, by consistency alone, the gauge symmetries must be there). One of the main advantages of pushing through with the covariant calculation is that it does allow us to see how the spacetime gauge symmetry emerges from the string worldsheet. Details can be found in Green, Schwarz and Witten. We'll also briefly return to this issue in Section 5.

2.3.3 Higher Excited States

We rescued the Lorentz invariance of the first excited states by choosing $D = 26$ to ensure that they are massless. But now we've used this trick once, we still have to worry about all the other excited states. These also carry indices that take the range $i, j = 1, \ldots, D-2 = 24$ and, from the mass formula (2.26), they will all be massive and so must form representations of $SO(D-1)$. It looks like we're in trouble again.

Let's examine the string at level $N = \tilde{N} = 2$. In the right-moving sector, we now have two different states: $\alpha^i_{-1}\alpha^j_{-1}|0\rangle$ and $\alpha^i_{-2}|0\rangle$. The same is true for the left-moving sector, meaning that the total set of states at level 2 is (in notation that is hopefully obvious, but probably technically wrong)

$$(\alpha^i_{-1}\alpha^j_{-1} \oplus \alpha^i_{-2}) \otimes (\tilde{\alpha}^i_{-1}\tilde{\alpha}^j_{-1} \oplus \tilde{\alpha}^i_{-2})|0; p\rangle \ .$$

These states have mass $M^2 = 4/\alpha'$. How many states do we have? In the left-moving sector, we have,

$$\tfrac{1}{2}(D-2)(D-1) + (D-2) = \tfrac{1}{2}D(D-1) - 1 \ .$$

But, remarkably, that does fit nicely into a representation of $SO(D-1)$, namely the traceless symmetric tensor representation.

In fact, one can show that all excited states of the string fit nicely into $SO(D-1)$ representations. The only consistency requirement that we need for Lorentz invariance is to fix up the first excited states: $D = 26$.

Note that if we are interested in a fundamental theory of quantum gravity, then all these excited states will have masses close to the Planck scale so are unlikely to be observable in particle physics experiments. Nonetheless, as we shall see when we come to discuss scattering amplitudes, it is the presence of this infinite tower of states that tames the ultra-violet behaviour of gravity.

2.4 Lorentz Invariance Revisited

The previous discussion allowed to us to derive both the critical dimension and the spectrum of string theory in the quickest fashion. But the derivation creaks a little in places. The calculation of the Casimir energy is unsatisfactory the first time one sees it. Similarly, the explanation of the need for massless particles at the first excited level is correct, but seems rather cheap considering the huge importance that we're placing on the result.

As I've mentioned a few times already, we'll shortly do better and gain some physical insight into these issues, in particular the critical dimension. But here I would just like to briefly sketch how one can be a little more rigorous within the framework of lightcone quantization. The question, as we've seen, is whether one preserves spacetime Lorentz symmetry when we quantize in lightcone gauge. We can examine this more closely.

Firstly, let's go back to the action for free scalar fields (1.30) before we imposed lightcone gauge fixing. Here the full Poincaré symmetry was manifest: it appears as a global symmetry on the worldsheet,

$$X^\mu \to \Lambda^\mu{}_\nu X^\nu + c^\mu \tag{2.32}$$

But recall that in field theory, global symmetries give rise to Noether currents and their associated conserved charges. What are the Noether currents associated to this Poincaré transformation? We can start with the translations $X^\mu \to X^\mu + c^\mu$. A quick computation shows that the current is,

$$P^\alpha_\mu = T\partial^\alpha X_\mu \tag{2.33}$$

which is indeed a conserved current since $\partial_\alpha P^\alpha_\mu = 0$ is simply the equation of motion. Similarly, we can compute the $\frac{1}{2}D(D-1)$ currents associated to Lorentz transformations. They are,

$$J^\alpha_{\mu\nu} = P^\alpha_\mu X_\nu - P^\alpha_\nu X_\mu$$

It's not hard to check that $\partial_\alpha J^\alpha_{\mu\nu} = 0$ when the equations of motion are obeyed.

The conserved charges arising from this current are given by $M_{\mu\nu} = \int d\sigma J^{\tau}_{\mu\nu}$. Using the mode expansion (1.36) for X^μ, these can be written as

$$\begin{aligned}\mathcal{M}^{\mu\nu} &= (p^\mu x^\nu - p^\nu x^\mu) - i\sum_{n=1}^{\infty}\frac{1}{n}\left(\alpha^\nu_{-n}\alpha^\mu_n - \alpha^\mu_{-n}\alpha^\nu_n\right) - i\sum_{n=1}^{\infty}\frac{1}{n}\left(\tilde\alpha^\nu_{-n}\tilde\alpha^\mu_n - \tilde\alpha^\mu_{-n}\tilde\alpha^\nu_n\right)\\ &\equiv l^{\mu\nu} + S^{\mu\nu} + \tilde S^{\mu\nu}\end{aligned}$$

The first piece, $l^{\mu\nu}$, is the orbital angular momentum of the string while the remaining pieces $S^{\mu\nu}$ and $\tilde S^{\mu\nu}$ tell us the angular momentum due to excited oscillator modes. Classically, these obey the Poisson brackets of the Lorentz algebra. Moreover, if we quantize in the covariant approach, the corresponding operators obey the commutation relations of the Lorentz Lie algebra, namely

$$[\mathcal{M}^{\rho\sigma}, \mathcal{M}^{\tau\nu}] = \eta^{\sigma\tau}\mathcal{M}^{\rho\nu} - \eta^{\rho\tau}\mathcal{M}^{\sigma\nu} + \eta^{\rho\nu}\mathcal{M}^{\sigma\tau} - \eta^{\sigma\nu}\mathcal{M}^{\rho\tau}$$

However, things aren't so easy in lightcone gauge. Lorentz invariance is not guaranteed and, in general, is not there. The right way to go about looking for it is to make sure that the Lorentz algebra above is reproduced by the generators $\mathcal{M}^{\mu\nu}$. It turns out that the smoking gun lies in the commutation relation,

$$[\mathcal{M}^{i-}, \mathcal{M}^{j-}] = 0$$

Does this equation hold in lightcone gauge? The problem is that it involves the operators p^- and α^-_n, both of which are fixed by (2.17) and (2.18) in terms of the other operators. So the task is to compute this commutation relation $[\mathcal{M}^{i-}, \mathcal{M}^{j-}]$, given the commutation relations (2.21) for the physical degrees of freedom, and check that it vanishes. To do this, we re-instate the ordering ambiguity a and the number of spacetime dimension D as arbitrary variables and proceed.

The part involving orbital angular momenta l^{i-} is fairly straightforward. (Actually, there's a small subtlety because we must first make sure that the operator $l^{\mu\nu}$ is Hermitian by replacing $x^\mu p^\nu$ with $\frac{1}{2}(x^\mu p^\nu + p^\nu x^\mu)$). The real difficulty comes from computing the commutation relations $[S^{i-}, S^{j-}]$. This is messy[3]. After a tedious computation, one finds,

$$[\mathcal{M}^{i-}, \mathcal{M}^{j-}] = \frac{2}{(p^+)^2}\sum_{n>0}\left(\left[\frac{D-2}{24} - 1\right]n + \frac{1}{n}\left[a - \frac{D-2}{24}\right]\right)\left(\alpha^i_{-n}\alpha^j_n - \alpha^j_{-n}\alpha^i_n\right) + (\alpha \leftrightarrow \tilde\alpha)$$

[3]The original, classic, paper where lightcone quantization was first implemented is Goddard, Goldstone, Rebbi and Thorn "*Quantum Dynamics of a Massless Relativistic String*", Nucl. Phys. B56 (1973). A pedestrian walkthrough of this calculation can be found in the lecture notes by Gleb Arutyunov. A link is given on the course webpage.

The right-hand side does not, in general, vanish. We learn that the relativistic string can only be quantized in flat Minkowski space if we pick,

$$D = 26 \quad \text{and} \quad a = 1 \; .$$

2.5 A Nod to the Superstring

We won't provide details of the superstring in this course, but will pause occasionally to make some pertinent comments. Although what follows is nothing more than a list of facts, it will hopefully be helpful in orienting you when you do come to study this material.

The key difference between the bosonic string and the superstring is the addition of fermionic modes on its worldsheet. The resulting worldsheet theory is supersymmetric. (At least in the so-called Neveu-Schwarz-Ramond formalism). Hence the name "superstring". Applying the kind of quantization procedure we've discussed in this section, one finds the following results:

- The critical dimension of the superstring is $D = 10$.
- There is no tachyon in the spectrum.
- The massless bosonic fields $G_{\mu\nu}$, $B_{\mu\nu}$ and Φ are all part of the spectrum of the superstring. In this context, $B_{\mu\nu}$ is sometimes referred to as the Neveu-Schwarz 2-form. There are also massless spacetime fermions, as well as further massless bosonic fields. As we now discuss, the exact form of these extra bosonic fields depends on exactly what superstring theory we consider.

While the bosonic string is unique, there are a number of discrete choices that one can make when adding fermions to the worldsheet. This gives rise to a handful of different perturbative superstring theories. (Although later developments reveal that they are actually all part of the same framework which sometimes goes by the name of *M-theory*). The most important of these discrete options is whether we add fermions in both the left-moving and right-moving sectors of the string, or whether we choose the fermions to move only in one direction, usually taken to be right-moving. This gives rise to two different classes of string theory.

- Type II strings have both left and right-moving worldsheet fermions. The resulting spacetime theory in $D = 10$ dimensions has $\mathcal{N} = 2$ supersymmetry, which means 32 supercharges.
- Heterotic strings have just right-moving fermions. The resulting spacetime theory has $\mathcal{N} = 1$ supersymmetry, or 16 supercharges.

In each of these cases, there is then one further discrete choice that we can make. This leaves us with four superstring theories. In each case, the massless bosonic fields include $G_{\mu\nu}$, $B_{\mu\nu}$ and Φ together with a number of extra fields. These are:

- **Type IIA**: In the type II theories, the extra massless bosonic excitations of the string are referred to as *Ramond-Ramond* fields. For Type IIA, they are a 1-form C_μ and a 3-form $C_{\mu\nu\rho}$. Each of these is to be thought of as a gauge field. The gauge invariant information lies in the field strengths which take the form $F = dC$.

- **Type IIB**: The Ramond-Ramond gauge fields consist of a scalar C, a 2-form $C_{\mu\nu}$ and a 4-form $C_{\mu\nu\rho\sigma}$. The 4-form is restricted to have a self-dual field strength: $F_5 = {}^\star F_5$. (Actually, this statement is almost true...we'll look a little closer at this in Section 7.3.3).

- **Heterotic** $SO(32)$: The heterotic strings do not have Ramond-Ramond fields. Instead, each comes with a non-Abelian gauge field in spacetime. The heterotic strings are named after the gauge group. For example, the Heterotic $SO(32)$ string gives rise to an $SO(32)$ Yang-Mills theory in ten dimensions.

- **Heterotic** $E_8 \times E_8$: The clue is in the name. This string gives rise to an $E_8 \times E_8$ Yang-Mills field in ten-dimensions.

It is sometimes said that there are five perturbative superstring theories in ten dimensions. Here we've only mentioned four. The remaining theory is called Type I and includes open strings moving in flat ten dimensional space as well as closed strings. We'll mention it in passing in the following section.

3. Open Strings and D-Branes

In this section we discuss the dynamics of open strings. Clearly their distinguishing feature is the existence of two end points. Our goal is to understand the effect of these end points. The spatial coordinate of the string is parameterized by

$$\sigma \in [0, \pi] \ .$$

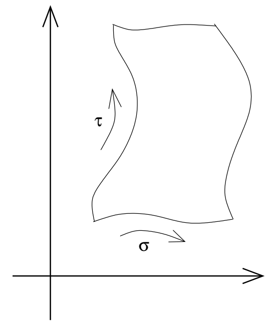


Figure 12:

The dynamics of a generic point on a string is governed by local physics. This means that a generic point has no idea if it is part of a closed string or an open string. The dynamics of an open string must therefore still be described by the Polyakov action. But this must now be supplemented by something else: boundary conditions to tell us how the end points move. To see this, let's look at the Polyakov action in conformal gauge

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

As usual, we derive the equations of motion by finding the extrema of the action. This involves an integration by parts. Let's consider the string evolving from some initial configuration at $\tau = \tau_i$ to some final configuration at $\tau = \tau_f$:

$$\begin{aligned} \delta S &= -\frac{1}{2\pi\alpha'} \int_{\tau_i}^{\tau_f} d\tau \int_0^\pi d\sigma \;\; \partial_\alpha X \cdot \partial^\alpha \delta X \\ &= \frac{1}{2\pi\alpha'} \int d^2\sigma \; (\partial^\alpha \partial_\alpha X) \cdot \delta X + \text{total derivative} \end{aligned}$$

For an open string the total derivative picks up the boundary contributions

$$\frac{1}{2\pi\alpha'} \left[\int_0^\pi d\sigma \; \dot{X} \cdot \delta X \right]_{\tau=\tau_i}^{\tau=\tau_f} - \frac{1}{2\pi\alpha'} \left[\int_{\tau_i}^{\tau_f} d\tau \; X' \cdot \delta X \right]_{\sigma=0}^{\sigma=\pi}$$

The first term is the kind that we always get when using the principle of least action. The equations of motion are derived by requiring that $\delta X^\mu = 0$ at $\tau = \tau_i$ and τ_f and so it vanishes. However, the second term is novel. In order for it too to vanish, we require

$$\partial_\sigma X^\mu \, \delta X_\mu = 0 \qquad \text{at } \sigma = 0, \pi$$

There are two different types of boundary conditions that we can impose to satisfy this:

- Neumann boundary conditions.

$$\partial_\sigma X^\mu = 0 \qquad \text{at } \sigma = 0, \pi \tag{3.1}$$

Because there is no restriction on δX^μ, this condition allows the end of the string to move freely. To see the consequences of this, it's useful to repeat what we did for the closed string and work in static gauge with $X^0 \equiv t = R\tau$, for some dimensionful constant R. Then, as in equations (1.34), the constraints read

$$\dot{\vec{x}} \cdot \vec{x}\,' = 0 \quad \text{and} \quad \dot{\vec{x}}^{\,2} + \vec{x}\,'^{\,2} = R^2$$

But at the end points of the string, $\vec{x}\,' = 0$. So the second equation tells us that $|d\vec{x}/dt| = 1$. Or, in other words, the end point of the string moves at the speed of light.

- Dirichlet boundary conditions

$$\delta X^\mu = 0 \qquad \text{at } \sigma = 0, \pi \tag{3.2}$$

This means that the end points of the string lie at some constant position, $X^\mu = c^\mu$, in space.

At first sight, Dirichlet boundary conditions may seem a little odd. Why on earth would the strings be fixed at some point c^μ? What is special about that point? Historically people were pretty hung up about this and Dirichlet boundary conditions were rarely considered until the mid-1990s. Then everything changed due to an insight of Polchinski...

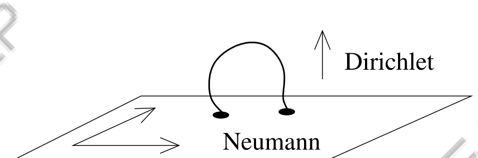


Figure 13:

Let's consider Dirichlet boundary conditions for some coordinates and Neumann for the others. This means that at both end points of the string, we have

$$\begin{aligned} \partial_\sigma X^a &= 0 \qquad \text{for } \; a = 0, \ldots, p \\ X^I &= c^I \qquad \text{for } \; I = p+1, \ldots, D-1 \end{aligned} \tag{3.3}$$

This fixes the end-points of the string to lie in a $(p+1)$-dimensional hypersurface in spacetime such that the $SO(1, D-1)$ Lorentz group is broken to,

$$SO(1, D-1) \to SO(1, p) \times SO(D-p-1) \ .$$

This hypersurface is called a *D-brane* or, when we want to specify its dimension, a Dp-brane. Here D stands for Dirichlet, while p is the number of spatial dimensions of the brane. So, in this language, a D0-brane is a particle; a D1-brane is itself a string; a D2-brane a membrane and so on. The brane sits at specific positions c^I in the transverse space. But what is the interpretation of this hypersurface?

It turns out that the D-brane hypersurface should be thought of as a new, dynamical object in its own right. This is a conceptual leap that is far from obvious. Indeed, it took decades for people to fully appreciate this fact. String theory is not just a theory of strings: it also contains higher dimensional branes. In Section 7.5 we will see how these D-branes develop a life of their own. Some comments:

- We've defined D-branes that are infinite in space. However, we could just as well define finite D-branes by specifying closed surfaces on which the string can end.
- There are many situations where we want to describe strings that have Neumann boundary conditions in all directions, meaning that the string is free to move throughout spacetime. It's best to understand this in terms of a space-filling D-brane. No Dirichlet conditions means D-branes are everywhere!
- The Dp-brane described above always has Neumann boundary conditions in the X^0 direction. What would it mean to have Dirichlet conditions for X^0? Obviously this is a little weird since the object is now localized at a fixed point in time. But there is an interpretation of such an object: it is an *instanton*. This "D-instanton" is usually referred to as a D(-1)-brane. It is related to tunneling effects in the quantum theory.

Mode Expansion

We take the usual mode expansion for the string, with $X^\mu = X^\mu_L(\sigma^+) + X^\mu_R(\sigma^-)$ and

$$X^\mu_L(\sigma^+) = \tfrac{1}{2}x^\mu + \alpha' p^\mu\,\sigma^+ + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\,\tilde{\alpha}^\mu_n\,e^{-in\sigma^+}\ ,$$

$$X^\mu_R(\sigma^-) = \tfrac{1}{2}x^\mu + \alpha' p^\mu\,\sigma^- + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\,\alpha^\mu_n\,e^{-in\sigma^-}\ . \tag{3.4}$$

The boundary conditions impose relations on the modes of the string. They are easily checked to be:

- Neumann boundary conditions, $\partial_\sigma X^a = 0$, at the end points require that
$$\alpha^a_n = \tilde{\alpha}^a_n \tag{3.5}$$
- Dirichlet boundary conditions, $X^I = c^I$, at the end points require that
$$x^I = c^I \quad , \quad p^I = 0 \quad , \quad \alpha^I_n = -\tilde{\alpha}^I_n$$

So for both boundary conditions, we only have one set of oscillators, say α_n. The $\tilde{\alpha}_n$ are then determined by the boundary conditions.

It's worth pointing out that there is a factor of 2 difference in the p^μ term between the open string (3.4) and the closed string (1.36). This is to ensure that p^μ for the open string retains the interpretation of the spacetime momentum of the string when $\sigma \in [0,\pi]$. To see this, one needs to check the Noether current associated to translations of X^μ on the worldsheet: it was given in (2.33). The conserved charge is then

$$P^\mu = \int_0^\pi d\sigma\, (P^\tau)^\mu = \frac{1}{2\pi\alpha'}\int_0^\pi d\sigma\, \dot{X}^\mu = p^\mu$$

as advertised. Note that we've needed to use the Neumann conditions (3.5) to ensure that the Fourier modes don't contribute to this integral.

3.1 Quantization

To quantize, we promote the fields x^a and p^a and α_n^μ to operators. The other elements in the mode expansion are fixed by the boundary conditions. An obvious, but important, point is that the position and momentum degrees of freedom, x^a and p^a, have a spacetime index that takes values $a = 0,\ldots,p$. This means that the spatial wavefunctions only depend on the coordinates of the brane not the whole spacetime. Said another, quantizing an open string gives rise to states which are restricted to lie on the brane.

To determine the spectrum, it is again simplest to work in lightcone gauge. The spacetime lightcone coordinate is chosen to lie within the brane,

$$X^\pm = \sqrt{\frac{1}{2}}\,(X^0 \pm X^p)$$

Quantization now proceeds in the same manner as for the closed string until we arrive at the mass formula for states which is a sum over the transverse modes of the string.

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

The first sum is over modes parallel to the brane, the second over modes perpendicular to the brane. It's worth commenting on the differences with the closed string formula. Firstly, there is an overall factor of 4 difference. This can be traced to the lack of the factor of $1/2$ in front of p^μ in the mode expansion that we discussed above. Secondly, there is a sum only over α modes. The $\tilde{\alpha}$ modes are not independent because of the boundary conditions.

Open and Closed

In the mass formula, we have once again left the normal ordering constant a ambiguous. As in the closed string case, requiring the Lorentz symmetry of the quantum theory — this time the reduced symmetry $SO(1,p) \times SO(D-p-1)$ — forces us to choose

$$D = 26 \quad \text{and} \quad a = 1 \; .$$

These are the same values that we found for the closed string. This reflects an important fact: the open string and closed string are not different theories. They are both different states inside the same theory.

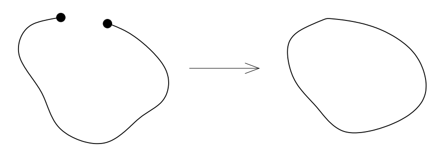


Figure 14:

More precisely, theories of open strings necessarily contain closed strings. This is because, once we consider interactions, an open string can join to form a closed string as shown in the figure. We'll look at interactions in Section 6. The question of whether this works the other way — meaning whether closed string theories require open strings — is a little more involved and is cleanest to state in the context of the superstring. For type II superstrings, the open strings and D-branes are necessary ingredients. For heterotic superstrings, there appear to be no open strings and no D-branes. For the bosonic theory, it seems likely that the open strings are a necessary ingredient although I don't know of a killer argument. But since we're not sure whether the theory exists due to the presence of the tachyon, the point is probably moot. In the remainder of these lectures, we'll view the bosonic string in the same manner as the type II string and assume that the theory includes both closed strings and open strings with their associated D-branes.

3.1.1 The Ground State

The ground state is defined by

$$\alpha_n^i |0;p\rangle = 0 \qquad\qquad n > 0$$

The spatial index now runs over $i = 1, \ldots, p-1, p+1, \ldots, D-1$. The ground state has mass

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

It is again tachyonic. Its mass is half that of the closed string tachyon. As we commented above, this time the tachyon is confined to the brane. In contrast to the closed string tachyon, the open string tachyon is now fairly well understood and its potential is of the form shown in the figure. The interpretation is that the brane is unstable. It will decay, much like a resonance state in field theory. It does this by dissolving into closed string modes. The end point of this process – corresponding to the minimum at $T > 0$ in the figure – is simply a state with no D-brane. The difference between the value of the potential at the minimum and at $T = 0$ is the tension of the D-brane.

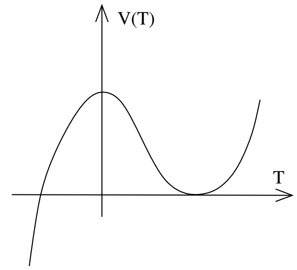


Figure 15:

Notice that although there is a minimum of the potential at $T > 0$, it is not a global minimum. The potential seems to drop off without bound to the left. This is still not well understood. There are suggestions that it is related in some way to the closed string tachyon.

3.1.2 First Excited States: A World of Light

The first excited states are massless. They fall into two classes:

- Oscillators longitudinal to the brane,

$$\alpha^a_{-1}|0;p\rangle \qquad a = 1,\ldots,p-1$$

  The spacetime indices a lie within the brane so this state transforms under the $SO(1,p)$ Lorentz group. It is a spin 1 particle on the brane or, in other words, it is a photon. We introduce a gauge field A_a with $a = 0,\ldots,p$ lying on the brane whose quanta are identified with this photon.

- Oscillators transverse to the brane,

$$\alpha^I_{-1}|0;p\rangle \qquad I = p+1,\ldots,D-1$$

  These states are scalars under the $SO(1,p)$ Lorentz group of the brane. They can be thought of as arising from scalar fields ϕ^I living on the brane. These scalars have a nice interpretation: they are fluctuations of the brane in the transverse directions. This is our first hint that the D-brane is a dynamical object. Note that although the ϕ^I are scalar fields under the $SO(1,p)$ Lorentz group of the brane, they do transform as a vector under the $SO(D-p-1)$ rotation group transverse to the brane. This appears as a global symmetry on the brane worldvolume.

3.1.3 Higher Excited States and Regge Trajectories

At level N, the mass of the string state is

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

The maximal spin of these states arises from the symmetric tensor. It is

$$J_{max} = N = \alpha' M^2 + 1$$

Plotting the spin vs. the mass-squared, we find straight lines. These are usually called *Regge trajectories*. (Or sometimes Chew-Frautschi trajectories). They are seen in Nature in both the spectrum of mesons and baryons. Some examples involving ρ-mesons are shown in the figure. These stringy Regge trajectories suggest a naive cartoon picture of mesons as two rotating quarks connected by a confining flux tube.

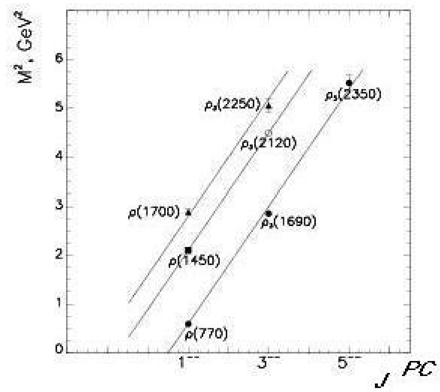


Figure 16:

The value of the string tension required to match the hadron spectrum of QCD is $T \sim 1$ GeV. This relationship between the strong interaction and the open string was one of the original motivations for the development of string theory and it is from here that the parameter α' gets its (admittedly rarely used) name "Regge slope". In these enlightened modern times, the connection between the open string and quarks lives on in the AdS/CFT correspondence.

3.1.4 Another Nod to the Superstring

Just as supersymmetry eliminates the closed string tachyon, so it removes the open string tachyon. Open strings are an ingredient of the type II string theories. The possible D-branes are

- Type IIA string theory has stable Dp-branes with p even.
- Type IIB string theory has stable Dp-branes with p odd.

The most important reason that D-branes are stable in the type II string theories is that they are charged under the Ramond-Ramond fields. (This was actually Polchinski's insight that made people take D-branes seriously). However, type II string theories also contain unstable branes, with p odd in type IIA and p even in type IIB.

The fifth string theory (which was actually the first to be discovered) is called Type I. Unlike the other string theories, it contains both open and closed strings moving in flat ten-dimensional Lorentz-invariant spacetime. It can be thought of as the Type IIB theory with a bunch of space-filling D9-branes, together with something called an orientifold plane. You can read about this in Polchinski.

As we mentioned above, the heterotic string doesn't have (finite energy) D-branes. This is due to an inconsistency in any attempt to reflect left-moving modes into right-moving modes.

3.2 Brane Dynamics: The Dirac Action

We have introduced D-branes as fixed boundary conditions for the open string. However, we've already seen a hint that these objects are dynamical in their own right, since the massless scalar excitations ϕ^I have a natural interpretation as transverse fluctuations of the brane. Indeed, if a theory includes both open strings and closed strings, then the D-branes have to be dynamical because there can be no rigid objects in a theory of gravity. The dynamical nature of D-branes will become clearer as the course progresses.

But any dynamical object should have an action which describes how it moves. Moreover, after our discussion in Section 1, we already know what this is! On grounds of Lorentz invariance and reparameterization invariance alone, the action must be a higher dimensional extension of the Nambu-Goto action. This is

$$S_{Dp} = -T_p \int d^{p+1}\xi \, \sqrt{-\det \gamma} \tag{3.6}$$

where T_p is the tension of the Dp-brane which we will determine later, while ξ^a, $a = 0, \ldots, p$, are the worldvolume coordinates of the brane. γ_{ab} is the pull back of the spacetime metric onto the worldvolume,

$$\gamma_{ab} = \frac{\partial X^\mu}{\partial \xi^a} \frac{\partial X^\nu}{\partial \xi^b} \eta_{\mu\nu} \ .$$

This is called the *Dirac action*. It was first written down by Dirac for a membrane some time before Nambu and Goto rediscovered it in the context of the string.

To make contact with the fields ϕ^I, we can use the reparameterization invariance of the Dirac action to go to static gauge. For an infinite, flat Dp-brane we can choose

$$X^a = \xi^a \qquad a = 0, \ldots, p \ .$$

The dynamical transverse coordinates are then identified with the fluctuations ϕ^I through

$$X^I(\xi) = 2\pi\alpha'\,\phi^I(\xi) \qquad I = p+1,\ldots,D-1$$

However, the Dirac action can't be the whole story. It describes the transverse fluctuations of the D-brane, but has nothing to say about the $U(1)$ gauge field A_μ which lives on the D-brane. There must be some action which describes how this gauge field moves as well. We will return to this in Section 7.

What's Special About Strings?

We could try to quantize the Dirac action (3.6) for a D-brane in the same manner that we quantized the action for the string. Is this possible? The answer, at present, is no. There appear to be both technical and conceptual obstacles . The technical issue is just that it's hard. Weyl invariance was one of our chief weapons in attacking the string, but it doesn't hold for higher dimensional objects.

The conceptual issue is that quantizing a membrane, or higher dimensional object, would not give rise to a discrete spectrum of states which have the interpretation of particles. In this way, they appear to be fundamentally different from the string.

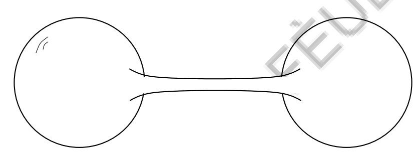


Figure 17:

Let's get some intuition for why this is the case. The energy of a string is proportional to its length. This ensures that strings behave more or less like familiar elastic bands. What about D2-branes? Now the energy is proportional to the area. In the back of your mind, you might be thinking of a rubber-like sheet. But membranes, and higher dimensional objects, governed by the Dirac action don't behave as household rubber sheets. They are more flexible. This is because a membrane can form many different shapes with the same area. For example, a tubular membrane of length L and radius $1/L$ has the same area for all values of L; short and stubby, or long and thin. This means that long thin spikes can develop on a membrane at no extra cost of energy. In particular, objects connected by long thin tubes have the same energy, regardless of their separation. After quantization, this property gives rise to a continuous spectrum of states. A quantum membrane, or higher dimensional object, does not have the single particle interpretation that we saw for the string. The expectation is that the quantum membrane should describe multi-particle states.

3.3 Multiple Branes: A World of Glue

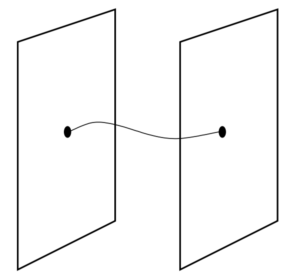


Figure 18:

Consider two parallel Dp-branes. An open string now has options. It could either end on the same brane, or stretch between the two branes. Let's consider the string that stretches between the two. It obeys

$$X^I(0,\tau) = c^I \quad \text{and} \quad X^I(\pi,\tau) = d^I$$

where c^I and d^I are the positions of the two branes. In terms of the mode expansion, this requires

$$X^I = c^I + \frac{(d^I - c^I)\sigma}{\pi} + \text{oscillator modes}$$

The classical constraints then read

$$\partial_+ X \cdot \partial_+ X = \alpha'^2 p^2 + \frac{|\vec{d} - \vec{c}|^2}{4\pi^2} + \text{oscillator modes} = 0$$

which means the classical mass-shell condition is

$$M^2 = \frac{|\vec{d} - \vec{c}|^2}{(2\pi\alpha')^2} + \text{oscillator modes}$$

The extra term has an obvious interpretation: it is the mass of a classical string stretched between the two branes. The quantization of this string proceeds as before. After we include the normal ordering constant, the ground state of this string is only tachyonic if $|\vec{d} - \vec{c}|^2 < 4\pi^2\alpha'$. Or in other words, the ground state is tachyonic if the branes approach to a sub-stringy distance.

There is an obvious generalization of this to the case of N parallel branes. Each end point of the string has N possible places on which to end. We can label each end point with a number $m, n = 1, \ldots, N$ which tell us which brane it ends on. This label is sometimes referred to as a *Chan-Paton factor*.

Consider now the situation where all branes lie at the same position in spacetime. Each end point can lie on one of N different branes, giving N^2 possibilities in total. Each of these strings has the mass spectrum of an open string, meaning that there are now N^2 different particles of each type. It's natural to arrange the associated fields to sit inside $N \times N$ Hermitian matrices. We then have the open string tachyon $T^m_{\ n}$ and the massless fields

$$(\phi^I)^m_{\ n} \quad , \quad (A_a)^m_{\ n} \tag{3.7}$$

Here the components of the matrix tell us which string the field came from. Diagonal components arise from strings which have both ends on the same brane.

The gauge field A_a is particularly interesting. Written in this way, it looks like a $U(N)$ gauge connection. We will later see that this is indeed the case. One can show that as N branes coincide, the $U(1)^N$ gauge symmetry of the branes is enhanced to $U(N)$. The scalar fields ϕ^I transform in the adjoint of this symmetry.

4. Introducing Conformal Field Theory

The purpose of this section is to get comfortable with the basic language of two dimensional conformal field theory[4]. This is a topic which has many applications outside of string theory, most notably in statistical physics where it offers a description of critical phenomena. Moreover, it turns out that conformal field theories in two dimensions provide rare examples of interacting, yet exactly solvable, quantum field theories. In recent years, attention has focussed on conformal field theories in higher dimensions due to their role in the AdS/CFT correspondence.

A *conformal transformation* is a change of coordinates $\sigma^\alpha \to \tilde{\sigma}^\alpha(\sigma)$ such that the metric changes by

$$g_{\alpha\beta}(\sigma) \to \Omega^2(\sigma) g_{\alpha\beta}(\sigma) \tag{4.1}$$

A *conformal field theory* (CFT) is a field theory which is invariant under these transformations. This means that the physics of the theory looks the same at all length scales. Conformal field theories care about angles, but not about distances.

A transformation of the form (4.1) has a different interpretation depending on whether we are considering a fixed background metric $g_{\alpha\beta}$, or a dynamical background metric. When the metric is dynamical, the transformation is a diffeomorphism; this is a gauge symmetry. When the background is fixed, the transformation should be thought of as an honest, physical symmetry, taking the point σ^α to point $\tilde{\sigma}^\alpha$. This is now a global symmetry with the corresponding conserved currents.

In the context of string theory in the Polyakov formalism, the metric is dynamical and the transformations (4.1) are residual gauge transformations: diffeomorphisms which can be undone by a Weyl transformation.

In contrast, in this section we will be primarily interested in theories defined on fixed backgrounds. Apart from a few noticeable exceptions, we will usually take this background to be flat. This is the situation that we are used to when studying quantum field theory.

[4]Much of the material covered in this section was first described in the ground breaking paper by Belavin, Polyakov and Zamalodchikov, "*Infinite Conformal Symmetry in Two-Dimensional Quantum Field Theory*", Nucl. Phys. B241 (1984). The application to string theory was explained by Friedan, Martinec and Shenker in "*Conformal Invariance, Supersymmetry and String Theory*", Nucl. Phys. B271 (1986). The canonical reference for learning conformal field theory is the excellent review by Ginsparg. A link can be found on the course webpage.

Of course, we can alternate between thinking of theories as defined on fixed or fluctuating backgrounds. Any theory of 2d gravity which enjoys both diffeomorphism and Weyl invariance will reduce to a conformally invariant theory when the background metric is fixed. Similarly, any conformally invariant theory can be coupled to 2d gravity where it will give rise to a classical theory which enjoys both diffeomorphism and Weyl invariance. Notice the caveat "classical"! In some sense, the whole point of this course is to understand when this last statement also holds at the quantum level.

Even though conformal field theories are a subset of quantum field theories, the language used to describe them is a little different. This is partly out of necessity. Invariance under the transformation (4.1) can only hold if the theory has no preferred length scale. But this means that there can be nothing in the theory like a mass or a Compton wavelength. In other words, conformal field theories only support massless excitations. The questions that we ask are not those of particles and S-matrices. Instead we will be concerned with correlation functions and the behaviour of different operators under conformal transformations.

4.0.1 Euclidean Space

Although we're ultimately interested in Minkowski signature worldsheets, it will be much simpler and elegant if we work instead with Euclidean worldsheets. There's no funny business here — everything we do could also be formulated in Minkowski space.

The Euclidean worldsheet coordinates are $(\sigma^1, \sigma^2) = (\sigma^1, i\sigma^0)$ and it will prove useful to form the complex coordinates,

$$z = \sigma^1 + i\sigma^2 \quad \text{and} \quad \bar{z} = \sigma^1 - i\sigma^2$$

which are the Euclidean analogue of the lightcone coordinates. Motivated by this analogy, it is common to refer to holomorphic functions as "left-moving" and anti-holomorphic functions as "right-moving".

The holomorphic derivatives are

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

These obey $\partial z = \bar{\partial}\bar{z} = 1$ and $\partial\bar{z} = \bar{\partial}z = 0$. We will usually work in flat Euclidean space, with metric

$$ds^2 = (d\sigma^1)^2 + (d\sigma^2)^2 = dz\, d\bar{z} \tag{4.2}$$

In components, this flat metric reads

$$g_{zz} = g_{\bar{z}\bar{z}} = 0 \quad \text{and} \quad g_{z\bar{z}} = \frac{1}{2}$$

With this convention, the measure factor is $dzd\bar{z} = 2d\sigma^1 d\sigma^2$. We define the delta-function such that $\int d^2z\, \delta(z,\bar{z}) = 1$. Notice that because we also have $\int d^2\sigma\, \delta(\sigma) = 1$, this means that there is a factor of 2 difference between the two delta functions. Vectors naturally have their indices up: $v^z = (v^1 + iv^2)$ and $v^{\bar{z}} = (v^1 - iv^2)$. When indices are down, the vectors are $v_z = \frac{1}{2}(v^1 - iv^2)$ and $v_{\bar{z}} = \frac{1}{2}(v^1 + iv^2)$.

4.0.2 The Holomorphy of Conformal Transformations

In the complex Euclidean coordinates z and $\bar{z}$, conformal transformations of flat space are simple: they are any holomorphic change of coordinates,

$$z \to z' = f(z) \quad \text{and} \quad \bar{z} \to \bar{z}' = \bar{f}(\bar{z})$$

Under this transformation, $ds^2 = dzd\bar{z} \;\to\; |df/dz|^2\, dzd\bar{z}$, which indeed takes the form (4.1). Note that we have an infinite number of conformal transformations — in fact, a whole functions worth $f(z)$. This is special to conformal field theories in two dimensions. In higher dimensions, the space of conformal transformations is a finite dimensional group. For theories defined on $\mathbf{R}^{p,q}$, the conformal group is $SO(p+1, q+1)$ when $p + q > 2$.

A couple of particularly simple and important examples of 2d conformal transformations are

- $z \to z + a$: This is a translation.
- $z \to \zeta z$: This is a rotation for $|\zeta| = 1$ and a scale transformation (also known as a *dilatation*) for real $\zeta \neq 1$.

For many purposes, it's simplest to treat z and $\bar{z}$ as independent variables. In doing this, we're really extending the worldsheet from $\mathbf{R}^2$ to $\mathbf{C}^2$. This will allow us to make use of various theorems from complex methods. However, at the end of the day we should remember that we're really sitting on the real slice $\mathbf{R}^2 \subset \mathbf{C}^2$ defined by $\bar{z} = z^\star$.

4.1 Classical Aspects

We start by deriving some properties of classical theories which are invariant under conformal transformations (4.1).

4.1.1 The Stress-Energy Tensor

One of the most important objects in any field theory is the *stress-energy tensor* (also known as the energy-momentum tensor). This is defined in the usual way as the matrix of conserved currents which arise from translational invariance,

$$\delta\sigma^\alpha = \epsilon^\alpha \ .$$

In flat spacetime, a translation is a special case of a conformal transformation.

There's a cute way to derive the stress-energy tensor in any theory. Suppose for the moment that we are in flat space $g_{\alpha\beta} = \eta_{\alpha\beta}$. Recall that we can usually derive conserved currents by promoting the constant parameter ϵ that appears in the symmetry to a function of the spacetime coordinates. The change in the action must then be of the form,

$$\delta S = \int d^2\sigma \ J^\alpha \, \partial_\alpha \epsilon \tag{4.3}$$

for some function of the fields, J^α. This ensures that the variation of the action vanishes when ϵ is constant, which is of course the definition of a symmetry. But when the equations of motion are satisfied, we must have $\delta S = 0$ for all variations $\epsilon(\sigma)$, not just constant ϵ. This means that when the equations of motion are obeyed, J^α must satisfy

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

The function J^α is our conserved current.

Let's see how this works for translational invariance. If we promote ϵ to a function of the worldsheet variables, the change of the action must be of the form (4.3). But what is J^α? At this point we do the cute thing. Consider the same theory, but now coupled to a dynamical background metric $g_{\alpha\beta}(\sigma)$. In other words, coupled to gravity. Then we could view the transformation

$$\delta\sigma^\alpha = \epsilon^\alpha(\sigma)$$

as a diffeomorphism and we know that the theory is invariant as long as we make the corresponding change to the metric

$$\delta g_{\alpha\beta} = \partial_\alpha \epsilon_\beta + \partial_\beta \epsilon_\alpha \ .$$

This means that if we just make the transformation of the coordinates in our original theory, then the change in the action must be the opposite of what we get if we just transform the metric. (Because doing both together leaves the action invariant). So we have

$$\delta S = -\int d^2\sigma \, \frac{\partial S}{\partial g_{\alpha\beta}} \, \delta g_{\alpha\beta} = -2 \int d^2\sigma \, \frac{\partial S}{\partial g_{\alpha\beta}} \, \partial_\alpha \epsilon_\beta$$

Note that $\partial S/\partial g_{\alpha\beta}$ in this expression is really a functional derivatives but we won't be careful about using notation to indicate this. We now have the conserved current arising from translational invariance. We will add a normalization constant which is standard in string theory (although not necessarily in other areas) and define the stress-energy tensor to be

$$T_{\alpha\beta} = -\frac{4\pi}{\sqrt{g}} \frac{\partial S}{\partial g^{\alpha\beta}} \tag{4.4}$$

If we have a flat worldsheet, we evaluate $T_{\alpha\beta}$ on $g_{\alpha\beta} = \delta_{\alpha\beta}$ and the resulting expression obeys $\partial^\alpha T_{\alpha\beta} = 0$. If we're working on a curved worldsheet, then the energy-momentum tensor is covariantly conserved, $\nabla^\alpha T_{\alpha\beta} = 0$.

The Stress-Energy Tensor is Traceless

In conformal theories, $T_{\alpha\beta}$ has a very important property: its trace vanishes. To see this, let's vary the action with respect to a scale transformation which is a special case of a conformal transformation,

$$\delta g_{\alpha\beta} = \epsilon g_{\alpha\beta} \tag{4.5}$$

Then we have

$$\delta S = \int d^2\sigma \, \frac{\partial S}{\partial g_{\alpha\beta}} \, \delta g_{\alpha\beta} = -\frac{1}{4\pi} \int d^2\sigma \sqrt{g} \, \epsilon \, T^\alpha_{\ \alpha}$$

But this must vanish in a conformal theory because scaling transformations are a symmetry. So

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

This is the key feature of a conformal field theory in any dimension. Many theories have this feature at the classical level, including Maxwell theory and Yang-Mills theory in four-dimensions. However, it is much harder to preserve at the quantum level. (The weight of the world rests on the fact that Yang-Mills theory fails to be conformal at the quantum level). Technically the difficulty arises due to the need to introduce a scale when regulating the theories. Here we will be interested in two-dimensional theories 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,

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

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,

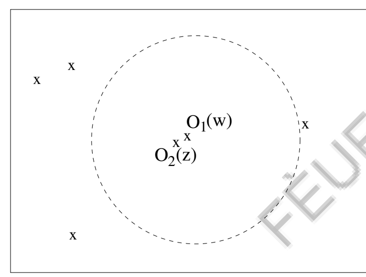


Figure 19:

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

where the $\ldots$ 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

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

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

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

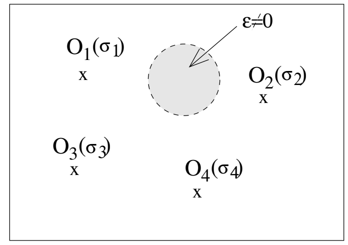


Figure 20:

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,

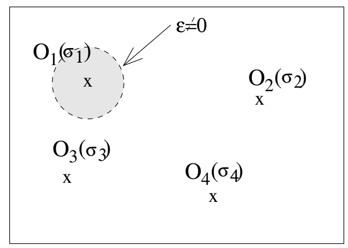


Figure 21:

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

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,

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

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

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

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

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

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

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

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

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

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

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

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

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

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 $h = 1$. □

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

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

[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. $\Box$.

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?

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

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

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

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

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

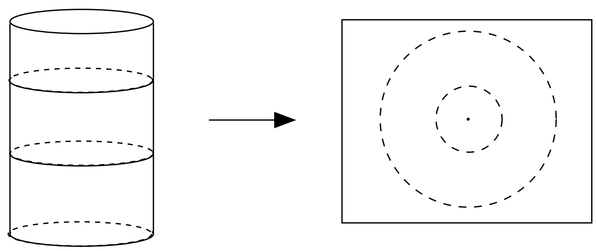


Figure 22:

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

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

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

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

  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

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

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

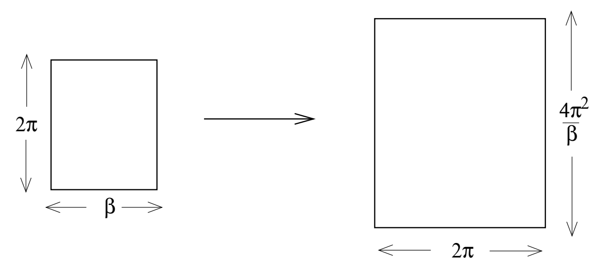


Figure 23:

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

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

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

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

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

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.

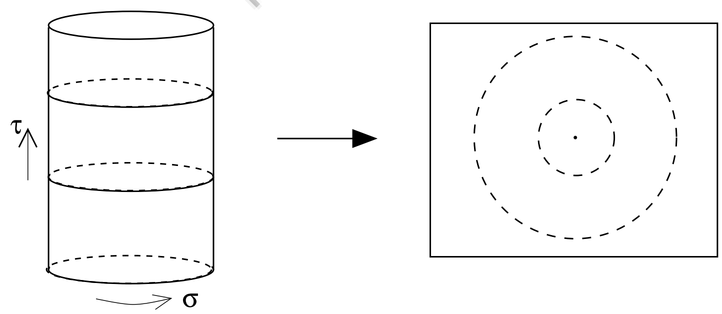


Figure 25: The map from the cylinder to the plane.

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,

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

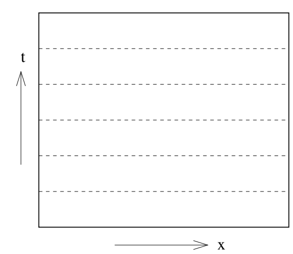


Figure 24:

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

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

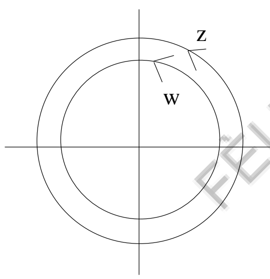
4.5.2 The Virasoro Algebra

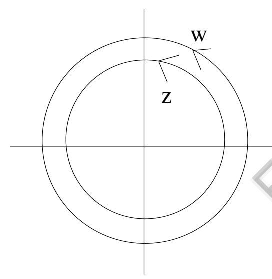
If we have some number of conserved charges, the first thing that we should do is compute their algebra. Representations of this algebra then classify the states of the theory. (For example, think angular momentum in the hydrogen atom). For conformal symmetry, we want to determine the algebra obeyed 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

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

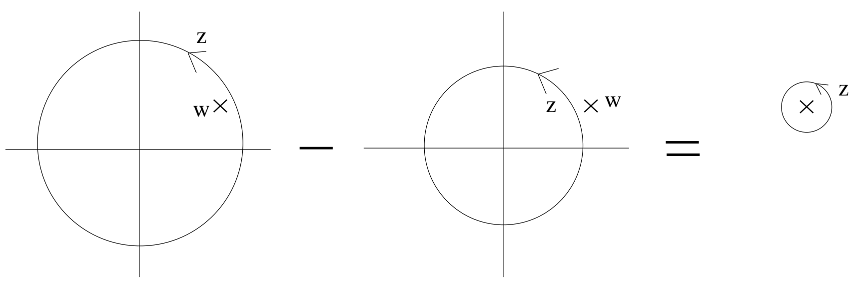
What does this actually mean?! We need to remember that all operator equations are to be viewed as living inside time-ordered correlation functions. Except, now we're working on the z-plane, this statement has transmuted into radially ordered correlation functions: outies to the left, innies to the right.

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

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

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

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

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

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

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

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 $|z| = r_i$.

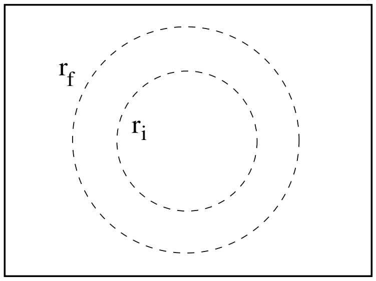


Figure 26:

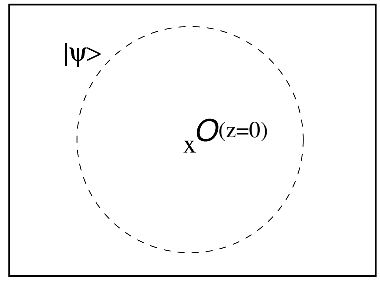


Figure 27:

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.

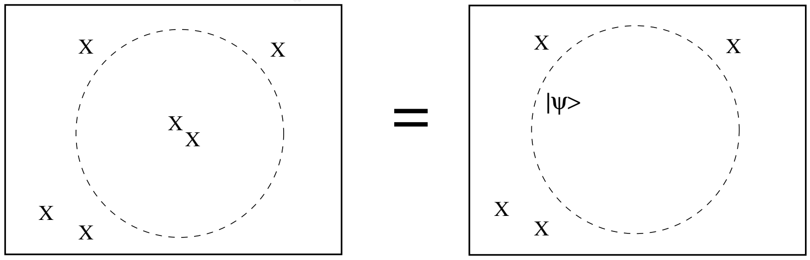


Figure 28:

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,

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

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. $\square$

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.

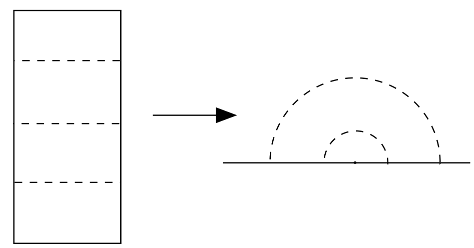


Figure 29:

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: $\mathrm{Im}z \geq 0$. The end points of the string are mapped to the real axis, $\mathrm{Im}z = 0$.

Much of our previous discussion goes through as before. But now we need to take care of boundary conditions at $\mathrm{Im}z = 0$. 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

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

In complex coordinates, this becomes

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

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 $\mathrm{Im} z = 0$. (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,

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

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.

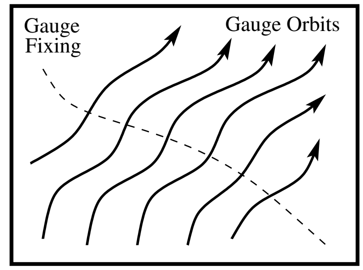


Figure 30:

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,

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

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.

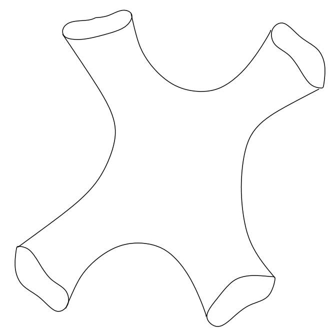


Figure 31:

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,

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

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).

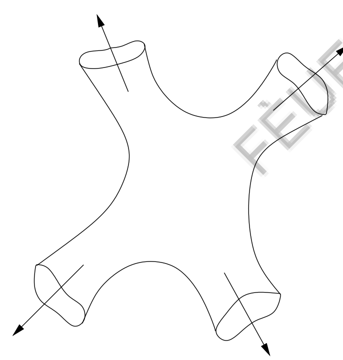


Figure 32:

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.

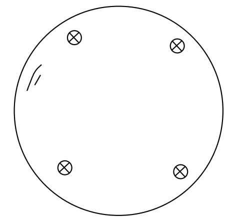
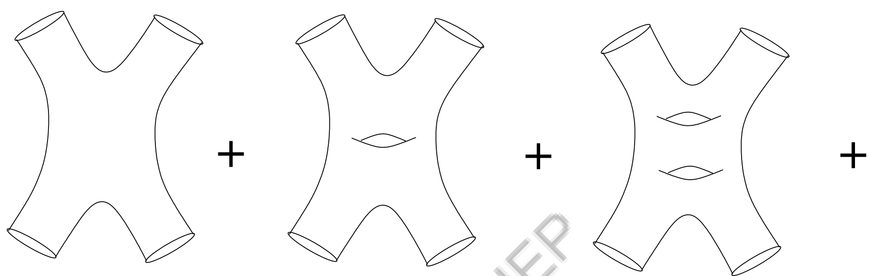


Figure 33:

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.

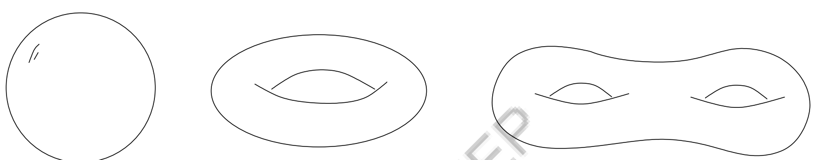


Figure 34: Examples of increasingly poorly drawn Riemann surfaces with $\chi = 2, 0$ and -2.

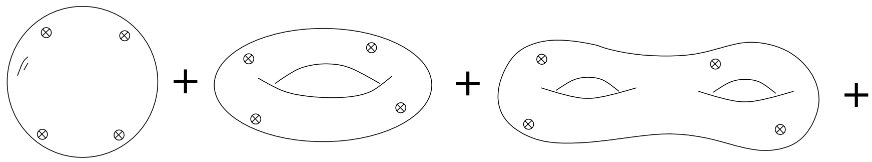
Now we see that the number λ — or, more precisely, e^λ — plays the role of the string coupling. The integral over worldsheets is weighted by,

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

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

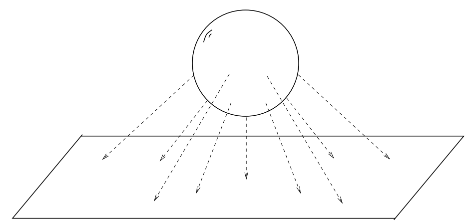


Figure 35:

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

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

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

where the tachyon vertex operator is given by,

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

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,

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

where the expectation value $\langle \ldots \rangle$ 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,

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

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 $d^2 z = 2 d^2 \sigma$.

The Gaussian Integral

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

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

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

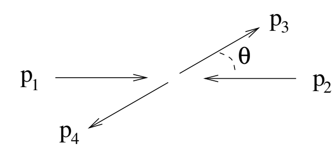


Figure 36:

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

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

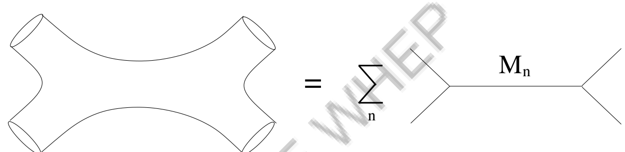
$$-1-\frac{\alpha' s}{4}=0 \qquad \Rightarrow \qquad s=-\frac{4}{\alpha'}$$

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,

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

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,



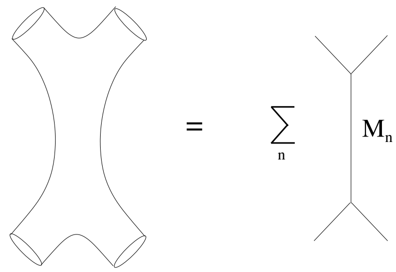
where the exchanged particles are all the different states of the free string.

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 (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,

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

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.

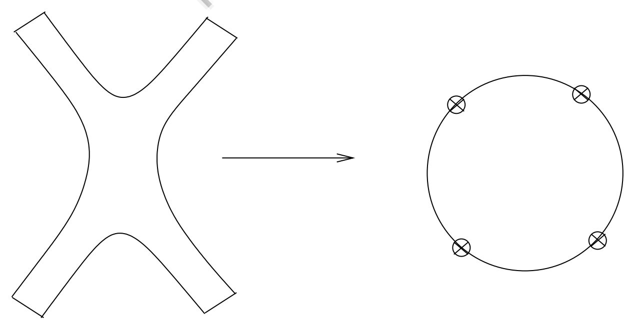


Figure 37: The conformal map from the open string worldsheet to the disc.

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

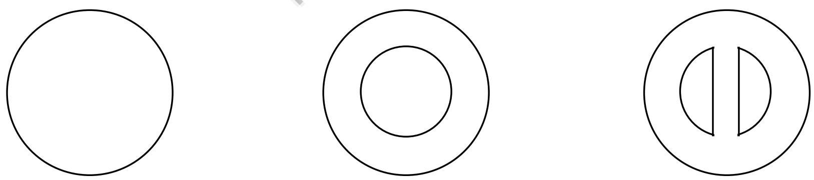


Figure 38: Riemann surfaces with boundary with $\chi = 1, 0$ and -1.

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 $\mathrm{Im}z = 0$. 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 $\mathrm{Im}z = 0$ 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,

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

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.

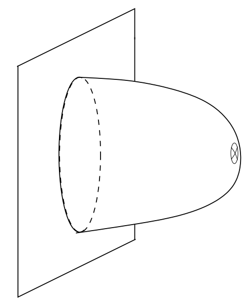


Figure 39:

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.

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* $\mathcal{M}$.

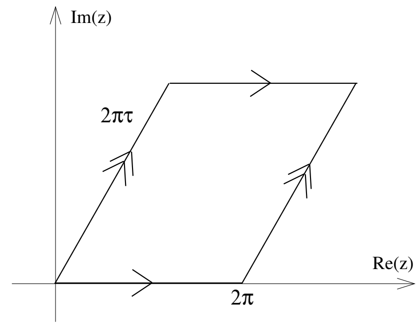


Figure 40:

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

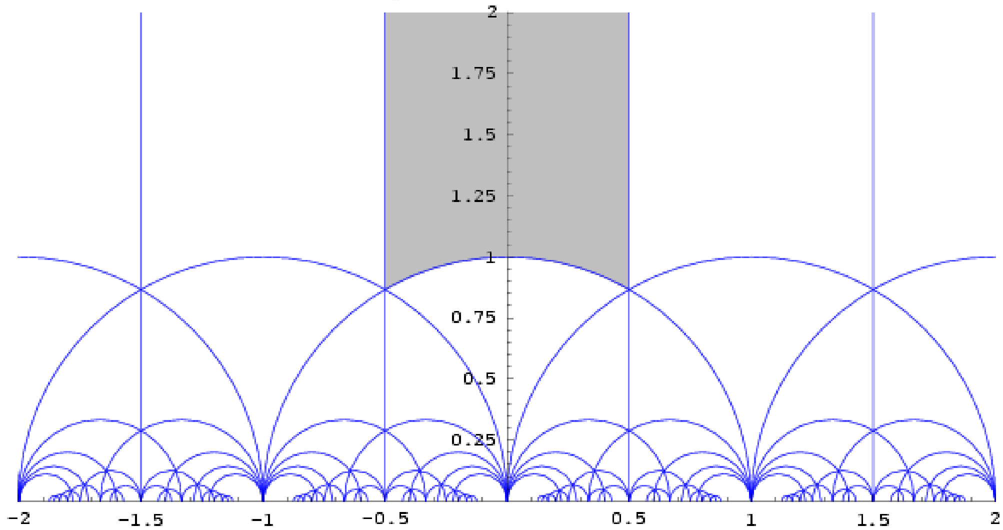


Figure 41: The fundamental domain.

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

$$d^2\tau \;\to\; \frac{d^2\tau}{|c\tau+d|^4} \quad \text{and} \quad \mathrm{Im}\,\tau \;\to\; \frac{\mathrm{Im}\,\tau}{|c\tau+d|^2}$$

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

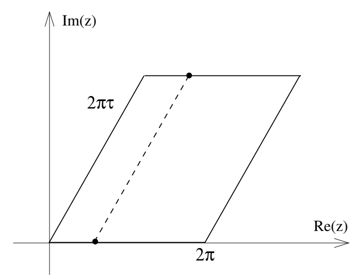


Figure 42:

$$P = L_0 - \tilde{L}_0$$

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,

$$Z_{\mathrm{scalar}}[\tau] \sim \frac{1}{\sqrt{\alpha' \mathrm{Im}\,\tau}} \frac{1}{(q\bar{q})^{1/24}} \prod_{n=1}^{\infty} \frac{1}{1-q^n} \;\; \prod_{n=1}^{\infty} \frac{1}{1-\bar{q}^n} \tag{6.20}$$

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

$$Z_{\mathrm{string}} = \int d^2\tau \; \frac{1}{(\mathrm{Im}\,\tau)} \frac{1}{(\alpha' \mathrm{Im}\,\tau)^{13}} \;\; \frac{1}{q\bar{q}} \left(\prod_{n=1}^{\infty} \frac{1}{1-q^n} \right)^{24} \left(\prod_{n=1}^{\infty} \frac{1}{1-\bar{q}^n} \right)^{24} \tag{6.21}$$

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

$$Z_{\text{string}} = \int \frac{d^2\tau}{(\text{Im}\,\tau)^2} \left(\frac{1}{\sqrt{\text{Im}\,\tau}} \frac{1}{\eta(q)} \frac{1}{\bar{\eta}(\bar{q})} \right)^{24}$$

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

$$Z_1 = \int_0^\infty dl\;\frac{1}{l^{1+D/2}}\;e^{-m^2l} \tag{6.23}$$

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

$$\Gamma(z) = \int_0^\infty dt\ t^{z-1} e^{-t} \qquad (6.26)$$

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.

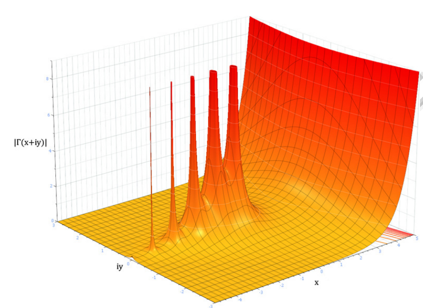


Figure 43:

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 $S_{\rm Poly}$ 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.

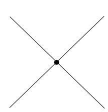
The analysis is greatly simplified by a cunning choice of coordinates. Around any point $\bar{x}$, we can always pick Riemann normal coordinates such that the expansion in $X^\mu = \bar{x}^\mu + \sqrt{\alpha'}\, Y^\mu$ gives

$$G_{\mu\nu}(X) = \delta_{\mu\nu} - \frac{\alpha'}{3}\mathcal{R}_{\mu\lambda\nu\kappa}(\bar{x})Y^\lambda Y^\kappa + \mathcal{O}(Y^3)$$

To quartic order in the fluctuations, the action becomes

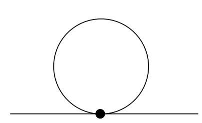
$$S = \frac{1}{4\pi}\int d^2\sigma\; \partial Y^\mu \partial Y^\nu \delta_{\mu\nu} - \frac{\alpha'}{3}\mathcal{R}_{\mu\lambda\nu\kappa}\, Y^\lambda Y^\kappa \partial Y^\mu \partial Y^\nu$$

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,

$$S = \frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{g} \, \left(G_{\mu\nu}(X)\, \partial_\alpha X^\mu \, \partial_\beta X^\nu g^{\alpha\beta} + i B_{\mu\nu}(X)\, \partial_\alpha X^\mu \, \partial_\beta X^\nu \, \epsilon^{\alpha\beta} \right) \tag{7.8}$$

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

$$S = \frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{g} \;\; \big(G_{\mu\nu}(X)\, \partial_\alpha X^\mu \, \partial_\beta X^\nu g^{\alpha\beta} + i B_{\mu\nu}(X)\, \partial_\alpha X^\mu \, \partial_\beta X^\nu \, \epsilon^{\alpha\beta} \\ + \alpha' \, \Phi(X)\, R^{(2)} \big) \tag{7.12}$$

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,

$$\Phi_0 = \operatorname*{limit}_{X\to\infty} \Phi(X) \tag{7.13}$$

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