Lecture Notes on
String Theory

Kevin Zhou
kzhou7@gmail.com

I should preface these notes by stating that I don't actually know any string theory at all. These "notes on string theory" are really just a collection of the prerequisites one needs to know before *starting* to learn string theory, beyond the usual ones of quantum field theory and general relativity. Topics covered include the quantization of constrained systems, 2D conformal field theory, and the quantization of the bosonic string. Nothing in these notes is original; they have been compiled from a variety of sources. The primary sources were:

- David Tong's String Theory lecture notes. These clear lecture notes follow the first volume of Polchinski, but with a much more approachable style, and hence are accessible after a good course on quantum field theory.
- Timo Weigand's String Theory lecture notes. As with the quantum field theory notes, these notes are slightly more comprehensive and precise, at the cost of being slightly drier; also contains a short introduction to the superstring.
- Zwiebach, *A First Course in String Theory*. A very basic introduction, covering bosonic string theory in the first part, and a wide variety of developments, such as D-branes and AdS/CFT in the second. The book is clearly written and accessible even without any field theory background, and in fact might be useful as indirect preparation for field theory. It has the benefit of explicitly showing many steps of logic skipped in most books. The downside is that the entire 300 pages of the first part barely covers the first 30 pages of Polchinski.
- Henneaux and Teitelboim, *Quantization of Gauge Systems*. The most thorough book by far on the subject; if you're wondering how constrained quantization, Grassmann variables, or BRST symmetry really work, this is the place to go. Naturally, more formal than any of the other books on this list. The most relevant chapters for these notes are 1, 4, 6, and 13.

The most recent version is here; please report any errors found to kzhou7@gmail.com.

Contents

1 Preliminaries

1.1 Miscellanea

As a first pass, we'll follow Zwiebach's text. In this first section, we establish some conventions.

- There are three fundamental dimensions: mass, length, and time. Charge is not a independent dimension; this is clearest in the cgs system, where the Coulomb force law is $F = q_1 q_2 / r^2$. That is, one can define charges purely in terms of the forces they produce. In the SI system, all occurrences of the unit of charge in measurable quantities are canceled out by ϵ_0 or μ_0.

- String theory is said to have no adjustable parameters. This means that it has no dimensionless parameters; there is a single dimensionful parameter, the string length ℓ_s. When string theory was considered as a theory of hadrons, ℓ_s was thought to be on the nuclear scale, but now it is viewed as much smaller.

- We use the $(-+++)$ metric convention, but define the interval ds^2 to measure proper time,

$$-ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu.$$

- We will perform the quantization of the relativistic string in light cone coordinates, where

$$x^\pm = \frac{1}{\sqrt{2}}(x^0 \pm x^1).$$

  The other coordinates stay the same, so in x^μ, the μ index runs over $(+, -, 2, 3)$. The light cone coordinate axes are tilted so that they lie on the light cone, e.g. a photon moving to the right has $x^- = 0$.

- The metric in light cone coordinates is

$$-ds^2 = -2dx^+ dx^- + (dx^2)^2 + (dx^3)^2, \quad \eta_{\mu\nu} = \begin{pmatrix} & -1 & & \\ -1 & & & \\ & & 1 & \\ & & & 1 \end{pmatrix}.$$

  For example, we have

$$a_+ = -a^-, \quad a_- = -a^+, \quad a^\mu b_\mu = -a^- b^+ - a^+ b^- + a^2 b^2 + a^3 b^3.$$

- None of the coordinates are true time coordinates, but we will conventionally think of x^+ as the light cone time coordinate. This isn't completely unnatural, since $dx^\pm / d\tau > 0$ for almost all particles.

- We think of dx^- / dx^+ as a "light cone velocity". For a particle with speed v, the light cone velocity is $(1 - v)/(1 + v)$. In particular, it is zero for a massless particle moving to the right, infinite for a massless particle moving to the left, and one when $v = 0$.

- Next we must define the light cone energy. Energy is the conjugate variable to time, and

$$p_\mu x^\mu = p_+ x^+ + p_- x^- + p_2 x^2 + p_3 x^3$$

  which motivates us to define the light cone energy as $-p_+ = p^-$. Defining energy as a conjugate variable is useful as, e.g. the Schrodinger equation $i\,\partial\psi/\partial x^0 = p^0 \psi$ becomes $i\,\partial\psi/\partial x^+ = p^- \psi$ in lightcone coordinates.

Next, we consider the possibility of extra spatial dimensions. We maintain only one time dimension, since it is difficult to construct a consistent theory with more than one.

- The extra dimensions can be topologically nontrivial; for example, for a single dimension we may identify $x \sim x + 2\pi R$. The interval $0 \leq x < 2\pi R$ is a fundamental domain for this identification, as every point is identified with exactly one point in the fundamental domain. We then construct the space by identifying points on the boundary, getting a circle.
- Sometimes identifications have fixed points. The resulting space is not a manifold, since it is singular at the fixed point, but an orbifold. For example, identifying $x \sim -x$ gives the half-line $x \geq 0$, called the $\mathbb{R}^1/\mathbb{Z}_2$ orbifold.
- As another example, consider the identification $z \sim e^{2\pi i/N} z$ in the complex plane, giving the $\mathbb{C}/\mathbb{Z}_N$ orbifold. It is a cone which is singular at its vertex; one fundamental domain is $0 \leq \theta < 2\pi/N$.
- Physics on spaces with generic singularities is typically complicated and possibly even inconsistent. Orbifolds are interesting because they have "tractable" singularities, so we may quantize strings on them.
- For point particles, a small compactified dimension creates new energy levels, but they are very high if the dimension is small. By contrast, for a string, new low-lying states can appear if the dimension is much *smaller* than the string length, corresponding to the string wrapping around it. This is a consequence of T duality, as we'll see.
- Compactified dimensions lead to subtleties with gauge theory. It can be the case that two configurations with the same fields are not related by a gauge transformation; then we must consider the states physically distinct. It also also be the case that some configuration of fields cannot be realized by a potential, even if we allow the potential be defined on patches and related between patches with gauge transformations; such states are forbidden.

Note. A quick review of the nonrelativistic string. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}\mu \left(\frac{\partial y}{\partial t}\right)^2 - \frac{1}{2}T \left(\frac{\partial y}{\partial x}\right)^2.$$

When we attempt to extremize the action, we run into a boundary term

$$\delta S \supset -T \int dt \, \left(\frac{\partial y}{\partial x}\delta y\right)\Bigg|_{x=0}^{x=a}.$$

To remove this unwanted boundary term, we need to apply boundary conditions. One acceptable boundary condition is Neumann boundary conditions (free ends), where

$$\left.\frac{\partial y}{\partial x}\right|_{x=0,a} = 0.$$

Alternatively, we could use Dirichlet boundary conditions (fixed ends),

$$\left.\frac{\partial y}{\partial t}\right|_{x=0,a} = 0$$

which ensures that y is constant on the boundaries, and hence $\delta y = 0$ there. A third alternative is that the string is closed, so that there are no boundaries at all.

In the initial days of string theory, open strings were given Neumann boundary conditions, but it was later realized that they could have Dirichlet boundary conditions if they attached to an extended object called a Dp-brane, where D stands for Dirichlet and p is the number of spatial dimensions. (One can think of Neumann boundary conditions as the special case where the D-brane fills all space.) Remarkably, it turns out that D-branes are physical objects in their own right, and arise naturally from string theory without being introduced by hand.

Finally, it will be useful later to consider the conjugate momenta,

$$\mathcal{P}^t = \frac{\partial \mathcal{L}}{\partial \dot{y}} = \mu \frac{\partial y}{\partial t}, \quad \mathcal{P}^x = \frac{\partial \mathcal{L}}{\partial y'} = -T \frac{\partial y}{\partial x}.$$

Here $\mathcal{P}^t$ is simply the usual momentum density in the y-direction. The equation of motion is

$$\frac{\partial \mathcal{P}^t}{\partial t} + \frac{\partial \mathcal{P}^x}{\partial x} = 0.$$

In this notation, Neumann and Dirichlet boundary conditions set $\mathcal{P}^x = 0$ and $\mathcal{P}^t = 0$ respectively.

Note. Gravity and electromagnetism in higher dimensions. We will use

$$\mathrm{vol}(S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma(d/2)}, \quad \mathrm{vol}(B^d) = \frac{\pi^{d/2}}{\Gamma(1+d/2)}.$$

In d spatial dimensions, Gauss's law continues to hold, so the field of a point charge is

$$E(r) = \frac{1}{\mathrm{vol}(S^{d-1})} \frac{q}{r^{d-1}}$$

in cgs units. As for gravitation, we have

$$g(r) = \frac{Gm}{r^2}$$

in $D = 4$ spacetime dimensions. More generally, the gravitational field is determined by

$$\nabla^2 V^{(D)} = 4\pi G^{(D)} \rho, \quad \mathbf{g} = -\nabla V$$

where $G^{(D)}$ is the gravitational constant in D dimensions. The Planck length in D dimensions is related by dimensional analysis,

$$\frac{\hbar}{c^3} = \frac{\ell_P^2}{G} = \frac{(\ell_P^{(D)})^{D-2}}{G^{(D)}}.$$

In order to relate gravitational constants in different dimensions, we need to relate densities in different dimensions. Consider a single small compactified dimension of length ℓ_C. Then for distances much greater than ℓ_C we have the coarse-grained mass density

$$\rho^{(4)} = 2\pi R \rho^{(5)}.$$

Plugging this into Poisson's equation, we conclude

$$\frac{G^{(5)}}{G} = \ell_C$$

and more generally for multiple compactified dimensions with volume V_C,

$$\frac{G^{(D)}}{G} = V_C.$$

For a D-dimensional theory where each compactified dimension has length ℓ_C,

$$(\ell_P^{(D)})^{D-2} = (\ell_P)^2 \frac{G^{(D)}}{G} = (\ell_P)^2 (\ell_C)^{D-4}, \quad \ell_C = \ell_P^{(D)} \left(\frac{\ell_P^{(D)}}{\ell_P} \right)^{\frac{2}{D-4}}.$$

Intuitively, in the presence of compactified dimensions, gravity appears much weaker than it actually is, leading to an underestimate of the fundamental Planck length $\ell_P^{(D)}$. In particular, for two compactified dimensions ($D = 6$) we could have a TeV-scale Planck energy if $\ell_C \sim 10^{-5}$ m.

In the context of string theory, the macroscopic dimensions form a D3-brane; open strings must terminate on this brane while closed strings do not. Electromagnetism and other forces are associated with open strings and hence are not affected by the extra dimensions, while gravity is associated with closed strings. Then the hypothesis of large extra dimensions could only be tested by measuring gravitational effects at short distances, which is extremely difficult. It predicts the appearance of many new, stringy excitations at a scale much lower than the conventional Planck scale.

Note. The study of higher dimensions began with Kaluza–Klein theory, which is essentially general relativity in five dimensions. Upon compactifying the fifth dimension, and restricting to four dimensions, the metric tensor g_{MN} (with $M, N = 0, \ldots, 4$) decomposes into the ordinary metric $g_{\mu\nu}$, a massless vector field $g_{\mu 4}$ interpreted as the electromagnetic field, and a massless scalar field g_{44}. This scalar field was originally thought to be problematic, but was later incorporated into Brans–Dicke theory, an extension of general relativity. In string theory, the dilaton plays a similar role.

1.2 Classical Strings

We now arrive at the classical relativistic string.

- A relativistic string traces out a two-dimensional surface in spacetime, called a worldsheet. The Nambu–Goto action is the area of the worldsheet.
- As a warmup, consider a surface in space, parametrized as $\mathbf{x}(\xi^1, \xi^2)$. Since the area spanned by two vectors is the determinant of the 2×2 matrix that contains them,

  $$A = \int d\xi^1 d\xi^2 \sqrt{(\partial_1 \mathbf{x} \cdot \partial_1 \mathbf{x})(\partial_2 \mathbf{x} \cdot \partial_2 \mathbf{x}) - (\partial_1 \mathbf{x} \cdot \partial_2 \mathbf{x})^2}$$

  where ∂_i is a derivative with respect to ξ^i.
- It is simpler to express this in terms of the induced metric on the worldsheet,

  $$ds^2 = g_{ij}\, d\xi^i d\xi^j, \quad g_{ij} = \partial_i \mathbf{x} \cdot \partial_j \mathbf{x}.$$

  This is the metric in the (ξ^1, ξ^2) coordinates, and the area is

  $$A = \int d\xi^1 d\xi^2 \sqrt{g}$$

  which is manifestly reparametrization invariant.

- Now, for a spacetime surface, we parametrize the worldsheet as $X^\mu(\tau, \sigma)$. Note that conventionally one calls both the domain and image of this map the worldsheet, though we will stick to the latter usage. The letter X is capitalized to avoid confusing the string coordinates X^μ with spacetime coordinates x^μ.
- The parameter τ is roughly related to time on the string, while σ is roughly related to position along the string. As such, we define

$$\dot{X} = \frac{\partial X}{\partial \tau}, \quad X' = \frac{\partial X}{\partial \sigma}$$

and we require $\dot{X}^0 \neq 0$. We will usually choose X' to be spacelike and $\dot{X}$ to be timelike or null. It is tempting to identify $\dot{X}$ with the velocity of a piece of the string, but this is inappropriate as none of the points on the string are distinguishable.

- By analogy with the spatial case, the area is

$$A = \int d\tau d\sigma \sqrt{(\dot{X} \cdot X')^2 - (\dot{X})^2 (X')^2}.$$

This is an area in Lorentzian signature, which is somewhat less intuitive. There is a minus sign relative to the spatial case so the argument of the square root is positive, and the area is zero if the string is moving transverse to itself at the speed of light.

- The Nambu–Goto action is

$$S = -T_0 \int_{\tau_i}^{\tau_f} d\tau \int_0^{\sigma_1} d\sigma \sqrt{(\dot{X} \cdot X')^2 - (\dot{X})^2 (X')^2}$$

where T_0 is called the string tension. The easiest way to check reparametrization invariance is again to introduce a metric on the worldsheet,

$$-ds^2 = \eta_{\mu\nu} dX^\mu dx^\nu = \gamma_{\alpha\beta} d\xi^\alpha d\xi^\beta, \quad \gamma_{\alpha\beta} = \frac{\partial X}{\partial \xi^\alpha} \cdot \frac{\partial X}{\partial \xi^\beta} = \begin{pmatrix} \dot{X}^2 & \dot{X} \cdot X' \\ \dot{X} \cdot X' & (X')^2 \end{pmatrix}.$$

Then the Nambu–Goto action can be written in the manifestly reparametrization invariant form

$$S = -T_0 \int d\tau d\sigma \sqrt{-\gamma}.$$

Now we turn to the equations of motion.

- Writing the Nambu–Goto action as the integral of a Lagrangian density as usual, we define the canonical momenta

$$P^\tau_\mu = \frac{\partial \mathcal{L}}{\partial \dot{X}^\mu} = -T_0 \frac{(\dot{X} \cdot X') X'_\mu - (X')^2 \dot{X}_\mu}{\sqrt{(\dot{X} \cdot X')^2 - (\dot{X})^2 (X')^2}}, \quad P^\sigma_\mu = \frac{\partial \mathcal{L}}{\partial X'^\mu} = -T_0 \frac{(\dot{X} \cdot X') \dot{X}_\mu - (\dot{X})^2 X'_\mu}{\sqrt{(\dot{X} \cdot X')^2 - (\dot{X})^2 (X')^2}}.$$

Throwing away boundary terms, the Euler-Lagrange equation is

$$\frac{\partial P^\tau_\mu}{\partial \tau} + \frac{\partial P^\sigma_\mu}{\partial \sigma} = 0$$

which is quite complicated in terms of $\dot{X}$ and X'.

- As for the nonrelativistic string, the derivation only goes through if we can ignore boundary terms. The τ boundary terms vanish by fixing the initial condition $X^\mu(\tau_i, \sigma)$ and final condition $X^\mu(\tau_f, \sigma)$, leaving the σ boundary term,

$$\delta S \supset \int d\tau \, (\delta X^\mu \, P^\sigma_\mu) \Big|_0^{\sigma_1}.$$

- In the case of a closed string, these conditions are automatically satisfied by the periodicity of X^μ. For an open string, we may remove these terms by imposing Dirichlet boundary conditions,

$$\frac{\partial X^\mu}{\partial \tau} \bigg|_{\sigma = \sigma_*}$$

  or the free endpoint condition,

$$P^\sigma_\mu \Big|_{\sigma=\sigma_*} = 0.$$

  The Dirichlet boundary condition cannot be used for $\mu = 0$, so we must use the free endpoint condition there. The free endpoint condition is analogous to a Neumann boundary condition. We will show later that the Dirichlet boundary condition implies $P^\tau_\mu = 0$ at the string endpoints.

- In order to simplify the equations, we work in the static gauge, where

$$\tau = t$$

  where t is the time coordinate. This is called the static gauge because constant τ slices match constant t slices. We normalize σ by

$$\sigma \in [0, \sigma_1] \text{ for open string}, \quad \sigma \in [0, \sigma_c] \text{ for closed string}.$$

  In the static gauge we simply have

$$X' = \left(0, \frac{\partial \mathbf{X}}{\partial \sigma}\right), \quad \dot{X} = \left(1, \frac{\partial \mathbf{X}}{\partial t}\right).$$

Example. A stretched string. Suppose a string is static along the x^1 axis, with endpoints at $x^1 = 0$ and $x^1 = a$. Then $X^1(f, \sigma) = f(\sigma)$, and we have

$$(X')^1 = f'(\sigma), \quad (\dot{X})^0 = 1$$

with all other elements zero. Note that $f(\sigma)$ must assign every point on the string a unique parameter value, so $f'(\sigma) > 0$. The action is simply

$$S = -T_0 \int dt \int d\sigma \sqrt{f'^2} = -T_0 a \int_{t_i}^{t_f} dt.$$

Since generically $L = T - V$ and the kinetic energy vanishes here, this may be interpreted as a potential $T_0 a$, justifying the interpretation of T_0 as a tension. The equation of motion is

$$\frac{\partial P^\sigma_\mu}{\partial \sigma} = 0, \quad P^\sigma_\mu = -T_0 \frac{X'_\mu}{f'}$$

Then $P^\sigma_\mu = -T_0 \delta_{\mu 1}$ and hence the equation of motion is satisfied.

The longitudinal velocity of a string is not meaningful, but we can get further insight by writing the action in terms of transverse velocity.

- Let s measure length along the string, so that $\partial\mathbf{X}/\partial s$ is a spatial unit vector. We then project out the longitudinal part of the string velocity $\partial\mathbf{X}/\partial t$, defining

$$\mathbf{v}_\perp = \frac{\partial\mathbf{X}}{\partial t} - \left(\frac{\partial\mathbf{X}}{\partial t}\cdot\frac{\partial\mathbf{X}}{\partial s}\right)\frac{\partial\mathbf{X}}{\partial s}, \quad v_\perp^2 = \left(\frac{\partial\mathbf{X}}{\partial t}\right)^2 - \left(\frac{\partial\mathbf{X}}{\partial t}\cdot\frac{\partial\mathbf{X}}{\partial s}\right)^2.$$

  The Nambu–Goto action can then simply be written as

$$S = -T_0\int dt\int ds\,\sqrt{1-v_\perp^2}.$$

- This is closely analogous to the relativistic point particle action. To make the analogy even clearer, we imagine a string as starting with an infinitesimal length. It takes work $T_0 a$ to stretch it to a length a, which gives it a mass $T_0 a$. Hence T_0 is the mass per length of the string.
- A somewhat nasty calculation shows that

$$P^{\sigma\mu} = -\frac{T_0}{\sqrt{1-v_\perp^2}}\left(\left(\frac{\partial\mathbf{X}}{\partial s}\cdot\frac{\partial\mathbf{X}}{\partial t}\right)\dot{X}^\mu + \left(1-\left(\frac{\partial\mathbf{X}}{\partial t}\right)^2\right)\frac{\partial X^\mu}{\partial s}\right)$$

  as well as

$$P^{\sigma 0} = -T_0\frac{\left(\frac{\partial\mathbf{X}}{\partial s}\cdot\frac{\partial\mathbf{X}}{\partial t}\right)}{\sqrt{1-v_\perp^2}}, \quad P^{\tau 0} = T_0\frac{ds}{d\sigma}\frac{1}{\sqrt{1-v_\perp^2}}, \quad P^{\tau i} = T_0\frac{ds}{d\sigma}\frac{v_\perp^i}{\sqrt{1-v_\perp^2}}.$$

- Now consider the motion of the endpoints of an open string; these are distinguished points and hence have an unambiguous velocity. At the endpoints, P^σ_μ must vanish, and hence for $P^{\sigma 0}$ to vanish we must have $(\partial\mathbf{X}/\partial s)\cdot(\partial\mathbf{X}/\partial t) = 0$. Then the string endpoints move transverse to the string. Plugging this back into the general expression for $P^{\sigma\mu}$, we see that $(\partial\mathbf{X}/\partial t)^2 = 1$, so the endpoints move at the speed of light.
- Finally, by writing the Lagrangian as a function of $\partial_s\mathbf{X}$ and $\partial_t\mathbf{X}$ and performing a Legendre transformation with respect to the canonical momentum $\mathbf{P} = \partial\mathcal{L}/\partial(\partial_t\mathbf{X})$, we find

$$H = \int\frac{T_0\,ds}{\sqrt{1-v_\perp^2}}$$

  which is precisely the kinetic energy of the string, with rest mass T_0 per unit length. Similarly, integrating $P^{\tau i}$ gives the momentum. Note that the speed of waves on a string is $v = \sqrt{T_0/\rho}$ where ρ is the mass density, so excitations on the string travel at the speed of light.

Now we choose a useful parametrization for σ.

- One could replace the parameter σ with s, but it is more convenient to use a parameter with a fixed range. Note that in static gauge, the spatial configuration of a string is described by $\mathbf{X}(\sigma, t)$. This yields a two-dimensional spatial surface, which we'll call the string surface; this should not be confused with the worldsheet.

- We know that $\partial \mathbf{X}/\partial t$ is a unit vector on the string surface, so we choose σ so that

$$\frac{\partial \mathbf{X}}{\partial \sigma} \cdot \frac{\partial \mathbf{X}}{\partial t} = 0, \quad \mathbf{v}_\perp = \frac{\partial \mathbf{X}}{\partial t}.$$

  Since s and σ are related to a scaling, we also have $(\partial \mathbf{X}/\partial s) \cdot (\partial \mathbf{X}/\partial t) = 0$.

- As a result, our expressions for the momenta simplify,

$$P^{\tau\mu} = T_0 \frac{ds}{d\sigma} \frac{\partial X^\mu/\partial t}{\sqrt{1 - v_\perp^2}}, \quad P^{\sigma\mu} = -T_0 \sqrt{1 - v_\perp^2} \frac{\partial X^\mu}{\partial s}.$$

  In other words, the simplifications that we previously found held for the string endpoints hold everywhere on the string for this parametrization.

- Now we consider the equations of motion again. The $\mu = 0$ component of the equation of motion is simply

$$\frac{\partial P^{\tau 0}}{\partial t} = 0.$$

  However, $P^{\tau 0}$ is simply the energy density of the string, per unit σ. The conservation of this quantity makes sense because, unlike strings that actually exist, the string motion here is purely transverse, so the tension does zero work on any segment with endpoints of fixed σ. If the transverse speed of such a segment changes, the energy comes from the individual stretching or shrinking of that segment.

- The spatial components of the equation of motion read

$$\frac{\partial}{\partial \sigma} \left(T_0 \sqrt{1 - v_\perp^2} \frac{\partial \mathbf{X}}{\partial s} \right) = T_0 \frac{ds/d\sigma}{\sqrt{1 - v_\perp^2}} \frac{\partial \mathbf{v}_\perp}{\partial t}$$

  which rearranges to

$$\frac{T_0}{\sqrt{1 - v_\perp^2}} \frac{\partial \mathbf{v}_\perp}{\partial t} = \frac{\partial}{\partial s} \left(T_0 \sqrt{1 - v_\perp^2} \frac{\partial \mathbf{X}}{\partial s} \right).$$

  This is precisely the usual wave equation, if we interpret T_0 as the tension in the rest frame of a piece of string. Then the factor of $\sqrt{1 - v_\perp^2}$ on the right simply reflects the Lorentz transformation of force, while the factor of γ on the left reflects the fact that moving objects are harder to turn by a factor of γ.

- This also physically explains why the endpoints of an open string must move at the speed of light. The tension on a real open string goes to zero at endpoints because the endpoints have no mass and hence can experience no force. But the proper tension in these abstract strings is constant, so we only avoid a divergence if it Lorentz transforms to zero at the ends.

- We have still not used up all the freedom in σ parametrization. We already know that the energy of a piece of string whose endpoints have fixed σ is constant. Hence it is consistent to scale σ so the energy density is constant,

$$d\sigma = \frac{ds}{\sqrt{1 - v_\perp^2}} = \frac{dE}{T_0}.$$

Then we have $\sigma \in [0, \sigma_1]$ where $\sigma_1 = E/T_0$. This is equivalent to

$$\left(\frac{\partial \mathbf{X}}{\partial \sigma}\right)^2 + \left(\frac{\partial \mathbf{X}}{\partial t}\right)^2 = 1.$$

- In this case, the equation of motion reduces to

$$\frac{\partial^2 \mathbf{X}}{\partial t^2} = \frac{\partial^2 \mathbf{X}}{\partial \sigma^2}.$$

  The canonical momenta and boundary conditions are now

$$P^{\tau\mu} = T_0 \frac{\partial X^\mu}{\partial t}, \quad P^{\sigma\mu} = -T_0 \frac{\partial X^\mu}{\partial \sigma}, \quad \frac{\partial \mathbf{X}}{\partial \sigma} = 0 \text{ at endpoints.}$$

  Note that we no longer have $1/\sqrt{1 - v_\perp^2}$ Lorentz factors, because we're now parametrizing by energy rather than by length.

Note. As shown in the lecture notes on General Relativity, energy density and pressure contribute equally to gravitational attraction. A string has a negative pressure due to its tension, which is equal to its mass density, so a long static string would exert no gravitational force. Cosmic strings would instead be detected by the deficit angles they produce, which would yield multiple images of distant objects.

1.3 String Motion

With the wave equation, we may now solve for the general motion of open and closed strings.

- A solution to the wave equation can always be written as a superposition of "left moving" and "right moving" solutions,

$$\mathbf{X}(t, \sigma) = \frac{1}{2}(\mathbf{F}(t + \sigma) + \mathbf{G}(t - \sigma)).$$

  The boundary condition at the $\sigma = 0$ endpoint demands that $\mathbf{F}' = \mathbf{G}'$, so they differ by a constant. Hence the general solution can be written as

$$\mathbf{X}(t, \sigma) = \frac{1}{2}(\mathbf{F}(t + \sigma) + \mathbf{F}(t - \sigma)).$$

- Next, consider the boundary condition at $\sigma = \sigma_1$, which gives

$$\mathbf{F}'(t + \sigma_1) = \mathbf{F}'(t - \sigma_1).$$

  This implies that $\mathbf{F}$ is quasi-periodic,

$$\mathbf{F}(u + 2\sigma_1) = \mathbf{F}(u) + 2\sigma_1 \mathbf{v}_0.$$

  To interpret $\mathbf{F}$, note that $\mathbf{X}(t, 0) = \mathbf{F}(t)$, so the curve $\mathbf{F}(u)$ traces out the motion of the $\sigma = 0$ endpoint, and $\mathbf{v}_0$ is its average speed.

- By adding and subtracting the results
$$\left(\frac{\partial \mathbf{X}}{\partial \sigma}\right)^2 + \left(\frac{\partial \mathbf{X}}{\partial t}\right)^2 = 1, \quad \frac{\partial \mathbf{X}}{\partial \sigma} \cdot \frac{\partial \mathbf{X}}{\partial t} = 0$$
we have the equivalent set of two constraints
$$\left(\frac{\partial \mathbf{X}}{\partial \sigma} \pm \frac{\partial \mathbf{X}}{\partial t}\right)^2 = 1.$$
Plugging this into our general form shows that
$$\left|\frac{d\mathbf{F}}{du}\right|^2 = 1$$
so that u can be interpreted as a length parameter along the curve $\mathbf{F}(u)$. This constraint is equivalent to demanding the endpoint $\sigma = 0$ moves at the speed of light.

- We now consider the motion of a closed string. This is a bit more complicated because we can't use boundary conditions to eliminate $\mathbf{G}$. We start with the general solution
$$\mathbf{X}(t, \sigma) = \frac{1}{2}(\mathbf{F}(u) + \mathbf{G}(v)), \quad u = t + \sigma, \quad v = t - \sigma.$$
We may solve for the derivatives of $\mathbf{F}$ and $\mathbf{G}$ in terms of $\mathbf{X}$,
$$\frac{\partial \mathbf{X}}{\partial \sigma} + \frac{\partial \mathbf{X}}{\partial t} = \mathbf{F}'(u), \quad \frac{\partial \mathbf{X}}{\partial \sigma} - \frac{\partial \mathbf{X}}{\partial t} = -\mathbf{G}'(v).$$
As a result, the parametrization constraints give
$$\left|\frac{d\mathbf{F}}{du}\right|^2 = \left|\frac{d\mathbf{G}}{dv}\right|^2 = 1.$$

- Instead of a boundary condition, we have a periodicity condition $\sigma \sim \sigma + \sigma_1$, which means that $\mathbf{X}(t, \sigma + \sigma_1) = \mathbf{X}(t, \sigma)$. This is equivalent to
$$\mathbf{F}(u + \sigma_1) - \mathbf{F}(u) = \mathbf{G}(v) - \mathbf{G}(v - \sigma_1).$$
That is, both $\mathbf{F}$ and $\mathbf{G}$ are quasi-periodic with the same constant.

- One interesting generic phenomenon is the formation of cusps. Note that $\mathbf{F}'$ and $\mathbf{G}'$ are periodic functions on the unit sphere. Suppose that $\mathbf{F}'(u_0) = \mathbf{G}'(v_0)$, with corresponding coordinates (t_0, σ_0). Without loss of generality, we shift to set $t_0 = \sigma_0 = 0$ and hence $u_0 = v_0 = 0$.

- Now we Taylor expand the shape of the string about this point,
$$\mathbf{X}(0, \sigma) - \mathbf{X}(0, 0) = \sigma \frac{\partial \mathbf{X}}{\partial \sigma} + \frac{\sigma^2}{2} \frac{\partial^2 \mathbf{X}}{\partial \sigma^2} + \ldots.$$
We note that
$$\frac{\partial \mathbf{X}}{\partial \sigma} = \frac{1}{2}(\mathbf{F}' - \mathbf{G}') = \mathbf{0}.$$
Therefore, the leading term is quadratic, which means we have a cusp; the string enters in one direction and exits along the same direction. Generically cusps will appear and disappear periodically throughout the string. They are thought to be efficient sources of gravitational waves for cosmic strings.

Example. Consider a straight string of length ℓ rotating with angular velocity ω, with its midpoint at the origin of the xy plane. We can infer $\mathbf{F}(u)$ from the endpoint motion,

$$\mathbf{F}(u) = \frac{\ell}{2}(\cos\omega u, \sin\omega u).$$

Periodicity requires $\omega = \pi m/\sigma_1$ for an integer m, and

$$\mathbf{X}(0,\sigma) = \frac{\mathbf{F}(\sigma) + \mathbf{F}(-\sigma)}{2} = \frac{\ell}{2}(\cos(\pi m\sigma/\sigma_1), 0).$$

For $m > 1$, the string traces over itself multiple times, so we focus on the case $m = 1$. Since the $\sigma = 0$ endpoint moves at the speed of light,

$$\ell = \frac{2}{\omega} = \frac{2\sigma_1}{\pi} = \frac{2}{\pi}\frac{E}{T_0}$$

so the total energy is $(\pi/2)T_0\ell$. The tension provides $T_0\ell$, while the rest is due to the kinetic energy of the string. The complete solution as a function of time is

$$\mathbf{X}(t,\sigma) = \frac{\sigma_1}{\pi}\cos\frac{\pi\sigma}{\sigma_1}\left(\cos\frac{\pi t}{\sigma_1}, \sin\frac{\pi t}{\sigma_1}\right).$$

We now apply Noether's theorem to find conserved quantities.

- The results of Noether's theorem will look somewhat different. Usually, a classical field maps from spacetime, $\phi\colon \mathbb{R}^4 \to N$ where N is the field space. However, in string theory we deal with maps into spacetime, $X^\mu\colon M \to \mathbb{R}^4$ where M is the worldsheet. Hence the dynamics of a single string can be thought of as a two-dimensional field theory with a four-component field X.

- The action is
$$S = \int d\xi^i\, \mathcal{L}(\partial_i X^\mu), \quad (\xi^0, \xi^1) = (\tau, \sigma).$$
The fields X^μ have a continuous shift symmetry $\delta X^\mu = \epsilon^\mu$, which corresponds to translation in spacetime. Applying Noether's theorem, the conserved currents are
$$j^\alpha_\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\alpha X^\mu)} = P^\alpha_\mu.$$
These are simply the canonical momenta we defined earlier.

- The equation for current conservation is
$$\partial_\alpha P^\alpha_\mu = \frac{\partial P^\tau_\mu}{\partial\tau} + \frac{\partial P^\sigma_\mu}{\partial\sigma} = 0.$$
We know this holds on-shell, since it is precisely the equation of motion.

- The four conserved charges are found by integrating over "space" (i.e. σ) on the worldsheet,
$$p_\mu = \int_0^{\sigma_1} P^\tau_\mu(\tau,\sigma)\,d\sigma, \quad \frac{dp_\mu}{d\tau} = 0.$$

To check conservation, note that

$$\frac{dp_\mu}{d\tau} = \int_0^{\sigma_1} \frac{\partial P^\tau_\mu}{\partial \tau}\, d\sigma = -\int_0^{\sigma_1} \frac{\partial P^\sigma_\mu}{\partial \sigma}\, d\sigma = -P^\sigma_\mu \Big|_0^{\sigma_1}.$$

The right-hand side vanishes for closed strings and open strings with free endpoints. For open strings with Dirichlet boundary conditions, p_μ is not conserved, reflecting the fact that momentum can be transferred to the D-brane.

- Since p_μ it is associated with spacetime translation of the string, we expect it is the total energy-momentum of the string, and indeed in the static gauge $\tau = t$ we have $dp_\mu/dt = 0$ and p_μ is the total energy-momentum of the string, as we argued above. Since ϵ^μ is a vector, the index on p_μ is a vector index, so it transforms as expected under Lorentz transformations.
- To interpret p_μ for a general parametrization, we simply note that Noether charges $Q = \int_\Sigma (n \cdot j)$ are scalars which are independent of the surface Σ used to compute them, a fact proven in the lecture notes on General Relativity. Hence p_μ is always the conserved energy-momentum, in any parametrization. Furthermore, we need not compute it over a curve of constant τ. For an arbitrary curve γ, which is only required to wrap around the worldsheet once for a closed string or end at the endpoints for an open string, we have

$$p_\mu = \int_\gamma P_\mu \cdot dn = \int_\gamma P^\tau_\mu\, d\sigma - P^\sigma_\mu\, d\tau.$$

The relative sign is just because the normal to $(d\tau, d\sigma)$ is $dn = (d\sigma, -d\tau)$.

Next, we turn to Lorentz symmetry.

- As usual, Lorentz symmetry is generated by

$$\delta X^\mu = \epsilon^{\mu\nu} X_\nu, \quad \epsilon^{\mu\nu} = -\epsilon^{\nu\mu}.$$

The string Lagrangian contains terms of the form $\eta_{\mu\nu}(\partial X^\mu/\partial\xi^\alpha)(\partial X^\nu/\partial\xi^\beta)$. Since X^μ is a Lorentz vector and ξ is a Lorentz scalar, the two quantities in parentheses are Lorentz vectors, so the string Lagrangian is a Lorentz scalar as expected.

- Applying Noether's theorem, the conserved currents can be packaged into a Lorentz tensor,

$$\mathcal{M}^\alpha_{\mu\nu} = X_\mu P^\alpha_\nu - X_\nu P^\alpha_\mu.$$

The conserved charges can be computed similarly,

$$M_{\mu\nu} = \int_\gamma \mathcal{M}^\tau_{\mu\nu}\, d\sigma - \mathcal{M}^\sigma_{\mu\nu}\, d\tau.$$

As usual, these Lorentz tensors are antisymmetric. As with the momenta, these charges might not be conserved for an open string attached to a D-brane.

- Explicitly, integrating over a curve γ with constant τ,

$$M_{\mu\nu} = \int X_\mu P^\tau_\nu - X_\nu P^\tau_\mu\, d\sigma$$

and the charges associated with boosts are

$$M^{0i} = \int tP^{\tau i} - X^i P^{\tau 0}\, d\sigma = tp^i - X_{\text{CM}} E$$

where X_{CM} is the spatial center of mass of the string, and conservation of M^{0i} ensures the center of mass moves at a constant velocity.

Example. Consider the rotating string of the previous example. We have

$$\mathbf{P}^\tau = T_0 \frac{\partial \mathbf{X}}{\partial t} = T_0 \cos\frac{\pi\sigma}{\sigma_1}\left(-\sin\frac{\pi t}{\sigma_1}, \cos\frac{\pi t}{\sigma_1}\right).$$

The angular momentum is

$$J = M_{12} = \int_0^{\sigma_1} X_1 P_2^\tau - X_2 P_1^\tau\, d\sigma = \frac{\sigma_1 T_0}{\pi} \int_0^{\sigma_1} \cos^2\frac{\pi\sigma}{\sigma_1}\, d\sigma = \frac{\sigma_1^2 T_0}{2\pi}.$$

Finally, using $\sigma_1 = E/T_0$ we have the relation

$$\boxed{\frac{J}{\hbar} = \alpha' E^2, \quad \alpha' = \frac{1}{2\pi T_0 \hbar c}}$$

where we restored $\hbar$ and c. Here α' is called the slope parameter, and was important in ancient times when string theory was used to predict the masses of mesons, which lied on "Regge trajectories" with slope α'. The relation $J \sim E^2$ is quite unusual, and comes from the fact that the mass of a string scales with its length. Given α', we may also define a length scale

$$\boxed{\ell_s = \hbar c\sqrt{\alpha'}}$$

called the string length. Most modern work on string theory uses ℓ_s or α', rather than T_0. Originally, $\ell_s \sim 10^{-15}$ m to describe mesons, but this was invalidated by deep inelastic scattering experiments, which instead supported QCD. In the modern usage of string theory as a theory of quantum gravity, ℓ_s is around the Planck length.

1.4 Light Cone Gauge

We now introduce light cone gauge, which will be used to carry out the quantization of the string.

- Previously, we worked in static gauge, $X^0(\tau, \sigma) = \tau$. We now consider the more general gauge

  $$n_\mu X^\mu(\tau, \sigma) = \lambda\tau$$

  for a constant vector n_μ, which reduces to static gauge for $n_\mu = (1, 0, 0, 0)$ and $\lambda = 1$. In this gauge, the configuration of the string at worldsheet time τ is the intersection of the worldsheet and a hyperplane orthogonal to n_μ. To ensure this configuration is everywhere spacelike or null, we restrict n_μ to be spacelike or null.

- Like static gauge, this gauge is not Lorentz covariant. In a more advanced treatment, string quantization would be carried out in a Lorentz covariant gauge, but the choice made here will be simpler to understand.

- Restricting to open strings, we recall that the string momentum p^μ is conserved for free boundary conditions, as P^σ vanishes at the endpoints. For Dirichlet boundary conditions, $n \cdot p$ is conserved as long as $n \cdot P^\sigma$ vanishes at the endpoints, a weaker condition which we will assume holds.
- In either case, $n \cdot p$ is conserved, and we conventionally normalize τ to be dimensionless, with

$$n \cdot X = 2\alpha'(n \cdot p)\tau.$$

  Since both sides are proportional to n, the gauge no longer depends on the normalization of n. To check the dimensions work out, note that $[X] = -1$, $[p] = 1$, $[\tau] = 0$, and $[\alpha'] = -2$.
- Previously, we parametrized σ so that the energy density $P^{\tau 0}$ was constant. This is generalized to demanding the energy-momentum density $n \cdot P^\tau$ in the n direction is constant. Scaling so that $\sigma \in [0, \pi]$ for open strings, this means

$$n \cdot P^\tau = \frac{n \cdot p}{\pi}.$$

  That is, the left-hand side does not depend on σ or τ.
- Dotting both sides of the equation of motion with n^μ, we find

$$\frac{\partial}{\partial \tau}(n \cdot P^\tau) + \frac{\partial}{\partial \sigma}(n \cdot P^\sigma) = \frac{\partial}{\partial \sigma}(n \cdot P^\sigma) = 0$$

  However, by assumption $n \cdot P^\sigma = 0$ at the endpoints, and hence

$$n \cdot P^\sigma = 0$$

  everywhere on the string.
- For closed strings, the momentum p is conserved in all cases, so the same reasoning as above goes through, with slightly different conventions,

$$n \cdot X = \alpha'(n \cdot p)\tau, \quad \sigma \in [0, 2\pi], \quad n \cdot P^\tau = \frac{n \cdot p}{2\pi}.$$

  Since there are no endpoints, it is ambiguous where $\sigma = 0$ is; furthermore we cannot use the endpoints to show $n \cdot P^\sigma = 0$. These two problems may be solved simultaneously by choosing the curve $\sigma = 0$ so that $n \cdot P^\sigma = 0$ on it.
- More explicitly, using the explicit expression for P^σ we have

$$n \cdot P^\sigma \propto (\dot{X} \cdot X')\partial_\tau(n \cdot X) - (\dot{X}^2)\partial_\sigma(n \cdot X) = (\dot{X} \cdot X')\partial_\tau(n \cdot X) \propto \dot{X} \cdot X'.$$

  Hence picking $n \cdot P^\sigma = 0$ is equivalent to setting $\dot{X} \cdot X' = 0$. This generalizes the condition $\dot{\mathbf{X}} \cdot \mathbf{X}' = 0$ we imposed in static gauge, as there we had $(X^0)' = 0$.

We now turn to the associated constraints and wave equations.

- Using $\dot{X} \cdot X' = 0$, the expression for P^τ simplifies to

$$P^{\tau\mu} = \frac{1}{2\pi\alpha'} \frac{X'^2 \dot{X}^\mu}{\sqrt{-\dot{X}^2 X'^2}}.$$

Dotting n into both sides and applying the gauge condition, we have

$$1 = \frac{X'^2}{\sqrt{-\dot{X}^2 X'^2}}$$

for both open and closed strings, and using $X'^2 > 0$ we have

$$\dot{X}^2 + X'^2 = 0$$

which generalizes the normalization condition $(\partial \mathbf{X}/\partial\sigma)^2 + (\partial \mathbf{X}/\partial t)^2 = 1$ in static gauge. These results together imply the induced metric on the worldsheet is conformally equivalent to the flat metric, so this gauge is also called conformal gauge.

- Using this result, the expressions for the momenta simplify considerably,

$$P^{\tau\mu} = \frac{\dot{X}^\mu}{2\pi\alpha'}, \quad P^{\sigma\mu} = -\frac{X^{\mu\prime}}{2\pi\alpha'}.$$

The equation of motion are simply wave equations,

$$\ddot{X}^\mu - X^{\mu\prime\prime} = 0.$$

- Finally, adding and subtracting $2\dot{X} \cdot X' = 0$, our two constraints are equivalent to

$$(\dot{X} \pm X')^2 = 0.$$

For an open string with free endpoints, we have the further constraint $P^{\sigma\mu} = 0$ at the endpoints, so $X^{\mu\prime}$ vanishes at the endpoints.

We will now find the general solution for the open string motion, going further than in static gauge.

- Now consider an open string with free endpoints. The general solution of the wave equation is

$$X^\mu(\tau, \sigma) = \frac{1}{2}(f^\mu(\tau + \sigma) + g^\mu(\tau - \sigma)).$$

The constraint $X' = 0$ at $\sigma = 0$ yields $f^{\mu\prime} = g^{\mu\prime}$, so the two differ by a constant, which can be absorbed in a redefinition of f^μ. Hence

$$X^\mu(\tau, \sigma) = \frac{1}{2}(f^\mu(\tau + \sigma) + f^\mu(\tau - \sigma)).$$

- The constraint $X' = 0$ at $\sigma = \pi$ yields

$$f^{\mu\prime}(\tau + \pi) - f^{\mu\prime}(\tau - \pi) = 0$$

which shows that $f^{\mu\prime}$ is periodic with period 2π, justifying our earlier normalization convention.

- So far, this is familiar from static gauge; we now go further by expanding $f^{\mu\prime}$ in a Fourier series,

$$f^{\mu\prime}(u) = f_1^\mu + \sum_{n>0} a_n^\mu \cos nu + b_n^\mu \sin nu$$

which integrates to

$$f^\mu(u) = f_0^\mu + u f_1^\mu + \sum_{n>0} A_n^\mu \cos nu + B_n^\mu \sin nu$$

and gives a general solution of

$$X^\mu(\tau, \sigma) = f_0^\mu + f_1^\mu \tau + \sum_{n>0} (A_n^\mu \cos n\tau + B_n^\mu \sin n\tau) \cos n\sigma.$$

- It is useful to change variables to

$$A_n^\mu \cos n\tau + B_n^\mu \sin n\tau = -i\frac{\sqrt{2\alpha'}}{\sqrt{n}}(a_n^{\mu *} e^{in\tau} - a_n^\mu e^{-in\tau})$$

so that the a_n^μ are dimensionless; it will turn into an annihilation operator upon quantization. Moreover, since the momentum density is

$$P^{\tau\mu} = \frac{1}{2\pi\alpha'}\dot{X}^\mu = \frac{1}{2\pi\alpha'}f_1^\mu + \text{oscillatory terms}$$

we have $f_1^\mu = 2\alpha' p^\mu$. In these variables, the solution takes the form

$$X^\mu(\tau,\sigma) = x_0^\mu + 2\alpha' p^\mu \tau - i\sqrt{2\alpha'}\sum_{n>0}(a_n^{\mu *} e^{in\tau} - a_n^\mu e^{-in\tau})\frac{\cos n\alpha}{\sqrt{n}}.$$

This is the notation one might see in a standard string theory textbook.

- It is also useful to define the scaled variables

$$\alpha_0^\mu = \sqrt{2\alpha'}p^\mu, \quad \alpha_n^\mu = a_n^\mu \sqrt{n}, \quad \alpha_{-n}^\mu = (\alpha_n^\mu)^*$$

so the sum ranges over all nonzero n,

$$X^\mu(\tau,\sigma) = x_0^\mu + \sqrt{2\alpha'}\alpha_0^\mu \tau + i\sqrt{2\alpha'}\sum_{n\neq 0}\alpha_n^\mu e^{-in\tau}\frac{\cos n\sigma}{n}.$$

The first term is called the position zero mode.

- In particular, in this notation the derivatives of X^μ are particularly simple,

$$\dot{X}^\mu = \sqrt{2\alpha'}\sum_n \alpha_n^\mu e^{-in\tau}\cos n\sigma, \quad X^{\mu\prime} = -i\sqrt{2\alpha'}\sum_n \alpha_n^\mu e^{-in\tau}\sin n\sigma.$$

This gives two nice linear combinations,

$$\dot{X}^\mu \pm X^{\mu\prime} = \sqrt{2\alpha'}\sum_n \alpha_n^\mu e^{-in(\tau\pm\sigma)}.$$

Note that at this point we have not yet imposed the constraints $(\dot{X} \pm X')^2 = 0$.

Finally, we impose the constraints by specializing to light cone gauge.

- Light cone gauge is the choice

$$n_\mu = (1/\sqrt{2}, 1/\sqrt{2}, 0, 0), \quad n\cdot X = X^+, \quad n\cdot p = p^+$$

in which case the gauge conditions become

$$X^+ = \beta\alpha' p^+ \tau, \quad p^+ = \frac{2\pi}{\beta}P^{\tau +}, \quad \beta = \begin{cases} 2 & \text{open string,} \\ 1 & \text{closed string.} \end{cases}$$

We define all the coordinates but the first two to be transverse coordinates X^I, so the metric restricted to transverse coordinates is Euclidean.

- The constraints take the form

$$-2(\dot{X}^+ \pm X^{+\prime})(\dot{X}^- \pm X^{-\prime}) + (\dot{X}^I \pm X^{I\prime})^2 = 0.$$

In light cone gauge we have $X^{+\prime} = 0$ and $\dot{X}^+ = \beta\alpha p^+$, giving

$$\dot{X}^- \pm X^{-\prime} = \frac{1}{\beta\alpha'}\frac{1}{2p^+}(\dot{X}^I \pm X^{I\prime})^2.$$

Here we have assumed $p^+ \neq 0$, which holds for almost all configurations. Hence in light cone gauge, we may easily solve for the derivatives of X^-. The key reason this is easy is that in light cone coordinates, we get inner product expressions like X^+X^- rather than squares, so we avoid having to take square roots.

- We can hence solve for the derivatives of X^- in terms of X^I, so we know X^- up to an integration constant x_0^-. Note that for a closed string we also have the consistency condition

$$\int_0^{2\pi} d\sigma\, X^{-\prime} = 0.$$

Hence the full evolution of the string is determined by the X^I and the constants p^+ and x_0^-.

- Going back to our earlier solution for the motion of the open string, we have

$$X^I(\tau,\sigma) = x_0^I + \sqrt{2\alpha'}\,\alpha_0^I\tau + i\sqrt{2\alpha'}\sum_{n\neq 0}\alpha_n^I e^{-in\tau}\frac{\cos n\sigma}{n}$$

where we restricted to the transverse coordinates. The plus component is simply

$$X^+(\tau,\sigma) = 2\alpha' p^+\tau$$

which is equivalent to setting the position zero mode x_0^+ and the α_n^+ (for $n \neq 0$) all to zero.

- Now we expand the minus component in the same way,

$$X^-(\tau,\sigma) = x_0^- + \sqrt{2\alpha'}\,\alpha_0^-\tau + i\sqrt{2\alpha'}\sum_{n\neq 0}\alpha_n^- e^{-in\tau}\frac{\cos n\sigma}{n}.$$

Using our earlier identity for $\dot{X} \pm X'$, we may solve for the α_n^-,

$$\sqrt{2\alpha'}\,\alpha_n^- = \frac{1}{p^+}L_n^\perp, \quad L_n^\perp = \frac{1}{2}\sum_p \alpha_{n-p}^I\alpha_p^I$$

where $L_n^\perp$ is called the transverse Virasoro mode. We have now found the general solution for the motion of the open string.

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

$$\sqrt{2\alpha'}\,\alpha_0^- = 2\alpha' p^- = \frac{1}{p_+}L_0^\perp$$

which implies that

$$2p^+p^- = \frac{1}{\alpha'}L_0^\perp = \frac{1}{\alpha'}\left(\frac{1}{2}\alpha_0^I\alpha_0^I + \sum_{n>0}\alpha_n^{I*}\alpha_n^I\right) = p^Ip^I + \frac{1}{\alpha'}\sum_{n>0} n a_n^{I*}a_n^I.$$

In particular, the mass of a string is

$$M^2 = -p^2 = 2p^+p^- - p^Ip^I = \frac{1}{\alpha'}\sum_{n>0} n a_n^{I*} a_n^I.$$

This confirms that $M^2 > 0$, which is actually hard to show without light cone gauge. Upon quantization, this formula will yield a discrete spectrum of masses. As a check, note that when the a_n^I all vanish, the string collapses to a point, and $M = 0$ accordingly.

1.5 Light Cone Fields

We now briefly consider the quantization of fields in light cone coordinates and light cone gauge.

- Writing the spacetime coordinates as $(x^+, x^-, \mathbf{x}_T)$, the Klein–Gordan equation is

$$(\partial^2 - m^2)\phi = \left(-2\partial_{x^+}\partial_{x^-} + \partial_{x^I}\partial_{x^I} - m^2\right)\phi = 0$$

  where I indexes over transverse coordinates.

- As usual, we can simplify this by Fourier transforming the spatial coordinates only, which we take to be x^- and $\mathbf{x}_T$. The conjugate momentum to x^- is $-p^+$, while the conjugate momentum to $\mathbf{x}_T$ is $\mathbf{p}_T$, so we define

$$\phi(x^+, x^-, \mathbf{x}_T) = \int đp^+ đ\mathbf{p}_T\, e^{-ix^-p^+ + i\mathbf{x}_T\cdot\mathbf{p}_T}\phi(x^+, p^+, \mathbf{p}_T)$$

  which converts the Klein–Gordan equation to

$$\left(i\partial_{x^+} - \frac{p^Ip^I + m^2}{2p^+}\right)\phi = 0.$$

  As usual, this equation simply enforces the mass-shell condition $p^2 + m^2 = 0$, as it implies

$$p^- = \frac{p^Ip^I + m^2}{2p^+}.$$

- As usual, we can quantize the field, constructing a creation and annihilation operator for every spatial momentum $\mathbf{p}$, where E is constrained to be positive and on the mass shell,

$$\phi = \int \frac{đ\mathbf{p}}{\sqrt{2E_\mathbf{p}}}\, a_\mathbf{p} e^{ipx} + a_\mathbf{p}^\dagger e^{-ipx}, \quad P^\mu = \int đ\mathbf{p}\, P^\mu = \int đ\mathbf{p}\, p^\mu a_\mathbf{p}^\dagger a_\mathbf{p}.$$

  Similarly in light cone coordinates, we label the operators by $(p^+, \mathbf{p}_T)$, and pick out the physical half of the mass shell by $p^+ > 0$. Then we have, for example,

$$P^+ = \sum_{p^+,\mathbf{p}_T} p^+ a^\dagger_{p^+\mathbf{p}_T} a_{p^+\mathbf{p}_T}, \quad \mathbf{P}^I = \sum_{p^+,\mathbf{p}_T} \mathbf{p}^I a^\dagger_{p^+\mathbf{p}_T} a_{p^+\mathbf{p}_T}, \quad P^- = \sum_{p^+,\mathbf{p}_T} \frac{p^Ip^I + m^2}{2p^+} a^\dagger_{p^+\mathbf{p}_T} a_{p^+\mathbf{p}_T}.$$

- Now we consider massless vector fields, which have the gauge symmetry

$$\delta A_\mu = \partial_\mu \epsilon.$$

  Alternatively, for the Fourier transform of the field,

$$\delta A_\mu(p) = ip_\mu\epsilon(p), \quad A^\mu(x) = \int đp\, e^{ipx} A^\mu(p).$$

- In light cone coordinates we have $\delta A^+ = ip^+\epsilon$. As usual, neglecting the set of measure zero where $p^+ = 0$, we may use the gauge freedom to reach light cone gauge,

$$A^+(p) = 0.$$

  Just like for temporal or axial gauge, this fixes the gauge symmetry up to a set of measure zero.

- The equation of motion is $\partial_\mu(\partial \cdot A) = \partial^2 A_\mu$, so for $\mu = +$ we find $\partial \cdot A = 0$. Hence

$$p \cdot A = -p^+ A^- - p^- A^+ + p^I A^I = 0.$$

  This hence determines A^- in terms of the transverse A^I,

$$A^- = \frac{p^I A^I}{p^+}$$

  which is reminiscent of the string in light cone gauge.

- Furthermore, the field equation reduces to $\partial^2 A_\mu = 0$, so the degrees of freedom are massless. They are contained in the transverse components only, $p^2 A^I = 0$, so that a massless vector field in D dimensions has $D-2$ degrees of freedom for each momentum.

- This is consistent with group theory, as the little group is E_{D-2}. Translations must act trivially to get a finite-dimensional representation of the little group, so we only need worry about $SO(D-2) \subseteq E_{D-2}$, and evidently the states for a given momentum transform in the fundamental of $SO(D-2)$, indicating spin one.

- We can also see the degrees of freedom are massless without fixing a gauge. Suppose that $A(p)$ has support only on $p^2 \neq 0$. Then the equation of motion is

$$p^2 A_\mu = p_\mu (p \cdot A)$$

  which may be solved for A_μ,

$$A_\mu = ip_\mu \left(\frac{ip \cdot A}{p^2} \right).$$

  However, this implies A_μ is pure gauge, related to the zero field by $\epsilon = ip \cdot A/p^2$.

- The quantization of the field is similar to that of the real scalar field. The creation and annihilation operators are now labeled by p^+, $\mathbf{p}_T$, and a transverse index I. The one-photon states take the form

$$\sum_{I=2}^{D-1} \xi_I {a^I_{p^+\mathbf{p}_T}}^\dagger |\Omega\rangle.$$

Next, we consider the more complicated case of gravitational fields.

- We will linearize about flat spacetime, $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, in which case the linearized Einstein equations for $h_{\mu\nu}$ in momentum space are

$$S^{\mu\nu}(p) \equiv p^2 h^{\mu\nu} - p_\alpha (p^\mu h^{\nu\alpha} + p^\nu h^{\mu\alpha}) + p^\mu p^\nu h = 0, \quad h = h^\mu_\mu.$$

  Here, indices on h are raised and lowered by the Minkowski metric.

- The equation of motion is invariant under the infinitesimal gauge transformation
$$\delta h^{\mu\nu} = ip^\mu \epsilon^\nu + ip^\nu \epsilon^\mu$$
where the gauge parameter is now a vector. Physically the gauge symmetry is reparametrization invariance.

- In light cone coordinates, the components of the metric are
$$(h^{IJ}, h^{+I}, h^{-I}, h^{+-}, h^{++}, h^{--}).$$
Note that in particular, we have
$$\delta h^{++} = 2ip^+\epsilon^+, \quad \delta h^{+-} = i(p^+\epsilon^- + p^-\epsilon^+), \quad h^{+I} = i(p^+\epsilon^I + p^I\epsilon^+).$$
Again ignoring degrees of freedom with $p^+ = 0$, we can choose ϵ^+ to set $h^{++} = 0$, we can choose ϵ^- to set $h^{+-} = 0$, and we can choose ϵ^I to set $h^{+I} = 0$. Then in light cone gauge
$$h^{++} = h^{+-} = h^{+I} = 0.$$

- Setting $\mu = \nu = +$ in the equations of motion gives $(p^+)^2 h = 0$, which implies $h = 0$, or equivalently $h^{II} = 0$. Plugging this back into the equation of motion,
$$p^2 h^{\mu\nu} - p^\mu p_\alpha h^{\nu\alpha} - p^\nu p_\alpha h^{\mu\alpha} = 0.$$
In particular, setting $\mu = +$ we have $p_\alpha h^{\nu\alpha} = 0$. Plugging this back in, we have $p^2 h^{\mu\nu} = 0$. Therefore, the equations of motion boil down to
$$h^{II} = 0, \quad p_\alpha h^{\mu\alpha} = 0, \quad p^2 h^{\mu\nu} = 0.$$

- Setting $\mu = I$ in the second condition above, we have
$$-p^+ h^{I-} + p_J h^{IJ} = 0$$
while setting $\mu = -$ gives
$$-p^+ h^{--} + p_I h^{-I} = 0.$$
These indicate that the h^{I-} and h^{--} are determined in terms of h^{IJ}.

- We thus conclude the degrees of freedom are all massless, embedded in a symmetric, traceless, transverse tensor field in $D-2$ dimensions. There are hence
$$n(D) = \frac{1}{2}(D-2)(D-1) - 1 = \frac{1}{2}D(D-3)$$
degrees of freedom per momentum. As with the photons, this is consistent with a little group analysis, assuming the gravitons have spin two. Of course we could have gotten this much more easy by simply subtracting D from $D(D-1)/2$, but it's good to do an explicit check.

- The one-particle states can hence be written as
$$\sum_{I,J=2}^{D-1} \xi_{IJ} a^{IJ}_{p^+\mathbf{p}_T}{}^\dagger |\Omega\rangle, \quad \xi_{II} = 0$$
where ξ_{IJ} is the graviton polarization tensor.

Example. The Kalb–Ramond field is a two-form gauge field B, with an associated field strength $H = dB$ and gauge symmetry $\delta B = d\epsilon$. Explicitly, in components,

$$H_{\mu\nu\rho} = \partial_\mu B_{\nu\rho} + \partial_\nu B_{\rho\mu} + \partial_\rho B_{\mu\nu}, \quad \delta B_{\mu\nu} = \partial_\mu \epsilon_\nu - \partial_\nu \epsilon_\mu.$$

Note that the gauge parameter itself has a gauge symmetry, $\delta\epsilon = \partial_\mu \lambda$. By the same argument as above, we can use up this gauge symmetry to set $\epsilon^+ = 0$. The action and equation of motion are

$$S = -\frac{1}{6} \int dx\, H_{\mu\nu\rho} H^{\mu\nu\rho}, \quad \partial^\mu H_{\mu\nu\rho} = 0.$$

The degrees of freedom are in $(B^{+-}, B^{+I}, B^{-I}, B^{IJ})$ where B^{IJ} is antisymmetric, and the gauge freedom is enough to set the first two to zero. The equation of motion then simplifies to $p^2 B^{IJ} = 0$, giving $(D-2)(D-3)/2$ degrees of freedom per momentum.

1.6 The Veneziano Amplitude

In this section we'll cover the early history of string theory, when it was used as a theory of the strong interaction. This history is still regarded as important today, despite the advent of QCD, for it constitutes the first and only prediction of string theory.

- During the 1950s and 1960s, many hadrons were discovered. It was found that many particles lay on lines on a plot of M^2 versus spin J, called a Chew–Frautschi plot. These lines were called Regge trajectories. The Regge trajectory of lowest mass was parametrized as

$$J = \alpha' M^2 + \alpha(0).$$

  This was quite mysterious, and a theory that explained the presence of all of these particles was required.

- Furthermore, it was known that fundamental, high-spin particles have problematic features in the UV. Consider the scattering of scalar particles, $\phi\phi \to \phi\phi$, with four-momenta directed inward. The Mandelstam invariants are

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

- Now consider tree-level processes mediated by a particle of spin J. Roughly speaking, the interaction term must look like

$$\delta\mathcal{L} \sim g_J (\phi \partial_{\mu_1} \ldots \partial_{\mu_J} \phi) \sigma^{(\mu_1 \ldots \mu_J)}$$

  where the derivatives act both to the left and the right. That is, we require powers of momentum to contract with the indices. But then the t-channel exchange would yield

$$\mathcal{A} \sim \frac{g_J^2\, s^J}{t - m_\sigma^2}.$$

  In the Regge limit of fixed (negative) t and large s, the scattering amplitude grows as s^J. This is called "hard" behavior, but was not observed for the scattering of pions. It is also difficult to understand theoretically, as it violates perturbative unitarity bounds.

- Furthermore, loop amplitudes diverge very strongly. If we consider the one-loop diagram consisting of the exchange of two σ's, we have

$$\mathcal{A} \sim \int đp \, \frac{p^{4J}}{(p^2)^4}.$$

This suggests that we cannot write down a renormalizable theory involving the high-spin σ particles, which indeed is the case.

- With regard to the tree-level amplitude, we should really sum over all particles,

$$\mathcal{A}_t \sim \sum_{J=0}^{\infty} \frac{g_J^2 \, s^J}{t - m_J^2}.$$

Similarly, we can sum over all s-channel diagrams, yielding $\mathcal{A}_s$. For simplicity, we suppose the scalars are not identical, so we can ignore the u-channel.

- Dolan, Horn, and Schmid were inspired by data to guess the complete tree-level amplitude is

$$\mathcal{A} = \mathcal{A}_s = \mathcal{A}_t.$$

That is, one can sum over *either* s-channel diagrams or t-channel ones, and they will give the same result. This would be impossible for usual scattering processes, because $\mathcal{A}_s$ only has poles in s and $\mathcal{A}_t$ only has poles in t, but the infinite sum can change the analytic structure. This led to the development of "dual models".

- In 1968, Veneziano guessed that the amplitude had the form

$$\mathcal{A} = \frac{\Gamma(-\alpha(s))\Gamma(-\alpha(t))}{\Gamma(-\alpha(s) - \alpha(t))} = B(-\alpha(s), -\alpha(t)), \quad \alpha(x) = \alpha(0) + \alpha' x$$

where B is the Euler beta function. The gamma function has simple poles at zero and negative integers n with residues $(-1)^n/n!$.

- This implies that the Veneziano amplitude has only simple poles,

$$\text{s-channel poles: } s = \frac{-\alpha(0) + n}{\alpha'}, \quad \text{t-channel poles: } t = \frac{-\alpha(0) + n}{\alpha'}, \quad n = 0, 1, 2, \ldots.$$

Since $\mathcal{A}(s,t) = \mathcal{A}(t,s)$, the Taylor expansions in s and in t are identical, e.g.

$$\mathcal{A} = -\sum_{n=0}^{\infty} \frac{(\alpha(s)+1)(\alpha(s)+2)\ldots(\alpha(s)+n)}{n!} \frac{1}{\alpha(t) - n}$$

with the same expansion in powers of t, hence realizing Dolan–Horn–Schmid duality. Furthermore, the poles are just in the right places for the leading Regge trajectory.

- In the following year, Virasoro and Shapiro generalized the amplitude to display duality between the s, t, and u channels. Furthermore, these amplitudes do not have the undesirable "hard" behavior. In the Regge limit we instead have

$$A_{\text{dual}} \sim s^{\alpha(t)}$$

which is "soft" since $\alpha(t) < 0$. (In fact, one only finds this behavior if s is given a small imaginary part, which allows the amplitude to "average" over many resonances. This is physically correct, because the resonances really have finite lifetimes.)

- However, at higher energies, these amplitudes were ultimately found to be too soft; by 1974 QCD was recognized as the correct theory of the strong interaction. Another problem was that the amplitudes were not clearly unitary; as we can see from our explicit field theory amplitudes above, the residues must have the correct sign to represent scattering of physical particles; residues of the opposite sign indicate ghosts.
- It was found that the Veneziano amplitude is only unitary if spacetime has 26 dimensions and $\alpha(0) = 1$, which implies the existence of a massless spin 1 particle (i.e. a gauge boson) and a tachyonic scalar particle, which indicates an unstable vacuum.
- In 1974, Scherk and Schwarz found that in the low-energy limit $s\alpha' \ll 1$, a modified Veneziano amplitude (manually removing the tachyon), one finds Maxwell scattering of photons. Similarly, for the Virasoro–Shapiro amplitude, one finds a massless spin 2 particle undergoing gravitational scattering. Hence the theory giving rise to the Veneziano–Shapiro amplitude was a quantum theory of gravity with good UV behavior. Furthermore, the theory has only one parameter α', which could be taken to be on the order of the Planck scale.

Dolan-Horn-Schmid duality can be explained by the scattering of strings.

- First consider scattering of open strings, represented by the below diagram. This can be interpreted as either a tree-level s-channel or t-channel amplitude, by squeezing the diagram vertically or horizontally, explaining the duality.
- This should be compared to the Feynman diagrams in the worldline formulation of perturbative QFT, explained in the notes on Quantum Field Theory. Unlike in those Feynman diagrams, here there is no definite spacetime point where the strings split or merge. Each individual observer can define such a point, but it is not Lorentz invariant.
- The above diagram can't be interpreted as a u-channel diagram, as it would change the connectivity of the diagram. Instead, $s/t/u$-channel duality, as in the Veneziano–Shapiro amplitude, can be explained in terms of closed string scattering. This is easy to see by imagining shrinking the four tubes to points; we can then freely move the points around on the sphere.
- In QFT, vertices are distinguished points in the Feynman diagrams, which accounts for the diversity of QFTs; we can place a wide variety of factors at the vertices. However, in string theory there are no such points. Instead, about every point the string locally appears free. Hence the form of the free theory essentially determines the interactions, and anomaly cancellation strongly constrains the free theory, making string theory essentially unique.
- Another useful feature of string perturbation theory is that there are fewer string diagrams. For the four-point amplitude, there is only one closed string diagram at each order, as shown.
- As in worldline QFT, the diagrams above should be regarded as drawn in "worldsheet space", not physical space; the shapes are determined by the worldsheet metric. This allows us to use Weyl invariance on the worldsheet to deform the diagrams to a more convenient shape. For example, the tree-level closed string amplitude above can be deformed to a sphere with four punctures corresponding to external states.

2 Constrained Systems

2.1 Constrained Hamiltonian Systems

To understand the issues above more generally, we take a detour into classical mechanics.

- A gauge theory can be thought of as a theory where the dynamical variables are specified with respect to a reference frame whose choice is arbitrary at every instant of time. Physical observables are independent of this choice, but evolution is not deterministic; solutions to the equations of motion may contain arbitrary functions of time. It turns out that all gauge theories are constrained Hamiltonian systems (though not vice versa), as we will see shortly.

- In the Lagrangian formalism, we have

$$S_L = \int_{t_1}^{t_2} dt\, L(q, \dot{q}), \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^n} - \frac{\partial L}{\partial q^n} = 0.$$

  Using the chain rule, the Euler-Lagrange equation is equivalent to

$$\ddot{q}^{n'} \frac{\partial^2 L}{\partial \dot{q}^{n'} \partial \dot{q}^n} = \frac{\partial L}{\partial q^n} - \dot{q}^{n'} \frac{\partial^2 L}{\partial q^{n'} \partial \dot{q}^n}.$$

  This shows that the accelerations $\ddot{q}^{n'}$ are determined if and only if the matrix $J = \partial^2 L / \partial \dot{q}^{n'} \partial \dot{q}^n$ can be inverted. If J is singular, the solutions to the equations of motion contain arbitrary functions of time, i.e. we are dealing with a gauge system.

- To switch to the Hamiltonian formalism, we define the canonical momenta

$$p_n = \frac{\partial L}{\partial \dot{q}^n}.$$

  Then J is just the Jacobian matrix to switch from the $\dot{q}$ to the p, and its singularity means that different $\dot{q}$ will be mapped to the same p.

- Then the p are not independent, so there must be some "primary constraints"

$$\phi_m(q, p) = 0, \quad m = 1, \ldots, M.$$

  These constraints follow purely from the "kinematics", not from the equations of motion.

- We assume for simplicity that these constraints are independent everywhere and smooth functions of the q and p. When the rank of J is constant and equal to $N - M$, and the primary constraints define a submanifold of dimension $2N - M$ in phase space, called the primary constraint surface. Each point on this surface corresponds to a submanifold of dimension M in configuration space. Furthermore, we can set up local coordinates for phase space where M of the coordinates are the ϕ_m.

- If a phase space function G vanishes on the primary constraint surface, we say it weakly vanishes and write $G \approx 0$, which implies

$$G = g^m \phi_m$$

for some phase space functions g^m. Moreover, if $\lambda_n \delta q^n + \mu^n \delta p_n = 0$ for arbitrary variations on the primary constraint surface, then

$$\lambda_n = u^m \frac{\partial \phi_m}{\partial q^n}, \quad \mu^n = u^m \frac{\partial \phi_m}{\partial p_n}$$

for phase space functions u^m.

Next, we discuss the Hamiltonian.

- The Hamiltonian is only well-defined on the primary constraint surface, though we may extend it to the entire phase space arbitrarily. Then we expect the replacement

$$H \to H + c^m(q,p)\phi_m$$

should make no difference, as we'll see below.

- As usual, the variation of the Hamiltonian can be written in terms of δp and δq,

$$\delta H = \dot{q}^n \delta p_n - \delta q^n \frac{\partial L}{\partial q^n}.$$

The variations here are restricted to the primary constraint surface. The partial derivatives are a bit tricky; for functions defined on phase space, $\partial/\partial q$ keeps p constant, while for the Lagrangian, $\partial/\partial q$ keeps $\dot{q}$ constant.

- We may rewrite the equation above as

$$\left(\frac{\partial H}{\partial q^n} + \frac{\partial L}{\partial q^n}\right)\delta q^n + \left(\frac{\partial H}{\partial p_n} - \dot{q}^n\right)\delta p_n = 0.$$

Since δq^n and δp_n are arbitrary variations on the constraint surface,

$$\dot{q}^n = \frac{\partial H}{\partial p_n} + u^m \frac{\partial \phi_m}{\partial p_n}, \quad -\frac{\partial L}{\partial q^n} = \frac{\partial H}{\partial q^n} + u^m \frac{\partial \phi_m}{\partial q^n}.$$

The first equation lets us solve for $\dot{q}$ in terms of q, p, and the new functions u, which can be thought of as coordinates on the submanifold of inverse images of a given p_n. That is, the transformation from $(q, \dot{q})$ to (q, p, u) is invertible.

- The equations of motion are now

$$\dot{q}^n = \frac{\partial H}{\partial p_n} + u^m \frac{\partial \phi_m}{\partial p_n}, \quad \dot{p}_n = -\frac{\partial H}{\partial q^n} - u^m \frac{\partial \phi_m}{\partial q^n}, \quad \phi_m(q,p) = 0.$$

We recognize that these equations follow from the variational principle

$$\delta \int_{t_1}^{t_2} dt\, (\dot{q}^n p_n - H - u^m \phi_m) = 0$$

for arbitrary variations with fixed q endpoints; here u^m acts as a Lagrange multiplier that enforces the constraints. We can now see clearly that substituting $H \to H + c^m \phi_m$ changes nothing, since it just shifts the u^m.

- The equations of motion can also be written in terms of Poisson brackets,

$$\dot{F} \approx [F, H + u^m \phi_m] \approx [F, H] + u^m [F, \phi_m], \quad [F, G] = \frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i}$$

for phase space functions F and G.

- For consistency, the constraints must remain satisfied in time, $\dot{\phi}_m \approx 0$, which requires

$$[\phi_m, H] + u^{m'}[\phi_m, \phi_{m'}] \approx 0.$$

This may yield further constraints on the u's, called secondary constraints. The distinction between primary and secondary constraints is not important; we let there be J in total and write a generic one as $\phi_j \approx 0$.

- In any case, all J constraints lead to restrictions on the Lagrange multipliers u^m,

$$[\phi_j, H] + u^m [\phi_j, \phi_m] \approx 0.$$

These are linear in the u_m, and a general solution takes the form

$$u^m = U^m + V^m$$

where U^m is a particular solution of the inhomogeneous equations, and V^m solves the homogeneous equations $V^m[\phi_j, \phi_m] \approx 0$. Then a general solution can be expanded as

$$u^m = U^m + v^a(t) V_a^m$$

where the v^a are arbitrary functions of time and the V_a^m are a basis of homogeneous solutions. We have thus separated out the gauge freedom in the solution to the equations of motion.

- Therefore, the equation of motion on the constraint surface can be written as

$$\dot{F} \approx [F, H_T], \quad H_T = H' + v^a \phi_a = H + U^m \phi_m + v^a \phi_a, \quad \phi_a = V_a^m \phi_m$$

where H_T is called the total Hamiltonian.

- A phase space function F is first-class if

$$[F, \phi_j] \approx 0$$

for all j, and second-class otherwise. For example, H' is first-class, as are the ϕ_a, which give a complete set of first-class primary constraints.

- The Poisson bracket of first-class functions is first-class. To see this, note that if F and G are first-class, we may expand

$$[F, \phi_j] = f_j^{j'} \phi_{j'}, \quad [G, \phi_j] = g_j^{j'} \phi_{j'}.$$

By the Jacobi identity,

$$[[F, G], \phi_j] = [F, [G, \phi_j]] - [G, [F, \phi_j]] = [F, g_j^{j'} \phi_{j'}] - [G, f_j^{j'} \phi_{j'}] \approx 0.$$

Next, we connect first-class constraints to gauge transformations.

- When a physical state (q, p) is specified, the time evolution is not unique, because of the arbitrary functions of time v^a. We take as a postulate that time evolution should be unique; this implies that we must identify distinct points in phase space as being the same physical state, i.e. we have a gauge redundancy.
- For example, consider a dynamical variable F with some initial condition, with either some v^a or $\tilde{v}^a$. Then an infinitesimal time δt later, the final values of F differ by
$$\delta F = (v^a - \tilde{v}^a)\delta t[F, \phi_a].$$
Then we say the first-class primary constraints ϕ_a generate gauge transformations. It doesn't make sense to say second-class constraints do, because they take us off the constraint surface.
- Similarly, the Poisson bracket $[\phi_a, \phi_{a'}]$ of two first-class primary constraints generates a gauge transformation. This can be realized by "translating in a rectangular loop" in $(v^a, v^{a'})$ space over time.
- Finally, the Poisson bracket $[\phi_a, H']$ generates a gauge transformation; it is the difference of translating in time and incrementing v^a, or doing the same in reverse order.
- From the two previous points, we see that some secondary first-class constraints may also generate gauge transformations, if they are the results of such Poisson brackets. Dirac's conjecture says this is always the case, but it is false: there exist systems where we can get a deterministic time evolution without using all the secondary first-class constraints as gauge generators. On the other hand, from an axiomatic perspective it is useful to postulate Dirac's conjecture to be true. That is, we *define* all first-class constraints to be gauge generators. Thus from this point on we completely ignore the primary/secondary distinction.
- We denote all first-class constraints by γ and all second-class ones by χ. Then the most general time evolution allowing gauge transformations is generated by the extended Hamiltonian
$$H_E = H' + u^a \gamma_a$$
where the u^a are arbitrary functions of time. In comparison, the total Hamiltonian H_T only included first-class primary constraints.

2.2 Dirac Brackets and Gauge Fixing

Now we turn to the interpretation of second-class constraints. For simplicity, we consider the irreducible case, i.e. the case where all the constraints are independent.

- Define the matrix $C_{jj'} = [\phi_j, \phi_{j'}]$, whose elements are phase space functions. If we split the constraints into first-class and second-class constraints in order, the matrix has the block form
$$C_{jj'} \approx \begin{pmatrix} 0 & 0 \\ 0 & C_{\beta\alpha} \end{pmatrix}, \quad C_{\beta\alpha} = [\chi_\beta, \chi_\alpha].$$
The reduced matrix $C_{\beta\alpha}$ is invertible on the constraint surface, because if it were not, there would be a nonzero solution to $\lambda^\beta C_{\beta\alpha} \approx 0$, which would give a first-class constraint $\lambda^\beta \chi_\beta$. Since $C_{\beta\alpha}$ is antisymmetric, this implies there are an even number of second-class constraints.

- Consider the simple case where two conjugate variables q^1 and p_1 are constrained to be zero,

$$\chi_1 = q^1 \approx 0, \quad \chi_2 = p_1 \approx 0.$$

  They are second-class since $[\chi_1, \chi_2] = 1 \not\approx 0$. It is obvious in this case that the first pair of canonical variables simply plays no role.

- To make this manifest, we define the Dirac bracket, a modified Poisson bracket which does not include the first pair,

$$[F, G]^* = \sum_{n=2}^{N} \left(\frac{\partial F}{\partial q^n} \frac{\partial G}{\partial p_n} - \frac{\partial G}{\partial q^n} \frac{\partial F}{\partial p_n} \right).$$

  The Dirac bracket has the same nice properties as the Poisson bracket and still yields the time evolution, but the bracket of any phase space function with either of the χ_α is zero. Hence we can simply treat the χ_α as if they are strongly equal to zero, setting them to zero before evaluating the bracket.

- More generally, let $C^{\alpha\beta}$ be the inverse of $C_{\alpha\beta}$ and define the Dirac bracket

$$[F, G]^* = [F, G] - [F, \chi_\alpha] C^{\alpha\beta} [\chi_\beta, G].$$

  It can be verified explicitly that this bracket satisfies all the usual properties of the Poisson bracket, in addition to

$$[\chi_\alpha, F]^* = 0 \text{ for any } F.$$

- We also have

$$[F, R]^* \approx [F, R] \text{ for any first-class } F.$$

  This shows that H_E still generates time evolution under the Dirac bracket, and the γ_a still generate gauge transformations; the Dirac bracket replaces the Poisson bracket completely.

- Upon switching to the Dirac bracket, the second-class constraints effectively become strong relations between the canonical variables, so we may in principle eliminate the redundant ones. This was straightforward in our trivial example, but in practical situations it's often cleaner to keep them all.

We now turn to the question of gauge fixing.

- Similarly, for first-class constraints we may impose gauge-fixing conditions $C_b(q, p) \approx 0$ to remove the gauge freedom. Geometrically, for a complete gauge fixing, the gauge fixing surface must intersect each gauge orbit exactly once. However, in some circumstances this is impossible; this situation is called a Gribov obstruction.

- After gauge fixing, the original first-class constraints become second-class constraints, since they now take us off the new constraint surface. Conversely, one can think of all second-class constraints as arising from a gauge fixation. For example, in our trivial example with second-class constraints $q_1 = p_1 = 0$, we could regard $p_1 = 0$ as a first-class constraint which generates shifts in q_1, and $q_1 = 0$ as a gauge fixing condition. This is occasionally useful because it allows the use of Poisson brackets, which are simpler than Dirac brackets in quantization.

- In the infinite dimensional case, gauge fixing can become even more subtle. Consider the $u^a\gamma_a$ part of the extended Hamiltonian. In continuum mechanics, this becomes

$$\int dx\, u^a(x)\gamma_a(x)$$

and we must ask what function space the u^a live in; it must be large enough to impose the constraint $\gamma_a(x) \approx 0$ but no larger.

- To see what can go wrong, note that the u^a generate the gauge transformation

$$\delta F = \int dx\, u^a(x)[F, \gamma_a(x)].$$

In the case of electrodynamics, physical fields vanish at infinity; if we choose $u^a(x)$ to be constant we generate a global "charge rotation". The only states invariant under such a rotation are those of zero total charge.

- Another subtlety is the possibility of large gauge transformations, where the $u^a(x)$ are not continuously connected to the identity in function space. We may choose to regard them as proper gauge transformations, but this is an additional assumption, as everything we've done above is at the infinitesimal level.

- A classical observable F is a function on the constraint surface. It must be gauge-invariant, so

$$[F, \gamma_a]^* \approx 0.$$

Note that we also have $[F, \chi_a]^* = 0$ automatically.

Example. We consider the Lagrangian

$$L = \sum_{i=1}^{n-1} \frac{1}{2}(q_i - \dot{q}_{i+1})^2.$$

The canonical momenta are thus

$$\pi_i = \dot{q}^i - q^{i-1}, i \geq 2, \quad \pi_1 = 0, \quad H = \frac{1}{2}\sum_{i\geq 2} \pi_i^2 + \sum_{i \geq 2} \pi_i q_{i-1}.$$

The only primary constraint is $\pi_1 = 0$, so time evolution is generated by $H_T = H + u\pi_1$. Requiring $\dot{\pi}_1 = 0$ gives the secondary constraint $\pi_2 = 0$, which then gives $\pi_3 = 0$, and so on. Now all of the constraints π_i are first-class, so the extended Hamiltonian is

$$H_E = H + u^i \pi_i \approx 0$$

and the theory possesses no physical degrees of freedom. This system is equivalent to the system with Lagrangian $L = 0$, which has the same first-class constraints, except that they are all primary. This is another illustration of the fact that we need not distinguish between primary and secondary constraints.

2.3 General Covariance

Next, we consider the case of generally covariant systems, with no constraints for simplicity.

- Usually, one describes a Hamiltonian system by giving the canonical variables as a function of time t, where t is assumed to be directly physically measurable. In such cases, one can always promote t to a canonical variable by "parametrizing" the theory with a parameter τ, which then plays the same formal role that t originally did. The resulting system is generally covariant, having reparametrization invariance under τ.

- However, the interpretation of t and τ can be quite tricky. For example, general relativity is already generally covariant, as it is invariant under diffeomorphisms of spacetime, but we think of the τ–like coordinate as the physical time for some observer. For now we'll think of t as time and τ as a meaningless parameter, but will return to this point below.

- Explicitly, the action for a system with canonical variables q^i and p_i and Hamiltonian H_0 is

$$S[q^i(t), p_i(t)] = \int_{t_1}^{t_2} \left(p_i \frac{dq^i}{dt} - H_0 \right) dt.$$

  Now we let $t = q^0$ with conjugate momentum p_0. Then an equivalent action is

$$S[q^0(\tau), q^i(\tau), p_0(\tau), p_i(\tau), u^0(\tau)] = \int_{\tau_1}^{\tau_2} p_0 \dot{q}^0 + p_i \dot{q}^i - u^0 (p_0 + H_0)\, d\tau$$

  where the dot indicates a τ derivative.

- To show this, note that varying with respect to auxiliary variables u^0 and p_0 yields

$$\gamma_0 \equiv p_0 + H_0 = 0, \quad \dot{t} - u^0 = 0.$$

  These equations may be used to eliminate u^0 and p_0, to arrive at the action

$$\int_{\tau_1}^{\tau_2} p_i \dot{q}^i - H_0 \dot{t}\, d\tau = \int_{t_1}^{t_2} \left(p_i \frac{dq^i}{dt} - H_0 \right) dt$$

  as before. However, this equality only holds if t is monotonic in τ. Thus the covariant formulation is more general, as it can accommodate trajectories with $\dot{t} < 0$. In fact, even in the covariant path integral for a nonrelativistic particle, one must admit trajectories with $\dot{t} < 0$.

We now consider the consequences of our result.

- There is a single constraint, $\gamma_0 \approx 0$, which is thus first-class. The extended Hamiltonian above contains only the constraint term $-u^0\gamma_0$, so the Hamiltonian itself in this formalism is *zero*. This is not completely unreasonable, because physically systems evolve in time, not in the arbitrary parameter τ. The motion itself is solely "the unfolding of a gauge transformation".

- This procedure can be practically useful in systems with complicated explicit time dependence, since it always results in a system with no explicit dependence on τ.

- In this formalism, γ_0 generates a gauge transformation which is identified with time evolutions. Note that an infinitesimal reparametrization $\tau \to \overline{\tau} = \tau - \epsilon(\tau)$ induces the changes
$$\delta q = \dot{q}\epsilon, \quad \delta p = \dot{p}\epsilon, \quad \delta u^0 = \frac{d}{d\tau}(u^0\epsilon)$$
where ϵ must vanish at the endpoints. This is the gauge transformation generated by γ_0, up to a trivial "equation-of-motion symmetry".
- We will always require gauge transformations to vanish at the endpoints. This is really just an artifact of keeping the limits of integration fixed. The key point is that it sets total derivatives of terms proportional to ϵ to zero.
- One can argue very generally that H must vanish. We say that q and p transform as scalars under reparametrization invariance since they obey the equations above, while u^0 transforms as a scalar density. Then all terms in the integrand of the action transform as scalar densities, making the action a scalar. If a Hamiltonian were present as well, it would have to transform as a scalar density, but it must be a scalar since it is a function of q and p.
- However, there can be systems where q and p are not scalars, in which case H need not vanish. For example, this can be achieved by performing a τ-dependent canonical transformation.

Now we turn to the interpretation of the formalism.

- General covariance may be viewed as a special case of gauge symmetry, as in either case solutions to the equation of motion may contain arbitrary functions of the time τ. This implies that *something* about the system is unphysical, such as the time τ or some of the canonical variables, but we cannot decide which from the theory alone. Instead, additional information must come from outside.
- For example, in electromagnetism, we suppose the time parameter is physical while A_μ is not. This is justified because the electromagnetic field is just a subsystem of the universe, and we know we can build clocks that measure τ independently.
- On the other hand, for a classical point particle, we suppose the canonical variables $(t, \mathbf{x})$ are physical while τ is not; that is, we treat t and $\mathbf{x}$ as the measurable quantities.
- General relativity is the best-known generally covariant theory, but in this case there is no "outside perspective" we can take. In this case, the most symmetric formulation is one where the Hamiltonian is weakly zero, and all physical questions are formulated in terms of functions with zero brackets with the constraints; these first-class functions are gauge-invariant constants of the motion.
- Such functions suffice even to ask apparently time-dependent questions. For example, for the free particle, the quantity
$$q(\tau) - \frac{p(\tau)}{m}(t(\tau) - t_0)$$
does not depend on τ, and it is equal to the position of the particle at time t_0.

2.4 Constrained Quantization

Finally, we discuss the quantization of constrained Hamiltonian systems. There are many sophisticated quantization methods, such as BRST, but it will suffice to consider the simplest ones. First, we consider the case of second-class constraints.

- In canonical quantization, Poisson brackets are replaced with commutators. The resulting operators are then postulated to act irreducibly on a Hilbert space, allowing us to construct it. For example, a single (q, p) pair gives $\hat{q} = q$ and $\hat{p} = -i\hbar\,\partial/\partial q$ uniquely by the Stone-von Neumann theorem, leading to the Hilbert space $L^2(\mathbb{R})$.

- Second-class constraints are quantized by replacing the commutator with the Dirac bracket. For example, consider our trivial example again,

  $$\chi_1 = q_1 \approx 0, \quad \chi_2 = p_1 \approx 0.$$

  Naive canonical quantization would give $[\hat{p}_1, \hat{q}_1] = -i\hbar$ which is inconsistent with the constraint $\hat{q}_1 = \hat{p}_1 = 0$. But with the Dirac bracket, $[\hat{p}_1, \hat{q}_1] = 0$, and there is no issue in imposing the operators equations $\hat{q}_1 = \hat{p}_1 = 0$.

- The disadvantage of this method is that it may be difficult to find a representation of the Dirac brackets. After using the second-class constraints to eliminate redundant degrees of freedom, we will have independent variables $\hat{y}^i$ satisfying the commutation relations

  $$[\hat{y}^i, \hat{y}^j] = i\hbar\sigma^{ij}(\hat{y}^k).$$

  There is no general analogue of the Stone-von Neumann theorem that covers this case.

- However, as we've shown earlier, every second-class constraint can be turned into a first-class constraint by "undoing a gauge fixation", allowing us to return to Poisson brackets. Hence it also suffices to consider quantization of first-class constraints.

Next, we consider the quantization of first-class constraints.

- In reduced phase space quantization, we find a complete set of gauge-invariant functions and build the Hilbert space from those. For example, for a single first-class constraint $p_1 = 0$, a complete set of observables is $(q_2, p_2), \ldots, (q_N, p_N)$. All of these functions are gauge-invariant, and every function F obeying $[F, p_1] \approx 0$ is weakly equal to some function of them. Applying canonical quantization, the wavefunctions are functions of $q_2, \ldots, q_N$.

- In practice, finding such a complete set is quite difficult. Another way to carry out reduced phase space quantization is to perform a complete gauge fixing, reducing all remaining constraints to second class, which are handled with Dirac brackets. However, this has the same technical issues we saw above.

- The advantage of reduced phase space quantization is that every state in the Hilbert space is physical, and only gauge-invariant observables are realized as quantum mechanical operators. However, in practice this procedure is difficult and may destroy manifest invariance under an important symmetry, such as Lorentz symmetry. Furthermore, for field theories, the elimination of the gauge degrees of freedom generally destroys locality in space.

- In Dirac quantization, we simply naively canonically quantize everything, ignoring the constraints, then impose them by restricting to "physical states".
- Specifically, if the gauge generators are $\hat{G}_a$, then physical states should satisfy

$$e^{i\epsilon^a \hat{G}_a}|\psi\rangle = |\psi\rangle$$

or equivalently

$$\hat{G}_a|\psi\rangle = 0.$$

For example, for $p_1 = 0$, the Hilbert space contains wavefunctions $\psi(q_1, \ldots, q_N)$, and the physical state condition is $\partial\psi/\partial q_1 = 0$, equivalent to the reduced phase space result.
- At the classical level, the constraints G_a obey

$$[G_a, G_b] = C^c_{ab} G_c$$

and we expect this relation to be preserved quantum mechanically,

$$[\hat{G}_a, \hat{G}_b] = i\hbar \hat{C}^c_{ab} \hat{G}_c.$$

Then we will automatically have $[\hat{G}_a, \hat{G}_b]|\psi\rangle = 0$.
- However, it is possible that at the quantum level, there will be ordering ambiguities that make this impossible; instead we generally could have

$$[\hat{G}_a, \hat{G}_b] = i\hbar \hat{C}^c_{ab} \hat{G}_c + \hbar^2 \hat{D}_{ab}$$

and the physical states would have to obey $\hat{D}_{ab}|\psi\rangle = 0$ as well. This usually gives us far too few physical states; if we do not impose this condition, then we have a gauge anomaly: the gauge symmetry is broken at the quantum level, and the entire procedure above is not applicable.
- Similarly, at the classical level we have

$$[H_0, G_a] = V_a^b G_b$$

but at the quantum level we may have

$$[\hat{H}_0, \hat{G}_a] = i\hbar \hat{V}_a^b \hat{G}_b + \hbar^2 \hat{C}_a.$$

When $\hat{C}_a$ is nonzero, physical states are not closed under time evolution, spoiling the theory. However, the quantization may sometimes be carried out with a more advanced formalism such as BRST, where the ghosts play an essential role for consistency.

Dirac quantization can also be inconvenient because it is difficult to define a finite scalar product, as we can already see in our trivial example $p_1 = 0$ if q_1 has a noncompact range. The Dirac–Fock method avoids this issue, and works whenever there is an even number of first-class constraints. It is also called the Gupta–Bleuler method in field theory and string theory.

- We consider a system with N degrees of freedom and first-class constraints

$$p_1 = p_2 = 0.$$

If we define

$$a = p_1 + ip_2, \quad b = -\frac{i}{2}(q^1 + iq^2)$$

along with the conjugates a^* and b^*, then we have the Poisson brackets

$$[a, b^*] = [b, a^*] = -i$$

with all others zero, and the constraints are equivalent to $a = a^* = 0$.

- At the quantum level, defining $a_\mu = (a, b)$, we have

$$[a_\mu, a^*_\nu] = \eta_{\mu\nu}, \quad \eta_{\mu\nu} = \begin{pmatrix} 0 & 1 \\ 1 & 9 \end{pmatrix}$$

with all other commutators zero. Hence we have a set of two quantum harmonic oscillators with an indefinite metric.

- Defining the vacuum $|0\rangle$ to be annihilated by both a and b, we see that $a^* - b^*$ creates negative norm states while acting an odd number of times, $a^* + b^*$ creates positive norm states, and a^* and b^* each create states of zero norm. The creation operators generate an entire Fock space.

- The other physical degrees of freedom $(q^3, p_3), \ldots, (q^N, p_N)$ may be quantized as usual, giving wavefunctions $\psi(q^3, \ldots, q^N)$. A general state is the tensor product of one of these wavefunctions with a Fock state. There are hence no divergences when defining the norm of a state, but the norm may be negative.

- Next, we need to impose the constraints. Naively, we would demand

$$a|\psi\rangle = a^*|\psi\rangle = 0.$$

However, this leaves us with no physical states at all, because the raising operator a^* has no nullspace. Instead, we take the weaker condition

$$a|\psi\rangle = 0.$$

This is equivalent to demanding a physical state contains no b^* modes, and ensures that no negative-norm states are physical.

- We might wonder if this condition is sufficient. A general physical state may be written as

$$|\psi\rangle = f(q^3, \ldots, q^N)|0\rangle + |n\rangle$$

where $|n\rangle$ is a "null spurious" state containing a^* excitations but no b^* excitations. The $|n\rangle$ have zero norm, and in fact zero overlap with every physical state. This is because we may always write $|n\rangle = a^*|\chi\rangle$, and for any physical state $|\psi\rangle$,

$$\langle\psi|n\rangle = \langle\psi|a^*|\chi\rangle = 0.$$

- Therefore, the null spurious states completely disappear from any physical matrix element, so one can consistently factor them out. That is, one can identify two physical states that differ by a null spurious state as the same state. In particular, each distinct physical state has a representative of the form $f(q^3, \ldots, q^N)|0\rangle$, and hence a^* is equivalent to the zero operator, imposing the other half of the constraint. The inner product on the reduced state space is positive definite.

- More generally, in the context of field theory, the analogue of imposing $a|\psi\rangle = 0$ is to impose $\hat{G}^{(-)}|\psi\rangle = 0$, where $\hat{G}^{(-)}$ is the annihilation part of $\hat{G}$. For each situation, one must check that there are physical null spurious states that decouple, to recover the second half of the gauge invariant. This requirement fixes $D = 26$ in bosonic string theory.

2.5 Classical Point Particle

To warm up for quantizing the string, we quantize a relativistic particle.

- There are generally two routes to a quantum theory: we may canonically quantize particle or field degrees of freedom. These two routes are called first and second quantization, respectively. In second quantization, one ends up with a theory of many particles, where the one-particle sector matches the result of first quantization.

- In general, string theory takes the first approach. The downside is that this approach is necessarily perturbative. The analogous second quantized formalism is called string field theory, where strings arise as excitations of a string field; little is known about this complex subject.

- Viewing the configuration of a particle as a set of scalar fields on its worldline, the first quantized approach is formally analogous to a one-dimensional field theory. Similarly string theory is formally like a two-dimensional field theory.

- Formally, an elementary particle is a unitary irrep of the Poincare group, classified by its mass and spin. Physically, it is a particle without structure. Ignoring any internal degrees of freedom, the classical action of such a particle should hence only depend on properties of its worldline. Furthermore, dimensional analysis forbids any dependence on, e.g. the curvature of the worldline, as there are no other length scales.

- Given these assumptions, the unique relativistic particle action is the proper time,
$$S = -m \int dt\, \sqrt{1 - \dot{\mathbf{x}} \cdot \dot{\mathbf{x}}}.$$
To put time and space on an even footing, we can instead parametrize by τ,
$$S = -m \int d\tau\, \sqrt{-\dot{x}^\mu \dot{x}^\nu \eta_{\mu\nu}} = -m \int d\tau\, \sqrt{-\dot{x}^2}, \quad \dot{x} = \frac{dx}{d\tau}.$$
Here τ is an arbitrary, usually dimensionless parameter, and the action has reparametrization invariance, in the sense that $S[x'] = S[x]$ if $x'(\tau') = x(\tau)$ for any monotonic function $\tau'(\tau)$. In temporal gauge we set $\tau = t$, recovering our original action.

- The canonical momenta and equation of motion are
$$p_\mu = \frac{m\dot{x}_\mu}{\sqrt{-x^2}}, \quad \frac{dp_\mu}{d\tau} = 0.$$

In particular, this yields the primary first-class constraint

$$p^2 + m^2 = 0.$$

At this point, Dirac quantization yields wavefunctions $\phi(x^\mu)$ and the constraint $p^2 + m^2 = 0$ means the wavefunctions obey the Klein–Gordan equation.

- In addition, the Hamiltonian is identically zero,

$$H = \dot{x}^\mu p_\mu - L = \frac{m\dot{x}^2}{\sqrt{-\dot{x}^2}} + m\sqrt{-\dot{x}^2} = 0$$

because the "time" variable τ has reparametrization invariance.

- Alternatively, we may completely fix the gauge, i.e. use reduced phase space quantization. Light cone gauge is the choice

$$x^+ = \frac{1}{m^2} p^+ \tau.$$

In this case, the $+$ component of the equation of motion immediately gives

$$\dot{x}^2 = -\frac{1}{m^2}$$

which simplifies the momenta and equation of motion to

$$p_\mu = m^2 \dot{x}_\mu, \quad \ddot{x}_\mu = 0.$$

Since we have removed the reparametrization invariance, the Hamiltonian no longer vanishes.

- The primary constraint may be used to solve for p^-,

$$p^- = \frac{1}{2p^+}(p^I p^I + m^2).$$

The value of p^- then determines the evolution of x^-, up to an integration constant x_0^-. Furthermore, x^+ is determined by p^+. Hence the independent dynamical variables are (x^I, p^I, x_0^-, p^+). We can straightforwardly quantize these variables because we have removed the gauge freedom and accounted for all the constraints.

- The disadvantage of this method is that we lose explicit Lorentz invariance. If we pressed on with the gauge symmetry intact, with the accompanying constraints, then we must impose the constraints at the quantum level, as we saw in the previous section. In this context, this is called covariant quantization.

- It is also useful to rewrite the point particle action with an einbein $e(\tau)$,

$$S = \frac{1}{2}\int d\tau\, e^{-1}\dot{x}^2 - em^2.$$

This action has reparametrization invariance,

$$\tau \to \tau', \quad x(\tau) \to x'(\tau') = x(\tau), \quad e \to e' = \frac{d\tau}{d\tau'} e$$

which is, infinitesimally,

$$\delta\tau = -\eta, \quad \delta x^\mu = \frac{dx^\mu}{d\tau}\eta, \quad \delta e = \frac{d}{d\tau}(\eta e)$$

where $\eta(\tau)$ is arbitrary. It is advantageous because it has no square roots, which makes it easier to handle in the path integral, and it can handle massless particles just as well as massive ones.

- Naively, if we use the reparametrization invariance to set $e = 1$, then the equation of motion for x is simply $\ddot{x} = 0$. However, this isn't quite right, because we've forgotten about the equation of motion for e, which is
$$\dot{x}^2 + e^2 m^2 = 0.$$
In the massive case, this tells us $\dot{x}$ is normalized to be the four-momentum. In the massless case, it tells us that $\dot{x}$ is null.

- To return to our original action in the massive case, we solve the equation of motion of e for e, and plug it back into the action to eliminate it; this is possible since e is an auxiliary field.

- Formally, one can think of this action as corresponding to a one-dimensional quantum gravity theory. This is easier to see if we write $e = \sqrt{-g_{\tau\tau}}$, so
$$S = -\frac{1}{2}\int d\tau\,\sqrt{-g_{\tau\tau}}(g^{\tau\tau}\dot{x}^2 + m^2).$$
In other words, introducing the einbein was equivalent to introducing a worldline metric.

- When we quantize the string, our actions here will correspond to the Nambu–Goto and Polyakov actions. The Polyakov action can be thought of in terms of two-dimensional quantum gravity on the worldsheet, and we will try to quantize it both covariantly and in light cone gauge.

We now explicitly quantize the relativistic point particle in light cone gauge.

- Starting from the Lagrangian in light cone gauge, we can show that (x^I, p^I) and (x_0^-, p^+) are conjugate variable pairs, so that in canonical quantization we have
$$[x^I, p^J] = i\eta^{IJ}, \quad [x_0^-, p^+] = i\eta^{-+} = -i.$$
In Heisenberg picture, these commutators hold when the operators are evaluated at equal times. We can then define the redundant operators
$$x^+ = \frac{p^+}{m^2}\tau, \quad x^- = x_0^- + \frac{p^-}{m^2}\tau, \quad p^- = \frac{1}{2p^+}(p^I p^I + m^2).$$
Note that p^- has no explicit τ-dependence, though in Heisenberg picture it has τ-dependence via p^+ and p^I.

- We know that H generates τ translations, and we expect p^- to generate x^+ evolution. Since these are proportional, we have
$$H = \frac{p^+}{m^2}p^- = \frac{1}{2m^2}(p^I p^I + m^2).$$
Note that unusually, H is dimensionless, because τ is.

- It's easy to check the Heisenberg equations of motion match the classical Hamilton's equations. For example, we have

$$i\frac{dp^+}{d\tau} = [p^+, H] = 0, \quad i\frac{dp^I}{d\tau} = [p^I, H] = 0, \quad i\frac{dx^I}{d\tau} = [x^I, H] = i\frac{p^I}{m^2}$$

  where the last result gives

$$x^I = x_0^I + \frac{p^I}{m^2}\tau$$

  as expected. We also have

$$i\frac{dx_0^-}{d\tau} = [x_0^-, H] = 0$$

  which is as expected, since x_0^- is a constant of the motion.

- We can choose (p^+, p^I) as a maximum commuting set and hence label the states of the point particle by their eigenvalues, as $|p^+, \mathbf{p}_T\rangle$. In this basis, the Hamiltonian is diagonal.

- For a general state $|\psi\rangle$ we may define a wavefunction by

$$|\psi\rangle = \int dp^+ d\mathbf{p}_T\, \psi(p^+, \mathbf{p}_T)|p^+, \mathbf{p}_T\rangle$$

  and the wavefunction obeys the Schrodinger equation

$$i\frac{\partial}{\partial\tau}\psi = \frac{1}{2m^2}(p^I p^I + m^2)\psi.$$

  Of course, up to rescaling this matches the equation of motion for the Klein–Gordan field, providing an example of the equivalence of first and second quantization: the equation of motion for a classical field matches the equation of motion for the one-particle wavefunction of the second quantized field, which in turn matches the Schrodinger equation in first quantization.

- Historically, this coincidence of equations led to confusion, as physicists thought the classical field that was the starting point for second quantization was the first quantized wavefunction itself, leading to the name. This is conceptually incorrect since the first quantized theory is already quantum; there is no need to quantize it again. In the modern view, the equivalence of first and second quantization is so well-known that in condensed matter, the two are introduced as slightly different ways of describing the same theory, i.e. by many-body wavefunctions or occupation numbers.

To discuss conserved quantities, it will be useful to remove the gauge fixing.

- Without the gauge fixing, we have canonical commutators

$$[x^\mu, p^\nu] = i\eta^{\mu\nu}.$$

  This is quite different from quantization in light cone gauge. For instance, the commutator $[x^+, p^-]$ vanishes in light cone gauge but not here, where $[x^+, p^-] = i\eta^{+-} = -i$. In other words, the Poisson bracket structure in light cone gauge is not merely a restriction of the structure without gauge fixing. Conceptually, we must distinguish between objects in light cone gauge and objects merely written in light cone coordinates, which unfortunately have identical notation.

- As expected, the operators p^μ generate translations of the particle, so that

$$\delta x^\mu = [i\epsilon_\nu p^\nu, x^\mu] = \epsilon^\mu.$$

We would like to confirm the same thing occurs in light cone gauge. We may expand

$$i\epsilon_\nu p^\nu = -i\epsilon^- p^+ - i\epsilon^+ p^- + i\epsilon^I p^I.$$

It is clear that the p^I generate translations in x^I, and that p^+ generates translations in x^-.

- It is less clear that the same holds for p^-, since it is determined in terms of the other momenta. In this case we have

$$\delta x^\mu = -i\epsilon^+ [p^-, x^\mu].$$

Taking specific values for μ, we find

$$\delta x^+ = -i\epsilon^+[p^-, x^+] = 0, \quad \delta x^I = -i\epsilon^+[p^-, x^I] = -i\epsilon^+ \frac{1}{2p^+}(-2ip^I) = -\epsilon^+ \frac{p^I}{p^+}.$$

The trickiest commutator to compute is

$$\delta x^- = -i\epsilon^+ \left[p^-, x_0^- + \frac{p^-}{m^2}\tau\right] = -i\epsilon^+ [p^-, x_0^-] = \frac{p^I p^I + m^2}{2} \left[p^-, \frac{1}{p^+}\right].$$

To finish this evaluation, note that

$$\left[p^-, \frac{1}{p^+}\right] = \frac{1}{p^+}[p^+, x_0^-]\frac{1}{p^+} = \frac{i}{p^{+2}}.$$

In conclusion, we have

$$\delta x^+ = 0, \quad \delta x^I = -\epsilon^+ \frac{p^I}{p^+}, \quad \delta x^- = -\epsilon^+ \frac{p^-}{p^+}.$$

- This result is very different from what we expect. The resolution is that, even though we have removed the diffeomorphism symmetry, the action retains a symmetry under τ translations, which corresponds to

$$\delta x^\mu = \lambda \dot{x}^\mu.$$

In this case, the action of p^- generates a translation in x^+ plus a translation in τ by $\lambda = -m^2\epsilon^+/p^+$. This is necessary to set $\delta x^+ = 0$, which preserves the light cone gauge condition.

- Similarly, the infinitesimal Lorentz transformations and conserved charges take the form

$$\delta x^\mu = \epsilon^{\mu\nu} x_\nu, \quad M^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu.$$

It is straightforward to see the conserved charges generate the transformations in covariant quantization. In light cone gauge, we wish to construct similar operators which generate the same transformations (up to τ translations) and obey the Lorentz algebra.

- The calculation here is a bit involved, but it turns out to be possible for the point particle. However, it turns out that for the relativistic string, it is only possible if $D = 26$.

3 The Bosonic String

3.1 The Polyakov Action

We now introduce the Polyakov action, the analogue of the einbein action for strings.

- First, recall that the Nambu–Goto action can be written in terms of the worldsheet metric,

$$S = -\frac{1}{2\pi\alpha'} \int d\tau d\sigma \, \sqrt{-\gamma}.$$

  We can derive the equations of motion from this form directly, using

$$\delta\sqrt{-\gamma} = \frac{1}{2}\sqrt{-\gamma}\,\gamma^{\alpha\beta}\delta\gamma_{\alpha\beta}$$

  and the definition of $\gamma_{\alpha\beta}$, which gives

$$\partial_\alpha(\sqrt{-\gamma}\,\gamma^{\alpha\beta}\partial_\beta X^\mu) = 0.$$

- The Polyakov action removes the square root by introducing another field $g_{\alpha\beta}$ on the worldsheet,

$$S = -\frac{1}{4\pi\alpha'} \int d^2\sigma \, \sqrt{-g}\, g^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X^\nu \eta_{\mu\nu}.$$

  Note that here we restrict to flat spacetime, and σ conventionally stands for both worldsheet coordinates. The new field $g_{\alpha\beta}$ is a metric with signature $(-+)$, so this is a two-dimensional quantum gravity theory on the world sheet, interacting with worldsheet scalars X^μ.

- The symmetries allow us to include the Einstein term $\sqrt{-g}\,R$, but this is a total derivative in $1+1$ dimensions; it does not make the metric a dynamical field. We will ignore it for now, though it will have global consequences. A cosmological constant term $\sqrt{-g}$ is not allowed because it would break Weyl symmetry.

- The equation of motion for X^μ is simply

$$\partial_\alpha(\sqrt{-g}\,g^{\alpha\beta}\partial_\beta X^\mu) = \sqrt{-g}\,\nabla^2 X^\mu = 0$$

  which resembles the Nambu–Goto equation of motion, except that $g_{\alpha\beta}$ has its own dynamics.

- The equation of motion for $g_{\alpha\beta}$ is

$$\left(\sqrt{-g}\,\partial_\alpha X^\mu \partial_\beta X^\nu - \frac{1}{2}\sqrt{-g}\,g_{\alpha\beta}g^{\rho\sigma}\partial_\rho X^\mu \partial_\sigma X^\nu\right)\eta_{\mu\nu} = 0$$

  which allows us to solve for the worldsheet metric,

$$g_{\alpha\beta} = 2f(\sigma)\,\partial_\alpha X\cdot\partial_\beta X, \quad f^{-1} = g^{\rho\sigma}\,\partial_\rho X\cdot\partial_\sigma X.$$

  We see that $g_{\alpha\beta}$ matches $\gamma_{\alpha\beta}$ up to a conformal factor f. However, since the Polyakov action only depends on $g_{\alpha\beta}$ by the combination $\sqrt{-g}\,g^{\alpha\beta}$, f cancels out upon substituting it back in, recovering the Nambu–Goto action; note that this cancellation only holds in two dimensions.

- Like the Nambu–Goto action, the Polyakov action has Poincare invariance,

$$X^\mu \to \Lambda^\mu{}_\nu X^\nu + c^\mu.$$

Both actions also have reparametrization invariance, i.e. diffeomorphisms of the worldsheet. That is to say, the reparametrization $\sigma^\alpha \to \tilde{\sigma}^\alpha(\sigma)$ induces the transformations

$$X^\mu(\sigma) \to \tilde{X}^\mu(\tilde{\sigma}) = X^\mu(\sigma), \quad g_{\alpha\beta}(\sigma) \to \tilde{g}_{\alpha\beta}(\tilde{\sigma}) = \frac{\partial\sigma^\gamma}{\partial\tilde{\sigma}^\alpha}\frac{\partial\sigma^\delta}{\partial\tilde{\sigma}^\beta} g_{\gamma\delta}(\sigma)$$

which are to be regarded as gauge symmetries. The infinitesimal gauge transformation induced by $\sigma^\alpha \to \tilde{\sigma}_\alpha = \sigma^\alpha - \eta^\alpha(\sigma)$ is

$$\delta X^\mu(\sigma) = \eta^\alpha \partial_\alpha X^\mu, \quad \delta g_{\alpha\beta}(\sigma) = \nabla_\alpha \eta_\beta + \nabla_\beta \eta_\alpha$$

where the covariant derivative is defined by the Levi-Civita connection of the worldsheet metric.

- The Polyakov action further has Weyl invariance, special to a two-dimension worldsheet,

$$g_{\alpha\beta}(\sigma) \to \Omega^2(\sigma) g_{\alpha\beta}(\sigma).$$

Infinitesimally, writing $\Omega^2(\sigma) = e^{2\phi(\sigma)}$ we have

$$\delta g_{\alpha\beta}(\sigma) = 2\phi(\sigma) g_{\alpha\beta}(\sigma).$$

We have seen above why the Polyakov action is invariant under a Weyl transformation. Because the symmetry is local (i.e. Ω is a function on the worldsheet, not a constant) we interpret it as a gauge symmetry. As we'll see below, this choice ensures that $g_{\alpha\beta}$ doesn't introduce any new degrees of freedom. Weyl invariance is quite rare and strongly constrains interaction terms that can be added to the action; at the quantum level it also constrains $D = 26$.

- Like the Nambu–Goto action, we may fix a gauge to make concrete progress. The worldsheet metric has three independent components, so using reparametrization invariance we may fix

$$g_{\alpha\beta} = e^{2\phi} \eta_{\alpha\beta}$$

which is known as conformal gauge. We can further use Weyl transformations to set $g_{\alpha\beta} = \eta_{\alpha\beta}$, making the metric flat.

- Since the curvature of the metric isn't changed by reparametrizations, we should also be able to see that a Weyl transformation alone can make the metric flat. It can be shown that under a Weyl transformation $g'_{\alpha\beta} = e^{2\phi} g_{\alpha\beta}$ we have

$$\sqrt{g'} R' = \sqrt{g}(R - 2\nabla^2\phi)$$

which gives a differential equation for ϕ which may be used to set $R = 0$. Since the Riemann tensor has only one degree of freedom in two dimensions, this implies the metric is flat.

- Upon setting $g_{\alpha\beta} = \eta_{\alpha\beta}$ in the Polyakov action, we simply have

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

which gives the simple equation of motion

$$\partial_\alpha \partial^\alpha X^\mu = 0.$$

We've seen this equation of motion for several gauge choices in the Nambu–Goto action before.

- There are constraints due to the equation of motion for $g_{\alpha\beta}$. It is convenient to write them as

$$T_{\alpha\beta} = -\frac{2}{T_0}\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta g^{\alpha\beta}} = 0$$

  where T_0 is the string tension. The vanishing of the stress-energy tensor is due to reparametrization invariance, just like the vanishing of the Hamiltonian for the point particle.

- Setting $g_{\alpha\beta} = \eta_{\alpha\beta}$, we have

$$T_{\alpha\beta} = \partial_\alpha X \cdot \partial_\beta X - \frac{1}{2}\eta_{\alpha\beta}\eta^{\rho\sigma}\partial_\rho X \cdot \partial_\sigma X.$$

  The vanishing of the stress-energy tensor gives the constraints

$$T_{01} = T_{10} = \dot{X} \cdot X' = 0, \quad T_{00} = T_{11} = \frac{1}{2}(\dot{X}^2 + X'^2) = 0$$

  which are just what we found earlier in light cone gauge. In terms of components of $g_{\alpha\beta}$, they simply reiterate that the metric takes the required flat form; also note that Weyl invariance alone guarantees $\operatorname{tr} T = 0$ and hence $T_{00} = T_{11}$.

Next, we write down the mode expansion.

- For reference, we are taking the conventions

$$\ell^2 = 2\alpha' = \frac{1}{\pi T_0}.$$

  Later, we will set $\ell = 1$, so that $\alpha' = 1/2$.

- Ignoring the constraint $T_{\alpha\beta} = 0$ for now, for a closed string with $\sigma \in [0, \pi]$, decomposing into left-moving and right-moving solutions gives

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

  where we conventionally define

$$X_R^\mu(u) = \frac{x^\mu}{2} + \frac{\ell^2 p^\mu}{2}u + \frac{i\ell}{2}\sum_{n\neq 0}\frac{\alpha_n^\mu}{n}e^{-2inu}, \quad X_L^\mu(u) = \frac{x^\mu}{2} + \frac{\ell^2 p^\mu}{2}u + \frac{i\ell}{2}\sum_{n\neq 0}\frac{\tilde{\alpha}_n^\mu}{n}e^{-2inu}.$$

  Reality of X^μ implies x^μ and p^μ are real, and $(\alpha_n^\mu)^* = \alpha_{-n}^\mu$. The string length ℓ is related to the tension by $\ell^2 = 1/\pi T_0$, and later we will set it to one.

- For an open string with Neumann boundary conditions ($X' = 0$ at endpoints), the general solution is

$$X^\mu(\tau, \sigma) = x^\mu + \ell^2 p^\mu \tau + i\ell\sum_{n\neq 0}\frac{\alpha_n^\mu}{n}e^{-in\tau}\cos(n\sigma).$$

  That is, the left-moving and right-moving waves are forced to combine into standing waves, $\alpha_n^\mu = \tilde{\alpha}_n^\mu$. For now we put aside Dirichlet boundary conditions, returning to the subject later.

- With the definition of ℓ as above, the Noether charge for translational symmetry is simply

$$P^\mu = T_0 \partial_\tau X^\mu = p^\mu.$$

Next, we consider the Noether charge for Lorentz symmetry,

$$M^{\mu\nu} = \int J_\tau^{\mu\nu}(\tau,\sigma)\, d\sigma, \quad J_a^{\mu\nu}(\tau,\sigma) = T_0(X^\mu \partial_a X^\nu - X^\nu \partial_a X^\mu).$$

Evaluating this by using the above solutions, we find

$$M^{\mu\nu} = \begin{cases} \ell^{\mu\nu} + E^{\mu\nu} + \tilde{E}^{\mu\nu} & \text{closed} \\ \ell^{\mu\nu} + E^{\mu\nu} & \text{open} \end{cases}$$

where

$$\ell^{\mu\nu} = x^\mu p^\mu - x^\nu p^\mu, \quad E^{\mu\nu} = -i \sum_{n>0} \frac{1}{n}(\alpha_{-n}^\mu \alpha_n^\nu - \alpha_{-n}^\nu \alpha_n^\mu), \quad \tilde{E}^{\mu\nu} = -i \sum_{n>0} \frac{1}{n}(\tilde{\alpha}_{-n}^\mu \tilde{\alpha}_n^\nu - \tilde{\alpha}_{-n}^\nu \tilde{\alpha}_n^\mu).$$

- Next, we impose the constraint $T_{\alpha\beta} = 0$. This is easiest if we switch to light cone coordinates,

$$\sigma^\pm = \tau \pm \sigma, \quad \partial_\pm = \frac{1}{2}(\partial_\tau \pm \partial_\sigma), \quad \eta_{+-} = \eta_{-+} = -\frac{1}{2}, \quad \eta^{+-} = \eta^{-+} = -2.$$

By Weyl symmetry, our general solution above automatically satisfies $T_{00} = T_{11}$, which implies $T_{-+} = T_{+-} = 0$. As for the other components,

$$T_{++} = \partial_+ X \cdot \partial_+ X, \quad T_{--} = \partial_- X \cdot \partial_- X.$$

- By translational symmetry on the worldsheet, the stress-energy tensor is conserved for our general solutions above, so they obey

$$\partial_+ T_{-+} + \partial_- T_{++} = \partial_+ T_{--} + \partial_- T_{-+} = 0.$$

Combining with the previous result, we have

$$\partial_- T_{++} = \partial_+ T_{--} = 0.$$

Thinking of the worldsheet as the complex plane, we can think of T_{++} as a holomorphic function and T_{--} as anti-holomorphic.

- This result leads to an infinite number of conserved charges,

$$Q_f = \int d\sigma\, f(\sigma_+) T_{++}(\sigma_+)$$

for any function f, because $\partial_-(f T_{++}) = 0$, so

$$\frac{\partial Q_f}{\partial \tau} = \int d\sigma\, \partial_\tau(f T_{++}) = \int d\sigma\, \partial_\sigma(f T_{++}) = 0.$$

- Geometrically, the reason for these conserved quantities is that there is residual diffeomorphism invariance, namely conformal transformations whose effect on the metric can be cancelled by a Weyl rescaling. Such diffeomorphisms are generated by a vector field ξ satisfying

$$\partial^\alpha \xi^\beta + \partial^\beta \xi^\alpha = \Lambda \eta^{\alpha\beta}.$$

This doesn't violate our parameter counting earlier, as this remaining gauge freedom is of "measure zero" compared to the original freedom. However, it remains infinite-dimensional, which is special to two dimensions. These symmetries are generated by the Q_f with $f \sim \xi^+$.

At this point, we impose the constraint $T_{\alpha\beta} = 0$.

- Before continuing, it is useful to compute the Poisson brackets. For closed strings, starting with

$$[\dot{X}^\mu(\sigma), X^\nu(\sigma')] = \frac{1}{T_0}\delta(\sigma - \sigma')\eta^{\mu\nu}$$

with all other Poisson brackets zero, we easily find

$$[\alpha^\mu_m, \alpha^\nu_n] = [\tilde{\alpha}^\mu_m, \tilde{\alpha}^\nu_n] = im\delta_{m+n,0}\eta^{\mu\nu}, \quad [p^\mu, x^\nu] = \eta^{\mu\nu}$$

with all others zero. These also hold for $m, n = 0$, where we define

$$\alpha^\mu_0 = \begin{cases} \ell p^\mu & \text{open} \\ \ell p^\mu/2 & \text{closed} \end{cases}, \quad \tilde{\alpha}^\mu_0 = \alpha^\mu_0 \text{ for closed.}$$

We see the position and momentum of the string are canonically conjugate, and the Fourier modes α^μ_n for $n \neq 0$ are harmonic oscillator coordinates with conjugate variable α^μ_{-n}. The solution for open strings has been normalized so that it obeys the same set of Poisson brackets, without the extra $\tilde{\alpha}^\mu_n$.

- Another straightforward calculation shows that the Hamiltonian is

$$H = \frac{T_0}{2}\int_0^\pi (\dot{X}^2 + X'^2)\, d\sigma = \frac{1}{2}\sum_n \begin{cases} \alpha_{-n}\cdot\alpha_n & \text{open} \\ \alpha_{-n}\cdot\alpha_n + \tilde{\alpha}_{-n}\cdot\tilde{\alpha}_n & \text{closed} \end{cases}.$$

- The nontrivial content of the constraint $T_{\alpha\beta} = 0$ is $T_{++} = T_{--} = 0$. For closed strings, defining

$$L_m = \frac{T_0}{2}\int_0^\pi d\sigma\, e^{2im\sigma_-}T_{--}, \quad \tilde{L}_m = \frac{T_0}{2}\int_0^\pi d\sigma\, e^{2im\sigma_-}T_{++}$$

it is sufficient to show that the Fourier components L_m and $\tilde{L}_m$ all vanish. We have $T_{--} = \dot{X}_R^2$ and $T_{++} = \dot{X}_L^2$, so

$$L_m = \frac{1}{2}\sum_n \alpha_{m-n}\cdot\alpha_n, \quad \tilde{L}_m = \frac{1}{2}\sum_n \tilde{\alpha}_{m-n}\cdot\tilde{\alpha}_n.$$

- For open strings, we can get a similar expression if we formally extend the range of σ to $[0, 2\pi]$, defining $X_R(\sigma + \pi) = X_L(\sigma)$ and $X_L(\sigma + \pi) = X_R(\sigma)$. In this case, open string boundary conditions imply X_R is periodic with period 2π. The constraints imply that T_{++} vanishes on $[-\pi, \pi]$, which is equivalent to the vanishing of the Fourier components

$$L_m = T_0 \int_0^\pi e^{im\sigma}T_{++} + e^{-im\sigma}T_{--}\, d\sigma = \frac{1}{2}\sum_n \alpha_{m-n}\cdot\alpha_n.$$

The constraint for T_{--} is redundant.

- Note in particular that
$$H = \begin{cases} L_0 & \text{open} \\ L_0 + \tilde{L}_0 & \text{closed} \end{cases}.$$
The constraint $L_0 = \tilde{L}_0 = 0$ and definition $M^2 = -p_\mu p^\mu$ gives the mass shell conditions
$$M^2 = \frac{1}{\alpha'} \sum_{n>0} \begin{cases} \alpha_{-n} \cdot \alpha_n & \text{open} \\ 2(\alpha_{-n} \cdot \alpha_n + \tilde{\alpha}_{-n} \cdot \tilde{\alpha}_n) & \text{closed} \end{cases}$$
where the two terms in the closed case give equal contributions. At the quantum level, these results will be modified due to normal ordering effects.

- By another straightforward calculation, we find that
$$[L_m, \alpha_n^\mu] = -in\alpha_{m+n}^\mu$$
along with the Witt algebra
$$[L_m, L_n] = i(m-n)L_{m+n}.$$

- This appearance of this algebra has a simple interpretation. A complete basis for diffeomorphisms of the circle is
$$D_n = ie^{in\theta} \frac{d}{d\theta}$$
and these satisfy the Witt algebra, so it is the algebra of infinitesimal diffeomorphisms of the circle. In fact, the transformations generated by the L_n and $\tilde{L}_n$ correspond to the worldsheet diffeomorphisms generated by $e^{2in\sigma^\pm}\partial_\pm$, where the $\sigma^\pm$ behave like angular variables because solutions to the equations of motion are periodic in them.

3.2 Old Covariant Quantization

We now continue with the quantization of the string. There are several possible approaches.

- In light cone quantization, we fix all gauge symmetry by going to light cone gauge, and solve all of the constraints of the system to determine the space of physically distinct classical solutions. This is the analogue of Coulomb gauge in QED, but loses manifest Lorentz invariance.

- In old covariant quantization, one quantizes the string in conformal gauge, then imposes the constraints $T_{++} = T_{--} = 0$ at the quantum level on the operators. This is the analogue of Gupta–Bleuler quantization in Lorenz gauge in QED.

- In covariant BRST quantization, one uses the path integral instead. One must be careful to account for the diffeomorphism and Weyl gauge symmetries, which leads to Faddeev–Popov ghosts and BRST cohomology, as we saw for Yang–Mills theory.

In this section, we use old covariant quantization, focusing on the closed string.

- As usual in canonical quantization, we replace Poisson brackets with commutators. The relations in the previous section can be converted by multiplying the right-hand sides by $-i$, so
$$[\hat{p}^\mu, \hat{x}^\nu] = -i\eta^{\mu\nu}, \quad [\alpha_m^\mu, \alpha_n^\nu] = m\delta_{m+n}\eta^{\mu\nu}, \quad [\tilde{\alpha}_m^\mu, \tilde{\alpha}_n^\nu] = m\delta_{m+n}\eta^{\mu\nu}$$
where $\hat{x}$ and $\hat{p}$ are Hermitian, and $(\alpha_n^\mu)^\dagger = \alpha_{-n}^\mu$ and $(\tilde{\alpha}_n^\mu)^\dagger = \tilde{\alpha}_{-n}^\mu$.

- The α_n^μ for $n>0$ can be interpreted as annihilation operators for a harmonic oscillator, with α_{-n}^μ the corresponding creation operators. Ignoring $\hat{x}$ and $\hat{p}$ for now, we define the vacuum state $|0\rangle$ to be annihilated by all the α_n^μ for $n>0$, and build up the Fock space by acting with α_{-n}^μ, all as usual. However, unlike in quantum field theory, the vacuum state should be interpreted as the lowest energy state of a *single* string, not the absence of any strings.

- We define the right-moving "number operator"

$$N=\sum_{k>0}\alpha_{-k}\cdot\alpha_k$$

with a similar definition for $\tilde{N}$, and we say a state is at level n if its eigenvalue of N is n. Then

$$N\left(\prod_i \alpha_{n_i}^{\mu_i}|0\rangle\right)=\sum_i n_i.$$

- Now we need to account for the zero mode associated with $\hat{x}$ and $\hat{p}$. This should be interpreted as generating the Hilbert space for a free particle. We may define the states

$$\hat{p}^\mu|p\rangle=p^\mu|p\rangle,\quad \langle p|p'\rangle=\delta(p-p')$$

and the resulting Hilbert space is just $L^2(\mathbb{R}^{1,D-1})$. The full Hilbert space is the tensor product of this with the Fock space associated with the harmonic oscillators, and we write the ground state as $|0,p\rangle$.

- The Poincare charges are promoted to the operators

$$P^\mu=\hat{p}^\mu,\quad M^{\mu\nu}=\hat{x}^\mu\hat{p}^\nu-\hat{x}^\nu\hat{p}^\mu-i\sum_{n>0}\frac{\alpha_{-n}^\mu\alpha_n^\nu-\alpha_{-n}^\nu\alpha_n^\mu}{n}-i\sum_{n>0}\frac{\tilde{\alpha}_{-n}^\mu\tilde{\alpha}_n^\nu-\tilde{\alpha}_{-n}^\nu\tilde{\alpha}_n^\mu}{n}$$

which obey the expected Poincare algebra. This is the benefit of working covariantly. Note that demanding $M^{\mu\nu}$ be antisymmetric eliminates the ordering ambiguity.

- The drawback is the need to impose the constraints $L_n=0$. For $n\neq 0$, we have

$$L_n=\frac{1}{2}\sum_k\alpha_{n-k}\cdot\alpha_k,\quad L_{-n}=L_n^\dagger$$

unambiguously, but for $n=0$ there is an ordering ambiguity. The naive ordering above is unacceptable because it yields infinity when acting on the vacuum, so a better ordering is the normal ordering

$$L_0=\frac{1}{2}\alpha_0\cdot\alpha_0+\sum_{k>0}\alpha_{-k}\cdot\alpha_k.$$

There is a similar story for the $\tilde{L}_n$.

- Because of this correction, the Witt algebra becomes the Virasoro algebra,

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

This is a central extension of the Witt algebra, and the new term is sometimes called an "anomaly". Note that the algebra of $\{L_{-1},L_0,L_1\}$ is unmodified; these are the generators of $\mathfrak{sl}(2,\mathbb{R})$, the conformal transformations which also exist in $d>2$.

- Further results that will be useful below are

$$[L_m, \alpha_n^\mu] = -n\alpha_{m+n}^\mu, \quad [L_m, x^\mu] = \begin{cases} -i\ell\alpha_m^\mu & \text{open} \\ (-i\ell/2)\alpha_m^\mu & \text{closed} \end{cases}$$

which may be combined to show that for closed strings,

$$[L_m, X^\mu] = -ie^{2im\sigma_-}\partial_- X^\mu$$

which confirms the interpretation of the L_m given above.

- A problem related to the constraints is the presence of ghosts, i.e. negative norm states. These are generated by the α_n^0 because of the indefinite sign of the metric, just like we saw in QED. There are also zero norm states, such as $(\alpha_n^0 + \alpha_n^1)|0, p\rangle$, which arise generically in gauge theories. These are the states generated by gauge transformations, which must necessarily have zero norm because gauge transformations cannot affect probabilities. As in QED, the hope is that fixing a gauge, which removes the zero norm states, will simultaneously decouple the negative norm states, so that they cannot be produced in physical processes.

Note. A very heuristic way to understand the appearance of the $1/12$ is that the difference between our L_0 and the original ordering is an additive factor of $D\sum_{k>0} k$, which by zeta function regularization is $-D/12$.

At a slightly more respectable level, we must add a cosmological constant term to preserve conformal invariance at the quantum level; in a suitable regularization, this soaks up the divergent part of the sum $D\sum_{k>0} k$ but leaves behind the finite part $-D/12$.

Note. There is an annoying sign issue here: usually in canonical quantization we multiply the result of a Poisson bracket by i to get a commutator, not $-i$. The reason is that canonical momenta are naturally covectors, so the contravariant momenta p^μ pick up a relative minus sign due to our unusual $(-+++)$ metric convention, causing the sign flip.

Next, we impose the classical constraints $L_n = \tilde{L}_n = 0$.

- As for QED, imposing that $L_n = \tilde{L}_n = 0$ as an operator equation is too strong. For example, we would necessarily have $[L_m, L_n] = 0$, but then the Virasoro algebra cannot be satisfied.

- Instead, as in the Gupta–Bleuler quantization of QED, we only demand that the L_n and $\tilde{L}_n$ have vanishing matrix elements within the subspace of physical states. Since $L_n^\dagger = L_{-n}$, it is sufficient to require

$$L_n|\text{phys}\rangle = \tilde{L}_n|\text{phys}\rangle = 0, \quad n > 0.$$

For $n = 0$, we also add a normal ordering constant

$$(L_0 - a)|\text{phys}\rangle = (\tilde{L}_0 - a)|\text{phys}\rangle = 0$$

accounting for the fact that we don't know the proper ordering to define L_0 and $\tilde{L}_0$ a priori. Recall that we need to remove the timelike oscillators; since we have $L_n \sim p \cdot \alpha_n + \cdots$ where p is timelike, this procedure stands a chance of working.

- The introduction of a modifies the mass shell constraint for closed strings to

$$M^2 = \frac{4}{\alpha'}\left(-a + \sum_{k>0} \alpha_{-k}\cdot\alpha_k\right) = \frac{4}{\alpha'}\left(-a + \sum_{k>0} \tilde{\alpha}_{-k}\cdot\tilde{\alpha}_k\right)$$

  which constrains the allowed values of p for a state with given oscillator excitations. It also yields the "level matching" constraint $N = \tilde{N}$.

- Note that $[L_n, M^{\mu\nu}] = 0$. This implies the physical state conditions are invariant under Lorentz transformations, so the physical states will form Lorentz multiplets.

- For simplicity, we focus on the open string case. The reasoning for the closed string is very similar, with essentially two copies of the theory (left-moving and right-moving) plus the level matching constraint. For open strings, we instead have

$$M^2 = \frac{1}{\alpha'}\left(-a + \sum_{k>0} \alpha_{-k}\cdot\alpha_k\right).$$

  At level zero, the states $|0, k\rangle$ hence have mass squared $M^2 = -a/\alpha'$.

- Also note that the string Hamiltonian is modified to

$$H = \begin{cases} L_0 - a & \text{open} \\ L_0 + \tilde{L}_0 - 2a & \text{closed} \end{cases}.$$

- Now consider the states at level one, given by

$$\zeta\cdot\alpha_{-1}|0, k\rangle$$

  for a polarization vector $\zeta^\mu(k)$. These states have $M^2 = (1-a)/\alpha'$, and the L_1 condition implies that $\zeta\cdot k = 0$, giving $D-1$ allowed polarizations, where the norm of the state is $\zeta\cdot\zeta$.

- If $a > 1$, then these states are tachyonic, so it is possible to rotate k to have no time component. Then one of the physical states has a timelike polarization and negative norm, so we require

$$a \leq 1.$$

  When $a < 1$ the mass is positive, and we get $D-1$ spacelike polarizations, to be interpreted as a massive vector particle.

- In the boundary case $a = 1$ the particle is massless; accordingly one of the physical polarizations is $\zeta^\mu = k^\mu$ with zero norm. As in the Gupta–Bleuler quantization of QED, this state decouples from the S-matrix, as we will see below.

Next, we define spurious states.

- In general, we define a state $|\psi\rangle$ to be spurious if

$$(L_0 - a)|\psi\rangle = 0, \quad \langle\phi|\psi\rangle = 0$$

for all physical states $|\phi\rangle$. All spurious states can be written in the form

$$|\psi\rangle = \sum_{n>0} L_{-n}|\chi_n\rangle$$

where the $|\chi_n\rangle$ satisfy

$$(L_0 - a + n)|\chi_n\rangle = 0.$$

In fact, since all L_{-n} can be constructed as commutators of L_{-1} and L_{-2}, the sum can be stopped at $n = 2$. Hence the general spurious state is

$$|\psi\rangle = L_{-1}|\chi_1\rangle + L_{-2}|\chi_2\rangle.$$

- A state can be both spurious and physical, in which case they must be null. For example, consider states of the form

$$|\psi\rangle = L_{-1}|\tilde{\chi}\rangle, \quad L_m|\tilde{\chi}\rangle = 0 \text{ for } m > 0, \quad (L_0 - a + 1)|\tilde{\chi}\rangle = 0.$$

The physical state conditions are automatically satisfied, except for the L_1 condition, where

$$L_1|\psi\rangle = L_1 L_{-1}|\tilde{\chi}\rangle = 2L_0|\tilde{\chi}\rangle.$$

This only vanishes for $a = 1$. We interpret spurious physical states as gauge equivalent to zero. For example, when $a = 1$ we have seen there is an extra massless state at level one; this is rendered unphysical since we may take $|\tilde{\chi}\rangle = |0, k\rangle$. Hence at level one we have a massless vector particle, corresponding to a gauge field.

- Now fixing $a = 1$, consider spurious states with the structure

$$|\psi\rangle = (L_{-2} + \gamma L_{-1}^2)|\tilde{\chi}\rangle, \quad L_m|\tilde{\chi}\rangle \text{ for } m > 0, \quad (L_0 + 1)|\tilde{\chi}\rangle = 0.$$

The latter condition ensures that $(L_0 - 1)|\psi\rangle = 0$. The physical state conditions $L_m|\psi\rangle = 0$ for $m > 2$ are always satisfied, so we only need impose $L_1|\psi\rangle = L_2|\psi\rangle = 0$. It turns out these are satisfied when

$$\gamma = \frac{3}{2}, \quad D = 26$$

so that there are many more spurious physical states in $D = 26$.

- Furthermore, it is possible to construct physical states of negative norm in $D > 26$. In fact, one can show the spectrum is ghost-free provided that $a = 1$ and $D = 26$, or $a \leq 1$ and $D \leq 25$. In the former case, there are many more zero-norm states, and the physical spectrum corresponds to 24 sets of α oscillators, while in the latter case it corresponds to $D - 1$ oscillators.

- Physically, we say the string has only transverse excitations in $D = 26$ but also longitudinal oscillations in lower dimension. Since the gauge symmetry is evidently much larger in $D = 26$, we will focus on this case. This formally contains the cases with $D < 26$ by restricting the momenta.

3.3 Computing Spectra

Now we'll use the results above to investigate the low-lying spectra of open and closed strings.

Example. The physical Hilbert space at level two for the open string with $a = 1$. We parametrize the states as

$$|g, \epsilon, p\rangle = (g_{\mu\nu}\alpha_{-1}^{\mu}\alpha_{-1}^{\nu} + \epsilon_{\mu}\alpha_{-2}^{\mu})|0, p\rangle$$

where $g_{\mu\nu}$ may be taken symmetric, giving $D(D+1)/2 + D$ candidate physical states. Note that

$$L_0|0, p\rangle = \frac{1}{2}(\alpha_0 \cdot \alpha_0)|0, p\rangle = \frac{1}{2}p^2|0, p\rangle$$

where we have set $\ell = 1$. Now consider the physical state condition $(L_0 - 1)|g, \epsilon, p\rangle$. By commuting the L_0 to the right,

$$\left(\frac{1}{2}p^2 - 1 + 2\right)|g, \epsilon, p\rangle = 0$$

which shows that $m^2 = -p^2 = 2$, so the states have positive mass, and

$$L_0|0, p\rangle = -|0, p\rangle.$$

The physical state conditions $L_k|g, \epsilon, p\rangle = 0$ are trivial for $k > 2$. For $k = 1$ we find

$$\left(g_{\mu\nu}\alpha_0^{\mu}\alpha_{-1}^{\nu} + g_{\mu\nu}\alpha_{-1}^{\mu}\alpha_0^{\nu} + 2\epsilon_{\mu}\alpha_{-1}^{\mu}\right)|0, p\rangle = 0.$$

Since α_{-1} is a raising operator, it must act on the zero state, giving the constraint

$$g_{\mu\nu}p^{\nu} + \epsilon_{\mu} = 0.$$

Next, for $k = 2$ we have

$$\left(g_{\mu\nu}\alpha_1^{\mu}\alpha_{-1}^{\nu} + g_{\mu\nu}\alpha_{-1}^{\mu}\alpha_1^{\nu} + 2\epsilon_{\mu}\alpha_0^{\mu}\right)|0, p\rangle$$

and since $[\alpha_1^{\mu}, \alpha_{-1}^{\nu}] = \eta^{\mu\nu}$ this gives

$$g_{\mu\nu}\eta^{\mu\nu} + 2\epsilon_{\mu}p^{\mu} = 0.$$

These are the full physical state conditions, which give a total of $D + 1$ constraints on $g_{\mu\nu}$ and ϵ_{μ}.

Next, the most general spurious state at level two is

$$|\tilde{\epsilon}, \tilde{\gamma}, p\rangle = (L_{-1}\,\tilde{\epsilon} \cdot \alpha_{-1} + \tilde{\gamma}L_{-2})\,|0, p\rangle$$

by the reasoning above. The simplest way to impose the physical state condition is simply to expand the expression above in terms of oscillator modes and use our earlier result. This gives

$$|\tilde{\epsilon}, \tilde{\gamma}, p\rangle = \left[\frac{1}{2}(\tilde{\gamma}\eta_{\mu\nu} + \tilde{\epsilon}_{\mu}p_{\nu} + \tilde{\epsilon}_{\nu}p_{\mu})\alpha_{-1}^{\mu}\alpha_{-1}^{\nu} + (\tilde{\epsilon} + \tilde{\gamma}p) \cdot \alpha{-2}\right]|0, p\rangle.$$

The two physical state conditions are

$$3\tilde{\gamma} + \tilde{\epsilon} \cdot p = 0, \quad 3\tilde{\epsilon} \cdot p + \frac{D}{2}\tilde{\gamma} - 4\tilde{\gamma} = 0$$

for $k = 1$ and $k = 2$ respectively. For $D = 26$, they are redundant, so there are D spurious physical states; for $D < 26$ we have $\tilde{\gamma} = \tilde{\epsilon} \cdot p = 0$, giving $D - 1$ spurious physical states. Therefore, accounting for the constraints and the spurious physical states, we have $D(D-1)/2 - 1$ states in $D = 26$, and $D(D-1)/2$ states otherwise. These are the number of degrees of freedom in a symmetric $SO(D-1)$ tensor, which is traceless in $D = 26$. In $D = 26$, this is to be interpreted as a massive spin 2 particle.

The analysis for the closed string is more complicated, so we'll start with lightcone quantization.

- In lightcone quantization, the counting is more straightforward, as there are precisely $D-2$ physical oscillator modes and no spurious states to worry about. For example, for the open string we manifestly have $D-2$ physical states at level one. This implies the states must be massless if we wish to preserve Lorentz invariance, so $a=1$. Furthermore, it is possible to compute a by regularizing the zero-point energy, giving $a=(D-2)/24$. This forces $D=26$ to preserve Lorentz invariance.
- A somewhat more rigorous way to show this is to compute $[\mathcal{M}^{i-}, \mathcal{M}^{j-}]$, which must vanish for the Lorentz algebra to be satisfied; this only holds if $a=1$ and $D=26$. We hence get the same conditions as for covariant quantization, but for a different reason.
- Now consider the closed string. Fixing $a=1$ and setting $\ell=1$ again, we have:
  - At level zero, the states $|0,k\rangle$ have mass squared $M^2=-8$.
  - At level one, there are $(D-2)^2$ massless states, which corresponds to the rank two tensor of $SO(D-2)$, the homogeneous part of the little group for massless particles.
  - At level two, the counting gets a bit more complicated, so consider only the left-moving sector. The states can be built from two α_{-1}'s or from one α_{-2}, giving

    $$\frac{1}{2}(D-2)(D-1)+(D-2)=\frac{1}{2}D(D-1)-1$$

    states of mass squared $M^2=8$. This is precisely the traceless symmetric tensor of $SO(D-1)$. The full state space at level two fits in the square of this representation.

  It wouldn't be too hard to continue, but we would require Young tableaux. It also isn't too interesting, because typically anything beyond level one will be far too heavy to observe.
- Next, we recover the level one result in covariant quantization. The states are

  $$|\Omega,p\rangle=\Omega_{\mu\nu}\alpha^{\mu}_{-1}\alpha^{\nu}_{-1}|0,0,p\rangle$$

  where $\Omega_{\mu\nu}$ is a tensor in D-dimensional spacetime.
- The L_0 physical state condition gives $p^2=0$, while the L_1 condition gives

  $$p^{\mu}\Omega_{\mu\nu}=p^{\nu}\Omega_{\mu\nu}=0.$$

  Furthermore, spurious physical states have the form

  $$p_{\mu}\xi_{\nu}\alpha^{\mu}_{-1}\alpha^{\nu}_{-1}|0,0,p\rangle, \quad p_{\nu}\xi_{\mu}\alpha^{\mu}_{-1}\alpha^{\nu}_{-1}|0,0,p\rangle$$

  where the physical state condition requires $p\cdot\xi=0$.
- Now we count the total number of states. We start with D^2, and the L_1 conditions remove $D+(D-1)$, as they have one redundancy. Then the spurious states remove a further $(D-1)+(D-2)$, where there is another redundancy between the equations in the case $p\propto\xi$. This leaves a total of $D^2-4D+4=(D-2)^2$, just as we found in lightcone quantization.

- We can get more insight by decomposing the tensor $\Omega_{\mu\nu}$. Let $G_{\mu\nu}$ be the traceless symmetric part. Then the constraints above are

$$p^\mu G_{\mu\nu} = 0, \quad G_{\mu\nu} \sim G_{\mu\nu} + p_\mu \xi_\nu + p_\nu \xi_\mu, \quad p \cdot \xi = 0.$$

  We may interpret $G_{\mu\nu}$ as a spin 2 graviton field; it has precisely the same gauge symmetries as the linearized metric.

- Next, let $B_{\mu\nu}$ be the antisymmetric part. The constraints are

$$p^\mu B_{\mu\nu} = 0, \quad B_{\mu\nu} \sim B_{\mu\nu} + p_\mu \xi_\nu - p_\nu \xi_\mu, \quad p \cdot \xi = 0.$$

  However, note that the gauge redundancy itself has a gauge redundancy: we get the same spurious state if p is added to ξ, so

$$\xi \sim \xi + p.$$

  Hence the transverse constraint removes $D-1$ degrees of freedom and the gauge redundancy removes $D-2$, leaving the right number of degrees of freedom for a $(D-2)$-dimensional antisymmetric tensor.

- The field creating these states is a spin 1 Kalb–Ramond field; it has the same gauge redundancies as those we saw earlier. A Kalb–Ramond field can naturally couple to strings in the same way that a one-form gauge field couples to particles, and we will see examples later where strings carry Kalb–Ramond charge.

- There is one remaining spin 0 degree of freedom, which is heuristically the trace. However, the naive guess $\Omega_{\mu\nu} \sim \eta_{\mu\nu}$ doesn't work, because the physical state condition $p^\mu \eta_{\mu\nu} = 0$ can't be satisfied. It instead turns out that we may write the candidate dilaton states in terms of an arbitrary polarization vector ζ in a somewhat complicated way, and the spurious states ensure all values of ζ are gauge equivalent.

- Physically, the spin 0 degree of freedom is the dilaton ϕ. It turns out to be related to the value of the string coupling g by $g \sim e^\phi$. This string coupling is the only dimensionless parameter of string theory, but its relation with the dilaton implies it may be determined dynamically.

- We could continue to higher levels, but the massless particles at level one are far more interesting, because the higher particles are presumably too heavy to observe. As we'll see, the graviton, Kalb–Ramond field, and dilaton are common to all string theories.

Note. We have used the word "spin" above casually. Properly speaking, the spin of a Lorentz representation in $D > 4$ is the highest helicity of any one of the components. Under this definition, antisymmetric tensor fields (differential forms) have spin 1, while symmetric rank n tensor fields have spin n. These correspond to the maximum possible helicities of the particles they generate.

The key property of the above definition is that it is preserved upon compactifying some dimensions to leave $D = 4$, as in Kaluza–Klein theory. For example, the states created by the Kalb–Ramond field will have helicity ± 1 or 0 in the 4D theory, depending on how the helicity was oriented in the original D dimensions. Similarly, the graviton $G_{\mu\nu}$ contains "our" graviton $g_{\mu\nu}$ in the low-energy 4D theory, along with some particles of helicity 0 or ± 1.

Note. So far, all the strings we have encountered have been oriented, as increasing σ defines a preferred direction. An unoriented string can be constructed by quotienting out by orientation reversal, or equivalently by only working with symmetric superpositions of oriented strings. Orientation reversal swaps α and $\tilde{\alpha}$ and hence eliminates the Kalb–Ramond field. We will not use it much, but it plays a role in constructing the five superstring theories.

Finally, we give a brief previous of superstring theory.

- To pass to superstring theory, we add fermionic modes on the worldsheet. We find that the critical dimension becomes $D = 10$, there is no tachyon, and the massless bosonic fields $G_{\mu\nu}$, $B_{\mu\nu}$, and Φ all appear.
- In type II string theory, there are both left-moving and right-moving worldsheet fermions on a closed string. The resulting spacetime theory in $D = 10$ has $\mathcal{N} = 2$ supersymmetry. There are additional massless bosonic excitations called Ramond-Ramond fields.
  - In type IIA string theory, they are a 1-form C_μ and a 3-form $C_{\mu\nu\rho}$.
  - In type IIB string theory, they are a scalar C, a 2-form $C_{\mu\nu}$, and a 4-form $C_{\mu\nu\rho\sigma}$ with a self-dual field strength.

  All of these Ramond-Ramond fields are to be interpreted as gauge fields.
- In heterotic string theory, there are only right-moving worldsheet fermions on a closed string. There is $\mathcal{N} = 1$ spacetime supersymmetry. Instead of Ramond-Ramond fields, there is a non-Abelian gauge field whose gauge group is either $SO(32)$ or $E_8 \times E_8$.
- It turns out that theories of open strings necessarily contain closed strings, as an open string can join into a closed string. In type I string theory, there are both types. In type II string theory, there are also both types, but for heterotic string theory there are only closed strings.
- It also turns out that string theories can contain Dp-branes, dynamical objects with p spatial dimensions where the endpoints of open strings can attach. In fact, type IIA string theory has stable Dp-branes with p even, and type IIB string theory has stable Dp-branes with p odd.
- Note that strings themselves are D1-branes, while particles are D0-branes. Instantons also exist in string theory, and are sometimes called D(-1)-branes.

Note. The interpretation of the tachyon in bosonic string theory. In field theory, tachyons arise as excitations of a quantum field if we expand about a field value with a negative mass squared; this indicates we are expanding about a maximum of the potential, so the theory is unstable.

In open bosonic string theory, we can think of the string end points as attached to a space-filling D25-brane; the tachyon indicates an instability of this brane. String field theory techniques have been used to show that there is indeed a minimum of the potential. Alone the journey to this minimum, the D25-brane decays into closed strings, and only closed string excitations remain at the minimum. The theory about this minimum is called vacuum string field theory, and is not well-understood. It has also been shown that Dp-branes with $p < 25$ can be thought of as coherent states of the open string tachyon.

The closed bosonic string tachyon is even less well-understood. Physically, tachyons don't appear in the superstring theories because the D-branes carry charge and are hence stable against decay. However, refinements of these theories meant to describe the real world sometimes contain tachyons.

Example. The open string with a Dp-brane. For simplicity, we take the Dp-brane to be a hyperplane. The boundary conditions are

$$\partial_\sigma X^a = 0, \quad X^I = c^I, \quad a = 0, \ldots, p, \quad I = p+1, \ldots, D-1.$$

This breaks the $SO(1, D-1)$ Lorentz group to $SO(1, p) \times SO(D-p-1)$. We recall that Neumann boundary conditions ensure $\alpha_n^\mu = \tilde{\alpha}_n^\mu$. In this case, we only have $\alpha_n^a = \tilde{\alpha}_n^a$, while for the dimensions with Dirichlet boundary conditions,

$$x^I = c^I, \quad p^I = 0, \quad \alpha_n^I = -\tilde{\alpha}_n^I.$$

As before, the right-moving and left-moving modes are not independent, and the spectrum computation goes through mostly as before, with the same conditions $D = 26$ and $a = 1$. The main difference is that the zero mode x^μ must lie on the D-brane. That is, for low-lying excitations the strings are confined to be near the brane.

At level one, we can split the excitations into those longitudinal and transverse to the brane,

$$\alpha_{-1}^a |0, p\rangle, \quad \alpha_{-1}^I |0, p\rangle$$

respectively. The longitudinal states transform as a vector of the $SO(1, p)$ Lorentz group of the brane and hence correspond to a spin 1 particle, i.e. a gauge field A_a restricted to the brane. The transverse states transform as scalars under $SO(1, p)$ and hence can be thought of as scalar fields ϕ^I living on the brane. In fact, it turns out that the brane can be thought of as a nonperturbative composite state of strings, and these transverse states correspond to fluctuations of the brane. The transverse states transform as a vector under the $SO(D-p-1)$ group, which is a global internal symmetry of a field theory living on the brane.

Note. Presumably, branes would be described by the Dirac action, a generalization of the Nambu–Goto action equal to their volume. In particular, the transverse components may be identified with the fields ϕ^I above associated with transverse excitations of the open string. However, quantizing the brane is more difficult than quantizing the string. We do not have Weyl invariance to work with. Furthermore, hypersurfaces are “more flexible” than strings, with many very different configurations having the same volume; this results in a continuous spectrum of states. This could possibly be interpreted as describing multi-particle states in the full theory.

4 Conformal Field Theory

4.1 Conformal Transformations

Before beginning, we need to clear up a persistent confusion over what a conformal transformation precisely is. For simplicity, we'll consider a scalar field theory.

- Given a spacetime manifold M, consider a diffeomorphism $f\colon M \to M$. Fixing a single set of coordinates, f maps the point with coordinates x to the point with coordinates x', which we write as $x \to x'$.

- As covered in the notes on General Relativity, we can interpret f either actively or passively. In the active picture, if $f(p) = q$, then we imagine the point p being physically moved to q. All other fields are transformed by applying a pushforward or inverse pullback via f, so that

$$\phi(x) \to \phi'(x') = \phi(x), \quad g_{\mu\nu}(x) \to g'_{\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}(x).$$

  In the passive picture, we interpret each point p as staying in the same place, but change the coordinate description of that point from x to x'. In these new coordinates the fields are

$$\phi'(x') = \phi(x), \quad g'_{\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}(x).$$

  We will prefer to fix one coordinate system throughout and use the active interpretation.

- A Weyl transformation is an active rescaling of the fields of the form

$$\phi(x) \to \tilde{\phi}(x) = \Omega^{-\Delta}(x)\phi(x), \quad g_{\mu\nu}(x) \to \tilde{g}_{\mu\nu}(x) = \Omega^2(x) g_{\mu\nu}(x).$$

  For completeness, we also have

$$\partial_\mu \phi(x) \to \partial_\mu \tilde{\phi}(x) = \partial_\mu(\Omega^{-\Delta}(x)\phi(x)).$$

  The quantity Δ is the scaling/conformal dimension of the field, and is simply equal to its mass dimension. (However, this will be corrected in the quantum theory.)

- A conformal transformation is a special case of a diffeomorphism, where the net effect is to change the metric by a scale factor,

$$\phi'(x') = \phi(x), \quad g'_{\mu\nu}(x') = \Omega^{-2}(x) g_{\mu\nu}(x)$$

- A classical theory has a symmetry if its action remains the same after an active transformation, and *all* theories have diffeomorphism invariance, as it amounts to saying that physics is independent of the choice of coordinate system. Hence all theories trivially have conformal invariance under the definition above.

- When we speak of the conformal invariance of a theory, we always mean the composition of a conformal transformation and the Weyl transformation that cancels the rescaling of the metric,

$$\phi(x) \to \tilde{\phi}'(x') = \Omega^{-\Delta}(x)\phi(x), \quad g_{\mu\nu}(x) \to \tilde{g}'_{\mu\nu}(x') = g_{\mu\nu}(x)$$

  and

$$\partial_\mu \phi(x) \to \partial'_\mu \tilde{\phi}'(x') = \frac{\partial x^\nu}{\partial x'^\mu}\partial_\nu(\Omega^{-\Delta}(x)\phi(x)).$$

  From this point forward, we will call this transformation a conformal transformation.

- Thus, (nontrivial) conformal invariance is a special case of Weyl invariance, i.e. the case where the Weyl scaling factor Ω can be derived from a diffeomorphism. Sometimes one says that conformal invariance is only a "global" Weyl invariance. Weyl invariance implies conformal invariance, but not vice versa, but most commonly encountered theories have both or neither. (Note that general relativity texts conventionally call Weyl transformations conformal transformations.)

- Another way of making this distinction is to say that the trivially true conformal invariance uses a "dynamical" metric, while the nontrivial conformal invariance uses a fixed "background" metric. Unfortunately, both notions come up in string theory; we will be careful to specify which is being used.

- We will usually work in flat spacetime in Cartesian coordinates, where the distinction between conformal and Weyl is important. Weyl transformations can clearly change the curvature, but (nontrivial) conformal transformations on flat spacetime can't, as the metric is completely unchanged, $\overline{g}'_{\mu\nu}(x') = g_{\mu\nu}(x) = \eta_{\mu\nu}$. Intuitively, if we have a drawing on a plane, a conformal transformation picks it up off the plane, distorts it while preserving angles, then pastes it back on; the plane is a fixed background structure.

- A useful special case of a conformal transformation is a dilation transformation,
$$x \to x' = \Omega x, \quad \phi(x) \to \tilde{\phi}'(x') = \Omega^{-\Delta}\phi(x).$$
In this case Δ is called the scaling dimension of the field. Theories symmetric under dilation transformations are called scale invariant.

- Explicitly, conformal invariance can be checked as follows. Let
$$S = \int d^4x\, \mathcal{L}[g_{\mu\nu}(x), \phi(x), \partial_\mu\phi(x)], \quad S' = \int d^4x\, \mathcal{L}[\tilde{g}'_{\mu\nu}(x), \tilde{\phi}'(x), \partial'_\mu\tilde{\phi}'(x)].$$
Then the theory is conformally invariant if $S = S'$. Since we know the primed fields in terms of the primed variables, the simplest way to evaluate S is to first rename x to x', then transform back from x' to x,
$$\begin{aligned} S' &= \int d^4x'\, \mathcal{L}[\tilde{g}'_{\mu\nu}(x'), \tilde{\phi}'(x'), \partial'_\mu\tilde{\phi}'(x')] \\ &= \int d^4x\, |\Omega(x)|^4 \mathcal{L}\left[g_{\mu\nu}(x), \Omega^{-\Delta}(x)\phi(x), \frac{\partial x^\nu}{\partial x'^\mu}\partial_\nu(\Omega^{-\Delta}(x)\phi(x))\right] \end{aligned}$$
However, directly checking conformal invariance from the action is quite painful, even on flat spacetime. In this case, one must compute the transformed Riemann tensor for a general Weyl transformation and set it to zero, to get the required constraint on Ω.

- As such, it is conventional to only check scale invariance and then tacitly pretend we have checked conformal invariance, which will usually be correct for a flat spacetime background. Checking scale invariance is easy, because a theory is classically scale invariant if and only if its action contains no dimensionful couplings. This also shows that the scaling dimension Δ is simply the ordinary "engineering" dimension.

Now we provide some motivation for studying conformal field theory.

- As we've seen, a classical theory has scale/conformal/Weyl invariance if its action is invariant under appropriate transformations. A quantum theory has these symmetries if its partition function is invariant under the same transformations; as a result the classical consequences of these symmetries hold as Ward identities. For example, in classically conformally invariant theories the trace of the stress-energy tensor vanishes; at the quantum level the expectation value of its trace vanishes.

- One can show that source-free electromagnetism, and more generally Yang–Mills, are both Weyl invariant. One manifestation of this is the use of conformal mapping to solve two-dimensional electrostatics problems. **(how about sources?)**

- Another example is a free massless scalar field. This theory is Weyl invariant in $d = 2$. However, for general dimension, one must include another term proportional to $R\phi^2$, as stated in the notes on General Relativity.

- On the other hand, free scalar field theory is conformally invariant in flat spacetime for any d, as the $R\phi^2$ term vanishes identically. (This is also the case for massless ϕ^4 theory in $d = 4$, and massless ϕ^6 theory in $d = 6$.) Note that in both cases, the canonical stress-energy tensor is not traceless, but it can be made traceless.

- Given the previous example, it is tempting to conclude that all scale-invariant theories are conformally invariant in flat spacetime, but this simply isn't true. For example, electromagnetism in $d = 3$ is a counterexample, as shown here.

- Generally, interacting theories that are conformal at the classical level are not conformally invariant at the quantum level, as they have nonvanishing beta functions and hence scale dependence. Conceptually, this occurs because the path integral measure fails to be conformally invariant.

- Fixed points of Wilsonian RG flow have quantum scale invariance because they obey the appropriate Ward identities, as we showed in the notes on Quantum Field Theory. In all cases we will deal with, quantum scale invariance comes with full quantum conformal invariance, so we will treat the two as equivalent. For example, the Wilson-Fisher fixed point hence has conformal invariance at the quantum level. At this point, all particles are massless.

- String theory has Weyl invariance and diffeomorphism invariance on the worldsheet, and hence is a conformal field theory at the classical level. Since these two symmetries are gauged, they must survive at the quantum level, and requiring the Weyl anomaly to vanish fixes the spacetime dimension. Hence string theory is a 2D quantum conformal field theory.

- Conformal symmetry is thought to be the most general spacetime symmetry, unless one includes supersymmetry, in which case we have superconformal field theories (SCFTs). For example, $\mathcal{N} = 4$ super Yang–Mills (SYM) is an SCFT, both at the classical and quantum level. This is the CFT that appears in the most prominent version of the AdS/CFT correspondence.

- Formally, CFTs can be used to define quantum field theories without reference to a Lagrangian. In fact, some CFTs do not have any known Lagrangian description, such as the 6d $(2, 0)$ SCFT. In the conformal bootstrap program, one attempts to solve a theory using only conformal invariance and consistency conditions.

4.2 Elementary Aspects

We now consider some elementary aspects of conformal invariance. We begin with establishing notation and conventions.

- We will work on a Euclidean worldsheet with coordinates $(\sigma^1, \sigma^2) = (\sigma^1, i\sigma^0)$. It is useful to work with the complex coordinates

$$z = \sigma^1 + i\sigma^2, \quad \overline{z} = \sigma^1 - i\sigma^2$$

  which are the Euclidean analogues of the lightcone coordinates. As such, we will refer to holomorphic functions as "left-moving" and anti-holomorphic functions as "right-moving".

- The holomorphic derivatives are

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

  which are defined so that

$$\partial z = \overline{\partial}\overline{z} = 0, \quad \partial\overline{z} = \overline{\partial} z = 0.$$

  Also note that $\partial^2 \equiv \partial_1^2 + \partial_2^2 = 4\partial\overline{\partial}$.

- We work in flat Euclidean space, with metric

$$ds^2 = (d\sigma^1)^2 + (d\sigma^2)^2 = dz\, d\overline{z}.$$

  In components, this means

$$g_{zz} = g_{\overline{z}\overline{z}} = 0, \quad g_{z\overline{z}} = g_{\overline{z}z} = \frac{1}{2}.$$

- Note that there are two possible integration measures,

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

  We define two delta functions with different normalization,

$$\int d^2z\, \delta(z, \overline{z}) = \int d^2\sigma\, \delta(\sigma) = 1.$$

- Vectors naturally have their indices up, with

$$v^z = v^1 + iv^2, \quad v^{\overline{z}} = v^1 - iv^2$$

  and indices are lowered using the metric, giving

$$v_z = \frac{1}{2}(v^1 - iv^2), \quad v_{\overline{z}} = \frac{1}{2}(v^1 + iv^2).$$

- Note that we are heuristically treating z and $\overline{z}$ as independent complex variables. This is justified as in the notes on Quantum Field Theory. We may think of working in the larger space $\mathbb{C}^2$, and imposing the constraint that $\overline{z}$ is the conjugate of z at the end.

- In two dimensional Euclidean space, all holomorphic changes of coordinates

$$z \to z' = f(z), \quad \overline{z} \to \overline{z}' = \overline{f(z)}$$

yield conformal transformations, because

$$ds^2 = dz\, d\overline{z} \to |df/dz|^2\, dz\, d\overline{z}.$$

Hence the conformal group in two dimensions is infinite-dimensional, which makes conformal symmetry much more powerful.

- Finally, it is conventional to define the stress-energy tensor as

$$T_{\alpha\beta} = -\frac{4\pi}{\sqrt{g}} \frac{\delta S}{\delta g^{\alpha\beta}}$$

in string theory. We have $\nabla^\alpha T_{\alpha\beta} = 0$, which reduces to $\partial^\alpha T_{\alpha\beta} = 0$ for a flat worldsheet.

- As we saw in the notes on General Relativity, conformal invariance implies the stress-energy tensor is traceless on-shell. As a rough converse, if the stress-energy tensor is traceless identically, it can be shown that all matter fields have vanishing conformal weight, and that the full theory is conformally invariant. This holds for string theory, where the matter fields are the X^μ.

- In complex coordinates, the vanishing of the trace becomes

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

and the conservation equation becomes $\partial T^{zz} = \overline{\partial} T^{\overline{z}\overline{z}} = 0$. Lowering the indices,

$$\overline{\partial} T_{zz} = \partial T_{\overline{z}\overline{z}} = 0$$

so $T_{zz}(z) \equiv T(z)$ is holomorphic and $T_{\overline{z}\overline{z}}(\overline{z}) \equiv \overline{T}(\overline{z})$ is anti-holomorphic on-shell.

- Since there are infinitely many conformal transformations, there are infinitely many conserved quantities. In particular, consider a conformal transformation

$$z \to z' = z + \epsilon(z), \quad \overline{z}' \to \overline{z} + \overline{\epsilon}(\overline{z}).$$

As usual for a position-dependent translation, the change in the action is the stress-energy tensor weighted by the translation,

$$\delta S \propto \int d^2\sigma\, T_{\alpha\beta} \partial^\alpha \delta\sigma^\beta = \frac{1}{2} \int d^2 z\, T_{zz}(\partial^z \delta z) + T_{\overline{z}\overline{z}}(\partial^{\overline{z}} \delta\overline{z}) = \int d^2 z\, T(z) \partial_{\overline{z}} \epsilon + \overline{T}(\overline{z}) \partial_z \overline{\epsilon}.$$

where we used $T_{z\overline{z}} = 0$. This automatically vanishes since ϵ is holomorphic, as expected.

- In the usual Noether trick, we promote the parameter ϵ of an infinitesimal symmetry to a position-dependent $\epsilon(x)$, and the conserved current is the part of δS proportional to $\partial_\mu \epsilon$. In this case, conformal symmetry already gives a position-dependent $\epsilon(z)$, but we can think of z and $\overline{z}$ as independent and promote it to $\epsilon(z) f(\overline{z})$, then look for the coefficient of $\overline{\partial} f$.

- Taking $\delta z = \epsilon(z) f(\overline{z})$ and $\delta\overline{z} = 0$ gives

$$J^z = 0, \quad J^{\overline{z}} = T(z)\epsilon(z).$$

  This is conserved, as $\partial_\alpha J^\alpha = \partial_z J^z = \partial_{\overline{z}} J^{\overline{z}} = 0$. Similarly, we can consider $\delta\overline{z} = \overline{\epsilon}(\overline{z}) f(z)$ with $\delta z = 0$, giving the anti-holomorphic current

$$\overline{J}^z = \overline{T}(\overline{z})\overline{\epsilon}(\overline{z}), \quad \overline{J}^{\overline{z}} = 0.$$

Next, we discuss the operator product expansion, a useful tool in the quantum theory.

- In CFT, a field refers to any local operator. The operator product expansion (OPE) states that the product of two fields can be expanded as a series as

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

  Usually one thinks of both sides as implicitly underneath a path integral. In this case, the OPE states that the path integrals of both sides, times other (distant) operator insertions, should be equal; in other words, time-ordered correlation functions involving both sides match.

- One might wonder under what conditions the OPE holds. The general intuition is that as $z \to w$, the path integral can be split into an integral over field values near z and w, and other fields which are far away,

$$\int \mathcal{D}\phi_> e^{iS_>}(\text{other, distant operators}) \int \mathcal{D}\phi_< e^{iS_<} \mathcal{O}_i \mathcal{O}_j.$$

  The inner path integral is constrained so that the values of $\phi_<$ match onto those of $\phi_>$ at the boundary. Hence it is simply some function of those boundary values.

- On the other hand, these boundary field values in turn can be written in terms of the field values and their derivatives when extrapolated to w. **(how, precisely?)** Hence the inner path integral can be written in terms of these values and the separation $z - w$, giving,

$$\int \mathcal{D}\phi_> e^{iS_>}(\text{other, distant operators}) \sum_k C_{ij}^k \mathcal{O}_k.$$

  Finally, in the limit $z \to w$ can replace $\mathcal{D}\phi_>$ with $\mathcal{D}\phi$ and $S_>$ with S, justifying the OPE, which thus holds in the limit where all other operator insertions are distant from z and w.

- It turns out that in a CFT, the OPE is an exact statement; the radius of convergence is precisely the distance to the nearest other insertion.

- In general, the coefficient functions C_{ij}^k will exhibit singular behavior like $1/(z-w)$ as $z \to w$. This singular behavior is often the only behavior we care about for practical applications; it will give us the commutation relations, and perturbatively it is related to the high-momentum behavior of loop diagrams. By naive dimensional analysis, more complicated operators on the right come with less singular behavior, which makes truncating the OPE practical. (At weak coupling, quantum effects only slightly modify this dimensional analysis argument.) For example, the OPE may be used in this way to describe hard scattering in QCD.

- A useful tool will be the Ward–Takahashi identity from the notes on Quantum Field Theory,

$$\partial_\mu \left\langle j^\mu(x) \prod_{i=1}^{n} \mathcal{O}_i(x_i) \right\rangle = -i \sum_{i=1}^{n} \left\langle \mathcal{O}_1(x_1) \ldots \widehat{\mathcal{O}}_i(x_i) \delta\mathcal{O}_i(x)\delta(x - x_i) \ldots \mathcal{O}_n(x_n) \right\rangle .$$

The form we will most often use comes by integrating both sides against $f(\sigma)$, a piecewise constant function which is 1 near x_1 and 0 at all the other x_i. Switching to our conventions,

$$-\frac{1}{2\pi} \int_f \partial_\alpha \langle J^\alpha(\sigma)\mathcal{O}_1(\sigma_1)\ldots\rangle = \langle \delta\mathcal{O}_1(\sigma_1)\ldots\rangle, \qquad \int_f \equiv \int d\sigma\, f(\sigma).$$

We will call this the Ward identity.

- The left-hand side can be simplified by noting that

$$\int_f \partial_\alpha J^\alpha = \oint_{\partial f} J_1\, d\sigma^2 - J_2\, d\sigma^1 = -i \oint_{\partial f} J_z\, dz - J_{\overline{z}}\, d\overline{z}$$

using Stokes' theorem. Performing this manipulation, we have

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

- Specializing to conformal transformations, J_z is holomorphic and $J_{\overline{z}}$ is anti-holomorphic. This makes it tempting to conclude the left-hand side is zero, but we recall the OPE has singular terms, including one of the form

$$J_z(z)\mathcal{O}_1(w,\overline{w}) \supset \frac{\operatorname{Res} J_z \mathcal{O}_1(w,\overline{w})}{z-w}.$$

We can again consider a conformal transformation as having two independent pieces. From $\delta z = \epsilon(z)$, we get

$$\delta\mathcal{O}_1(\sigma_1) = -\operatorname{Res} J_z(z)\mathcal{O}_1(\sigma_1) = -\operatorname{Res} \epsilon(z) T(z)\mathcal{O}_1(\sigma_1)$$

while from $\delta\overline{z} = \overline{\epsilon}(\overline{z})$, we get

$$\delta\mathcal{O}_1(\sigma_1) = -\operatorname{Res} \overline{J}_z(\overline{z})\mathcal{O}_1(\sigma_1) = -\operatorname{Res} \overline{\epsilon}(\overline{z}) \overline{T}(\overline{z})\mathcal{O}_1(\sigma_1)$$

where an extra sign flip occurs because the $d\overline{z}$ integral is traversed the opposite direction. Again, all these identities are implicitly underneath a path integral. We conclude that knowing the OPE between an operator and $T(z)$ and $\overline{T}(\overline{z})$ gives how it transforms under conformal symmetry.

- For example, under translations $\delta z = \epsilon$ with ϵ constant, we must have

$$\delta\mathcal{O}(z) = -\epsilon\partial\mathcal{O}(z).$$

The Ward identity hence tells us that

$$T(z)\mathcal{O}(w,\overline{w}) \supset \frac{\partial\mathcal{O}(w,\overline{w})}{z-w}, \quad \overline{T}(\overline{z})\mathcal{O}(w,\overline{w}) \supset \frac{\overline{\partial}\mathcal{O}(w,\overline{w})}{\overline{z}-\overline{w}}.$$

- Rotations and scalings are given by

$$\delta z = \epsilon z, \quad \delta \overline{z} = \overline{\epsilon z}$$

with ϵ again constant. We can't just write down $\delta\mathcal{O}$ in this case, as even classically we get a scaling that depends on the mass dimension of $\mathcal{O}$. Furthermore, we would not have a definite transformation property for $\mathcal{O}$ if it were the sum of two operators with different mass dimensions. These problems are exacerbated in the quantum theory.

- Instead, we choose a basis of local operators with good transformation properties under rotations and dilations. We say an operator $\mathcal{O}$ is quasi-primary with weight $(h, \overline{h})$ if, under $\delta z = \epsilon z$ and $\delta\overline{z} = \overline{\epsilon z}$,

$$\delta\mathcal{O} = -\epsilon(h\mathcal{O} + z\partial\mathcal{O}) - \overline{\epsilon}(\tilde{h}\mathcal{O} + \overline{z}\overline{\partial}\mathcal{O}).$$

Note that the $\partial\mathcal{O}$ and $\overline{\partial}\mathcal{O}$ terms must always be present. Rotations correspond to imaginary ϵ, where $\overline{\epsilon} = -\epsilon$, so the eigenvalue under rotation is the spin

$$s = h - \overline{h}.$$

Dilations correspond to real ϵ, where $\overline{\epsilon} = \epsilon$, so the scaling dimension is

$$\Delta = h + \overline{h}.$$

We will show later that h and $\overline{h}$ are real, and nonnegative in a unitary CFT.

- The Noether current arising from rotations and scaling is simply $J(z) = zT(z)$. Hence the residue of the $J\mathcal{O}$ OPE determines the $1/z^2$ term in the $T\mathcal{O}$ OPE, so

$$T(z)\mathcal{O}(w,\overline{w}) \supset h\frac{\mathcal{O}(w,\overline{w})}{(z-w)^2} + \frac{\partial\mathcal{O}(w,\overline{w})}{z-w}, \quad \overline{T}(\overline{z})\mathcal{O}(w,\overline{w}) \supset \tilde{h}\frac{\mathcal{O}(w,\overline{w})}{(\overline{z}-\overline{w})^2} + \frac{\overline{\partial}\mathcal{O}(w,\overline{w})}{\overline{z}-\overline{w}}.$$

- We define a primary operator as one whose OPE with T and $\overline{T}$ has no further singular terms. Since we know all the singular terms in the OPE, we can reconstruct its transformation under all conformal transformations. Focusing on a transformation $\delta z = \epsilon(z)$, we have

$$\begin{aligned}
\delta\mathcal{O}(w,\overline{w}) &= -\operatorname{Res}\epsilon(z)T(z)\mathcal{O}(w,\overline{w}) \\
&= -\operatorname{Res}\left((\epsilon(w) + \epsilon'(w)(z-w) + \ldots)\left(h\frac{\mathcal{O}(w,\overline{w})}{(z-w)^2} + \frac{\partial\mathcal{O}(w,\overline{w})}{z-w} + \ldots\right)\right) \\
&= -h\epsilon'(w)\mathcal{O}(w,\overline{w}) - \epsilon(w)\partial\mathcal{O}(w,\overline{w}).
\end{aligned}$$

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

- Integrating this, we find that for a finite conformal transformation $z \to \tilde{z}$, $\overline{z} \to \overline{\tilde{z}}$,

$$\mathcal{O}(z,\overline{z}) \to \tilde{\mathcal{O}}(\tilde{z},\overline{\tilde{z}}) = \left(\frac{\partial\tilde{z}}{\partial z}\right)^{-h}\left(\frac{\partial\overline{\tilde{z}}}{\partial\overline{z}}\right)^{-\tilde{h}}\mathcal{O}(z,\overline{z}).$$

As discussed in more detail in the notes on Conformal Field Theory, there is a distinction between globally and locally defined conformal transformations; the latter are unique to $d = 2$, and are used whenever we consider transformations such as $\delta z = \epsilon z$ for a general holomorphic ϵ. Quasi-primary operators only satisfy the above transformation law for globally defined conformal transformations; they may have additional more-singular terms in the OPE.

- In CFT, one of the main objects of interest will be the spectrum of weights $(h, \tilde{h})$ of primary fields; this is equivalent to computing the particle mass spectrum in a quantum field theory. In statistical field theory, the weights of primary operators will yield the critical exponents.

4.3 Free Scalar Field

As an extended example, we consider the free scalar field in $d = 2$ and Euclidean signature.

- The action is
$$S = \frac{1}{4\pi\alpha'} \int d^2\sigma \, \partial_\alpha X \partial^\alpha X.$$
As stated earlier, this theory is classically conformally invariant. Varying the action with respect to the metric, the stress-energy tensor is
$$T_{\alpha\beta} = -\frac{1}{\alpha'} \left(\partial_\alpha X \partial_\beta X - \frac{1}{2} \delta_{\alpha\beta} (\partial X)^2 \right)$$
which is indeed traceless off-shell; this is special to $d = 2$.

- In complex coordinates, we have $T_{z\overline{z}} = 0$ and
$$T = -\frac{1}{\alpha'} \partial X \partial X, \quad \overline{T} = -\frac{1}{\alpha'} \overline{\partial} X \overline{\partial} X.$$
The equation of motion for X is $\partial\overline{\partial} X = 0$, and the general solution is
$$X(z, \overline{z}) = X(z) + \overline{X}(\overline{z}).$$
Then we see T and $\overline{T}$ are holomorphic and anti-holomorphic on-shell, as expected.

- To compute the propagator, note that
$$0 = \int \mathcal{D}X \frac{\delta}{\delta X(\sigma)} (e^{-S} X(\sigma')) = \int \mathcal{D}X \, e^{-S} \left(\frac{1}{2\pi\alpha'} \partial^2 X(\sigma) X(\sigma') + \delta(\sigma - \sigma') \right)$$
which tells us that
$$\partial^2 \langle X(\sigma) X(\sigma') \rangle = -2\pi\alpha' \delta(\sigma - \sigma').$$
This is a differential equation for the propagator.

- To solve it, note that $\partial^2 \log(\sigma^2)$ is simply the divergence of $2\sigma_\alpha/\sigma^2$, and the flux of this vector field about a curve around the origin is 2π. Then
$$\partial^2 \log(\sigma^2) = 4\pi\delta(\sigma)$$
from which we conclude
$$\langle X(\sigma) X(\sigma') \rangle = -\frac{\alpha'}{2} \log(\sigma - \sigma')^2.$$
The propagator has a UV singularity as $\sigma \to \sigma'$, as common to all field theories, but also a large-distance singularity. In the context of statistical field theory, this is another statement of the Mermin-Wagner theorem: fluctuations grow at large distances, so that spontaneous symmetry breaking is impossible.

- By repeating this derivation with other operator insertions in the path integral, we conclude

$$X(\sigma)X(\sigma') \supset -\frac{\alpha'}{2}\log(\sigma-\sigma')^2.$$

We can also consider the holomorphic piece of X, giving

$$X(z)X(w) \supset -\frac{\alpha'}{2}\log(z-w).$$

Evidently, X does not behave nicely under conformal transformations, but its derivative does,

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

- Now we investigate the OPE of T with other operators. However, we have just seen that the classical definition of T is singular in the quantum theory. In canonical quantization, we can fix this by normal ordering. Here, we do something effectively similar,

$$T \equiv -\frac{1}{\alpha'}:\partial X\partial X: \equiv -\frac{1}{\alpha'}\lim_{z\to w}\left(\partial X(z)\partial X(w) - \langle \partial X(z)\partial X(w)\rangle\right).$$

Like normal ordering, this enforces $\langle T\rangle = 0$.

- Now we would like to consider the product

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

Consider time-ordering both sides. (Unfortunately, there is no way to write this in our notation, because a time-ordered product looks exactly like a regular product; however, this is standard.) By Wick's theorem, the time-ordering of the right-hand side is a totally normal-ordered term plus all possible partial contractions.

- In this case we can contract either $\partial X(z)$ with $\partial X(w)$, giving a propagator; we don't contract the two $\partial X(z)$ factors because they are already normal ordered. Hence we have

$$T(z)\partial X(w) \sim -\frac{2}{\alpha'}\partial X(z)\left(-\frac{\alpha'}{2}\frac{1}{(z-w)^2}\right)$$

where $\sim$ indicates we are dropping non-singular terms. Note that the use of ∂ is somewhat ambiguous; if it acts on a single field, it is the derivative with respect to that field's argument, $\partial X(z) \equiv \partial_z X(z)$ and $\partial X(w) \equiv \partial_w X(w)$. When it is not clear, we will indicate it with a subscript.

- The right-hand side of the OPE needs fields only evaluated at w, so we expand

$$\partial X(z) = \partial X(w) + (z-w)\partial^2 X(w).$$

This gives

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

This shows that ∂X is a primary field with weight $h=1$. Since the $\langle X\overline{X}\rangle$ propagator vanishes **(why? this is a bit unintuitive)** , it has $\tilde{h}=0$. Similarly, $\overline{\partial X}$ is primary with weight $(0,1)$.

- Applying a derivative ∂ still increments h by one, as we would expect classically. Then
$$T(z)\partial^2 X(w) \sim \partial_w \left(\frac{\partial X(w)}{(z-w)^2} \right) \sim \frac{2\partial X(w)}{(z-w)^3} + \frac{\partial^2 X(w)}{(z-w)^2}$$
which indicates $\partial^2 X$ has weight $h = 2$, and similarly $\tilde{h} = 0$, but is not primary. It is tempting to conclude that X has weight $(0,0)$ as we would expect classically, but it has no well-defined weights at all.

- Another important field is $:e^{ikX}:$, which has the OPE
$$\partial X(z) :e^{ikX(w)}: = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \partial X(z) :X(w)^n: \sim \sum_{n=1}^{\infty} \frac{(ik)^n}{(n-1)!} :X(w)^{n-1}: \left(-\frac{\alpha'}{2} \frac{1}{z-w} \right)$$
where everything is implicitly time-ordered, and we used Wick's theorem. Shifting the sum,
$$\partial X(z) :e^{ikX(w)}: \sim -\frac{i\alpha' k}{2} \frac{:e^{ikX(w)}:}{z-w}.$$

- This allows us to compute the OPE with T,
$$T(z) :e^{ikX(w)}: = -\frac{1}{\alpha'} :\partial X(z) \partial X(z): :e^{ikX(w)}: \sim \frac{\alpha' k^2}{4} \frac{:e^{ikX(w)}:}{(z-w)^2} + ik \frac{:\partial X(z) e^{ikX(w)}:}{z-w}$$
where the two terms come from performing two or one contraction, respectively. Expanding the $\partial X(z)$ and throwing away a non-singular term,
$$T_z :e^{ikX(w)}: \sim \frac{\alpha' k^2}{4} \frac{:e^{ikX(w)}:}{(z-w)^2} + \frac{\partial :e^{ikX(w)}:}{z-w}.$$
This shows that $:e^{ikX(w)}:$ is primary with weight $(\alpha' k^2/4, 0)$. These nonzero weights are a quantum effect, as α' sits outside the action where $\hbar$ would. From this point on we drop the normal ordering symbols, leaving them implicit.

- Similar reasoning shows that $e^{ik\overline{X}}$ is primary with weight $(0, \alpha' k^2/4)$, while the full $e^{ikX(z,\overline{z})}$ is primary with weight $(\alpha' k^2/4, \alpha' k^2/4)$. This is a bit unintuitive; in some sense these results are due to the normal ordering.

- Finally, we can compute the OPE of T with itself,
$$\begin{aligned} T(z)T(w) &= \frac{1}{\alpha'^2} :\partial X(z) \partial X(z): :\partial X(w) \partial X(w): \\ &\sim \frac{2}{\alpha'^2} \left(-\frac{\alpha'}{2} \frac{1}{(z-w)^2} \right)^2 - \frac{4}{\alpha'^2} \frac{\alpha'}{2} \frac{:\partial X(z) \partial X(w):}{(z-w)^2} \\ &\sim \frac{1/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} - \frac{2}{\alpha'} \frac{\partial^2 X(w) \partial X(w)}{z-w} \\ &\sim \frac{1/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w} \end{aligned}$$
where we performed contractions (four single contractions, two double contractions), then Taylor expanded $\partial X(z)$. Then T has weight $(2,0)$ but is not a primary operator, because of the first term. This holds in all 2D CFTs, as T must have spin $s = 2$ since it is a symmetric rank 2 tensor, and dimension $\Delta = 2$ since it is integrated over space to get the energy. Similarly, $\overline{T}$ has weight $(0,2)$.

4.4 Central Charges

The extra term in the TT OPE is known as a central charge.

- In general, the TT OPE has the form

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

while the $\overline{T}\overline{T}$ OPE has the form

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

and c and $\tilde{c}$ are called the central charges, and are positive in all CFTs. Evidently, for a free scalar field $c = \tilde{c} = 1$. In general the central charges heuristically measure the number of degrees of freedom, though they are not always integers.

- This is the most general possible form. There are no more-singular terms because every term on the right-hand side must have dimension $\Delta = 4$, but in a unitary CFT there are no operators with $h, \tilde{h} < 0$. There is no $1/(z-w)^3$ term because the OPE must be symmetric; note that the given expression is symmetric because

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

- The OPE of T with T gives us the infinitesimal transformation of T,

$$\delta T(w) = -\operatorname{Res}(\epsilon(z)T(z)T(w)) = -\operatorname{Res}\left(\epsilon(z)\left(\frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w} + \ldots\right)\right).$$

Assuming $\epsilon(z)$ is non-singular, this may be computed by Taylor expanding it, giving

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

- It turns out the finite version of this transformation is

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

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

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

Note that the Schwarzian vanishes for globally defined conformal transformations; this means that T is merely a quasi-primary operator, not a primary operator. In fact, one can show that this is the most general possible transformation law for a dimension 2 quasi-primary operator.

We now turn to the interpretation of the central charge.

- Note that the extra term in the transformation of T does not depend on T itself. This means that it will be independent of the state; it can be interpreted as the energy of the "zero mode", i.e. the Casimir energy.

- As one example, consider the nontrivial conformal transformation corresponding to the diffeomorphism
$$z = e^{-iw}, \quad w = \sigma + i\tau.$$
This maps the complex plane (minus the origin) to itself, except that the coordinate $\sigma \in [0, 2\pi)$ is periodic. Using $\partial w/\partial z = i/z$, and so on, the transformation of T is
$$T_{\text{cyl}}(w) = -z^2 T_{\text{plane}}(z) + \frac{c}{24}.$$

- Suppose the ground state energy vanishes on the plane. Then on the cylinder,
$$H = \int d\sigma\, T_{\tau\tau} = -\int d\sigma\, T_{ww} + \overline{T}_{\overline{ww}} = -\frac{2\pi(c+\tilde{c})}{24}.$$
For example, for a free scalar field the energy density is $E/2\pi = -1/12$.

A closely related consequence of the central charge is the Weyl anomaly.

- The central charge also causes the stress-energy tensor to acquire a trace. We recall that classically, $T^\alpha{}_\alpha = 0$ in a Weyl-invariant theory. At the quantum level, $\langle T^\alpha{}_\alpha \rangle$ is not necessarily zero. We expect that it will be the same for every state of the theory, since it is a result of regulating short-distance divergences, which cannot see the long-distance behavior of a finite-energy state.

- As a result, $\langle T^\alpha{}_\alpha \rangle$ can only depend on the background metric, and it must be local with dimension 2. The only candidate in 2D is the Ricci scalar, and in fact the result is
$$\langle T^\alpha{}_\alpha \rangle = -\frac{c}{12} R.$$
Note that we are generalizing our CFT to curved backgrounds here. Our theory still has conformal symmetry on a flat background, but the presence of the central charge means it no longer has Weyl symmetry. Hence this result is called the Weyl anomaly or trace anomaly.

- In higher dimensions there are also analogous results, e.g. in 4D we have
$$\langle T^\mu{}_\mu \rangle = \frac{c}{16\pi^2} C_{\rho\sigma\kappa\lambda} C^{\rho\sigma\kappa\lambda} - \frac{a}{16\pi^2} \tilde{R}_{\rho\sigma\kappa\lambda} \tilde{R}^{\rho\sigma\kappa\lambda}$$
where C is the Weyl tensor and $\tilde{R}$ is the dual of the Riemann tensor.

- To maintain Weyl invariance, we would need $c = 0$. It also turns out that for a CFT to be consistent on curved spacetime, we need $c = \tilde{c}$. Having $c \neq \tilde{c}$ is perfectly consistent for CFTs on a flat background, but is an example of a gravitational anomaly.

- To prove the Weyl anomaly formula, first note that by a diffeomorphism we can write any 2D metric in the form $g_{\alpha\beta} = e^{2\omega}\delta_{\alpha\beta}$, in which case the Ricci scalar is
$$R = -2e^{-2\omega}\partial^2\omega.$$

- Now, we consider the $T_{z\overline{z}}T_{w\overline{w}}$ OPE. Energy conservation is $\partial T_{z\overline{z}} = -\overline{\partial} T_{zz}$, so we have

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

  It is tempting to conclude that the right-hand side is zero, because it contains anti-holomorphic derivatives of a holomorphic quantity, but the point of the OPE is that there are terms that are not holomorphic, but rather singular. The singular terms must be evaluated by integration.

- For example, earlier we saw that

$$\partial^2 \log(\sigma^2) = 4\pi\delta(\sigma)$$

  which in our complex notation is

$$\overline{\partial}\partial \log(z\overline{z}) = \overline{\partial}\frac{1}{z} = 2\pi\delta(z,\overline{z}).$$

  Therefore, we may compute

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

- Therefore, removing a ∂_z and ∂_w, we find

$$T_{z\overline{z}}(z,\overline{z})T_{w\overline{w}}(w,\overline{w}) \sim \frac{c\pi}{6}\,\partial_z\overline{\partial}_{\overline{w}}\delta(z-w,\overline{z}-\overline{w}) + \ldots.$$

  Hence the OPE vanishes for finite separations, but with singular behavior at $z = w$. Such a term is referred to as a contact term. There are more terms from the rest of the $T_{zz}T_{ww}$ OPE we ignored, but we won't need them since it turns out we only need the leading singularity. It is unclear in this approach how we would compute the subleading terms; a more general approach using an explicit regulator is given here.

- Now, starting with flat space and performing an infinitesimal Weyl transformation $\delta g_{\alpha\beta} = 2\omega\delta_{\alpha\beta}$, we have $\delta g^{\alpha\beta} = -2\omega\delta^{\alpha\beta}$ and hence

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

  The right-hand side can now be evaluated by the OPE we just computed. Carefully accounting for factors of 2 when transforming back to Cartesian coordinates, we get

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

- Plugging this in and integrating by parts, we have

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

  where in the last step we simply took the tensorial generalization of our result, which is now valid even after a finite Weyl transformation from flat space.

- This interpretation of the central charge is very closely related to the vacuum energy. Note that classically $T^\alpha{}_\alpha$ vanishes identically, without even using the equations of motion. Hence it seems impossible for any weighted average of $T^\alpha{}_\alpha$ to give a nonzero result. The resolution is that in passing to the quantum theory, we normal-ordered $T_{\alpha\beta}$, subtracting off the vacuum energy. It is precisely this subtraction that allowed the stress-energy tensor to pick up a trace.

Yet more interpretations of the central charge come from thinking of the CFT as a statistical field theory.

- We consider a conformal field theory on a Euclidean torus, with $\sigma \in [0, 2\pi)$ and $\tau \in [0, \beta)$. As shown in the notes on Quantum Field Theory, the partition function of such a theory is related to the free energy at temperature $T = 1/\beta$,
$$Z[\beta] = \operatorname{tr} e^{-\beta H} = e^{-\beta F}.$$
- At low temperatures, $\beta \to \infty$, the free energy is dominated by the lowest energy state. This is precisely the vacuum, with vacuum energy $H = -c/12$, and hence
$$\lim_{\beta\to\infty} Z[\beta] = e^{c\beta/12}.$$
- However, in Euclidean space, both directions of the torus are on a perfectly equal footing, so we can treat τ as the spatial coordinate instead. To scale the range of τ, perform the conformal transformation associated with the coordinate transformation
$$\tau \to \frac{2\pi}{\beta}\tau, \quad \sigma \to \frac{2\pi}{\beta}\sigma$$
at which point the coordinate ranges are
$$\sigma \in [0, 4\pi^2/\beta), \quad \tau \in [0, 2\pi).$$
This tells us that
$$Z[4\pi^2/\beta] = Z[\beta].$$
This is one consequence of modular invariance, a symmetry of CFTs on the Euclidean torus, which we will investigate in more detail later.
- Using our result, we have
$$\lim_{\beta\to 0} Z[\beta] = e^{c\pi^2/3\beta}.$$
To relate this with the density of states $\rho(E) = e^{S(E)}$, note that
$$e^{-\beta F} = \int dE\, e^{S(E)} e^{-\beta E}.$$
If there are N degrees of freedom, then in two dimensions $S(E) \sim N\sqrt{E}$ at high energies, and applying the saddle point approximation gives $F \sim N^2 T^2$. Hence we have $N \sim \sqrt{c}$ **(I thought $N \sim c$?)** and the asymptotic result
$$S(E) \sim \sqrt{cE},$$
known as Cardy's formula.

- Finally, we can think of Wilsonian RG flow on the space of theories. CFTs are fixed points of the RG flow, making them ubiquitous and important. For example, a critical statistical field theory is a CFT, with universality occurring as different theories flow to the same IR fixed point. In high energy physics, QFTs with a sensible known UV limit presumably flow from a UV fixed point and hence can be viewed as a deformation of a CFT.
- Zamalodchikov's c-theorem states that there is a function c on the space of all theories, which monotonically decreases along RG flows and coincides with the central charge at fixed points. This formalizes how c measures degrees of freedom, which are integrated out during RG flow. This was generalized to even dimensions in Cardy's a-theorem.

4.5 The Virasoro Algebra

Now we investigate the states in a CFT.

- It is useful to focus on the CFT on a cylinder. Here, the states live on slices of constant σ and evolve by the Hamiltonian $H = \partial_\tau$. After conformal mapping to the plane, the Hamiltonian becomes the dilation operator $D = z\partial + \overline{z}\overline{\partial}$, which means the states should live on circles of constant radius.
- Thus, to compute time-ordered quantities on the cylinder, we need to apply radial ordering on the plane. This general approach is called radial quantization.
- Now, we decompose the stress tensor $T(z)$ on the cylinder as

$$T_{\text{cyl}}(w) = -\sum_m L_m e^{imw} + \frac{c}{24}$$

  where the sum runs over all integer m. After conformal transformation to the plane,

$$T(z) = \sum_m \frac{L_m}{z^{m+2}}.$$

  Similarly, for the right-moving sector we have

$$\overline{T}(\overline{z}) = \sum_m \frac{\tilde{L}_m}{\overline{z}^{m+2}}.$$

- This can be inverted by a suitable contour integral,

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

  where $\oint dz$ means any counterclockwise contour encircling the origin once, and $\oint d\overline{z}$ means any clockwise contour encircling the origin once.
- Recall that the conserved current associated with $\delta z = z^{n+1}$ is $J(z) = z^{n+1}T(z)$. This indicates that L_n is the conserved charge associated with this conformal transformation, because it is the current integrated over a spatial (i.e. radial) slice. Similarly, $\tilde{L}_n$ is the conserved charged associated with $\delta\overline{z} = \overline{z}^{n+1}$.

- The L_n and $\tilde{L}_n$ are known as Virasoro generators, just as we saw earlier. The most important examples are L_{-1} and $\tilde{L}_{-1}$, which generate translations in the plane, and L_0 and $\tilde{L}_0$, which generate scalings and rotations. The Hamiltonian is a pure scaling, $H = D = L_0 + \tilde{L}_0$.

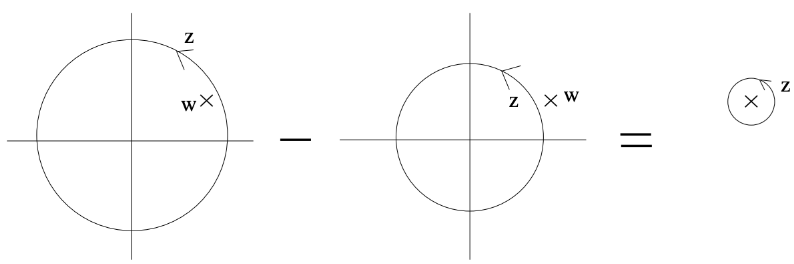
- Note the close similarity to what we encountered quantizing the bosonic string. In that case, we worked entirely on the cylinder: the closed string was automatically defined on a cylinder, while we could double the range of the open string to make it so. We defined the L_n and $\tilde{L}_n$ in the exact same way.

- We may compute the Virasoro algebra using the TT OPE. We note that

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

This notation is quite ambiguous. The point is that we are using the OPE only to refer to radially ordered correlation functions, so $T(z)$ must be at a larger radius than $T(w)$. So in the first term, the $\oint dz$ contour is at a greater radius than the $\oint dw$ contour, while in the second term it is the opposite.

- To compute this, consider fixing w. Then the $\oint dz$ integrations are:



Here we have assumed there are no further operator insertions. Then the two z integrals together can be deformed to a circle about w, picking up the residue at $z = w$,

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

- To compute the residue, we expand

$$z^{m+1} = w^{m+1} + (m+1) w^m (z-w) + \frac{1}{2} m(m+1) w^{m_1} (z-w)^2 + \dots.$$

Then we get

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

Integrating the first term by parts, the first term gives the expected $(m-n)L_{m+n}$, while the third term produces the extra term in the Virasoro algebra,

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

as we found by more elementary means before. The $\tilde{L}_n$ satisfy the same algebra with c replaced by $\tilde{c}$, and $[L_m, \tilde{L}_n] = 0$. The appearance of c here justifies its name as the central charge, as it is a new term in the algebra that commutes with everything.

- To understand the appearance of the central charge, note that the diffeomorphisms $\delta z = z^{n+1}$ give the Witt algebra, as we saw earlier. The extra term for conformal transformations arises because of the extra Weyl rescaling. **(is this right? how does c affect the Weyl rescaling?)**

Using the Virasoro algebra, we can generate states.

- Suppose we have a state $|\psi\rangle$ which is an eigenstate of L_0 and $\tilde{L}_0$,
$$L_0|\psi\rangle = h|\psi\rangle, \quad \tilde{L}_0|\psi\rangle = \tilde{h}|\psi\rangle.$$
On the cylinder, this correspond to a state of energy
$$\frac{E}{2\pi} = h + \tilde{h} - \frac{c+\tilde{c}}{24}.$$
Hence we refer to h and $\tilde{h}$ as the energy of the state. Furthermore, the angular momentum of the state is $h - \tilde{h}$. Note the tempting similarity with the conformal weight; we will make this more precise with the state-operator correspondence.

- By acting with L_n operators, we get further states with eigenvalues
$$L_0 L_n|\psi\rangle = (L_n L_0 - nL_n)|\psi\rangle = (h-n)L_n|\psi\rangle$$
so L_n lowers the energy h by n, while $\tilde{L}_n$ lowers the energy $\tilde{h}$ by n.

- If the spectrum is bounded below, there must be states annihilated by L_n and $\tilde{L}_n$ for all $n > 0$. These are called primary states; they are the states of lowest energy. By acting with the L_{-n} and $\tilde{L}_{-n}$, we can construct an infinite tower of higher-energy states, called the descendants. The primary state condition corresponds to the physical state condition in covariant quantization.

- In the language of representation theory, a primary state is a highest weight state, and the whole resulting set of states generated from a primary is called a Verma module **(formal definition?)** ; it is a representation of the Virasoro algebra.

- One might wonder how these representations decompose into Lorentz representations, as the Lorentz algebra is a subalgebra of the Virasoro algebra in any dimension. Generically there may be massive representations, but the masses must form a gapless, continuous spectrum.

- The vacuum state $|0\rangle$ has zero energy, $h = \tilde{h} = 0$, and is hence annihilated by L_n and $\tilde{L}_n$ for all $n \geq 0$. This accords with our intuition that the vacuum states should have the greatest symmetry. However, it is impossible for all of the L_n and $\tilde{L}_n$ to annihilate the vacuum, as this would leave no room for the central charge term.

- It is possible that the states in the Verma module are dependent. A linear combination of states that vanishes identically is called a null state. The existence of null states depends on the values of h and c.

- It is also physically important to impose unitarity. However, it isn't possible to talk about this in Euclidean signature, so we must return to the Euclidean cylinder, and from there to the Minkowski cylinder, where the Hamiltonian density is
$$\mathcal{H} = T_{ww} + T_{\overline{ww}} = \sum_n L_n e^{-in\sigma^+} + \tilde{L}_n e^{-in\sigma^-}.$$

For the Hamiltonian to be Hermitian, we require

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

just as we saw when quantizing the string.

- Furthermore, we must demand the Hilbert space does not contain negative norm states. We cannot simply dismiss these states as unphysical; we could only do this for the string because the conformal symmetry there was gauged. **(right?)** This leads to some tight constraints.
- First, note that for any primary state $|\psi\rangle$,

$$|L_{-1}|\psi\rangle|^2 = \langle\psi|L_1 L_{-1}|\psi\rangle = \langle\psi|[L_1, L_{-1}]|\psi\rangle = 2h\langle\psi|\psi\rangle$$

which must be nonnegative, so $h \geq 0$. The only state with $h = 0$ is the vacuum state $|0\rangle$.

- Now consider the norm

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

For this to be nonnegative, we require $c \geq 0$. When $c = 0$, the only state in the vacuum module is the vacuum itself; in fact, it turns out this is the only state in the entire theory. Any nontrivial CFT has $c > 0$. There are many more such requirements; the constraints are sufficient to classify and solve all CFTs with $c < 1$.

4.6 The State-Operator Correspondence

The states in a CFT can be related to operators by the state-operator correspondence.

- States and local operators are on a very different footing in quantum field theory. Local operators live at a point, while states are wavefunctionals over field configurations defined over an entire spatial slice. However, in a CFT the distant past in the cylinder is mapped to a single point $z = 0$, giving rise to a correspondence between them.
- To make this more precise, note that in Euclidean QFT, a state $\Psi_i[\phi_i(\sigma)]$ at time τ_i evolves to

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

For a CFT on the plane, we simply replace timeslices with surfaces of constant r,

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

- Now suppose we take the initial state to the infinite past, $z = 0$. Then we must integrate over the whole disc $|z| < r_f$, and the initial state wavefunctional which weights the path integral just translates to a weighting factor at $z = 0$, which can be expressed as a local operator. Hence the effect of the initial state can be equivalently written as an operator insertion,

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

This gives a correspondence between local operators and states in the infinite past. We take a Heisenberg-like picture where we imagine states as extending through time, or in this case through radii. Hence the state $\Psi[\phi_f, r]$ is just a snapshot of the state $|\mathcal{O}\rangle$ at a given moment.

- The state-operator correspondence works because we can map the infinite past goes to a single point. Similarly, CFTs on $\mathbb{R}^D$ can be conformally mapped to $\mathbb{R} \times S^{D-1}$. **(how to show this explicitly?)**
- The state-operator correspondence shouldn't be confused with the construction of the Hilbert space by creation operators in ordinary QFT, as these operators are not local. Furthermore, in ordinary QFT, we can always map a set of operators to a state, e.g. by having them all act on the vacuum at the same time. It is only in CFT that the map goes the other direction, since in CFT we can shrink a "timeslice" to a point.
- The Euclidean path integral projects out the vacuum state when integrated over infinite Euclidean time. As such, the vacuum state $|0\rangle$ corresponds to the identity operator and is hence sometimes written as $|1\rangle$.

We now consider some simple consequences of this result.

- First, note that for a primary operator $\mathcal{O}$ and corresponding state $|\mathcal{O}\rangle$,
$$L_n|\mathcal{O}\rangle = \oint \frac{dz}{2\pi i} z^{n+1} T(z)\mathcal{O}(z=0)$$
where the right-hand side is understood to be underneath a path integral. Applying the OPE,
$$L_n|\mathcal{O}\rangle \sim \oint \frac{dz}{2\pi i} z^{n+1} \left(\frac{h\mathcal{O}}{z^2} + \frac{\partial\mathcal{O}}{z} \right).$$
- As a result, we find that
$$L_{-1}|\mathcal{O}\rangle = |\partial\mathcal{O}\rangle$$
which in fact holds for all operators; it simply states that L_{-1} is a translation. We also have
$$L_0|\mathcal{O}\rangle = h|\mathcal{O}\rangle$$
which holds for any operators with well-defined transformations under scaling, and
$$L_n|\mathcal{O}\rangle = 0, \quad n > 0$$
which tells us that primary operators correspond to primary states.
- Hence computing the spectrum of weights of primary operators corresponds to computing the spectrum of energies and angular momenta of primary states. Going the other way, diagonalizing D and L on the set of states will give us a set of local operators with good transformation properties, which we merely postulated existed earlier.
- The state-operator correspondence also explains why the OPE works so well in CFT. Consider a correlation function with two local operators inserted nearby. As argued earlier, we can perform the path integral over a small disk containing them, yielding a weighting function for the remaining path integral, which behaves like a state specified on the boundary of the disk. Putting the origin at the center of this disk, the state may be viewed as the time evolution of a state specified in the infinite past, which then corresponds to a local operator, precisely the right-hand side of the OPE. This furthermore tells us the OPE is an exact statement as long as the operators are closer to each other than to any other operators.

We now apply the operator-state correspondence to the free scalar field.

- It's convenient to expand X in a Fourier series on the Euclidean cylinder,

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

following our treatment of the closed string. The constraint of reality of X in Minkowski space gives us $\alpha_n^\dagger = \alpha_{-n}$ and $\tilde{\alpha}_n^\dagger = \tilde{\alpha}_{-n}$.

- As we've seen, X is not primary, so we instead work with ∂X,

$$\partial_w X(w) = -\sqrt{\frac{\alpha'}{2}}\sum_n \alpha_n e^{inw}, \quad \alpha_0 = i\sqrt{\frac{\alpha'}{2}}\,p.$$

Note that we have restricted to the holomorphic component. Next, since ∂X is a primary field of weight $(1,0)$ we can easily transform to the plane,

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

- We can invert this expression with a contour integral,

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

The commutation relations can be found using the OPE,

$$[\alpha_m,\alpha_n] = -\frac{2}{\alpha'}\left(\oint\frac{dz}{2\pi i}\oint\frac{dw}{2\pi i} - \oint\frac{dw}{2\pi i}\oint\frac{dz}{2\pi i}\right) z^m w^n \partial X(z)\partial X(w)$$

which by the same trick as before, gives

$$[\alpha_m,\alpha_n] = -\frac{2}{\alpha'}\oint\frac{dw}{2\pi i}\mathrm{Res}_{z=w}\left(z^m w^n\left(\frac{-\alpha'/2}{(z-w)^2}+\ldots\right)\right) = m\oint\frac{dw}{2\pi i}\,w^{m+n-1} = m\delta_{m+n}$$

where we Taylor expanded $z^m = w^m + mw^{m-1}(z-w)+\ldots$. The final result is just what we found for the bosonic string.

- Now, we know that the Fock space is defined by acting with creation operators α_{-m} with $m>0$ on the vacuum $|0\rangle$, so a general state is given by

$$\prod_{m>0}\alpha_{-m}^{k_m}|0,p\rangle$$

where the value of p corresponds to the zero mode α_0. We would like to explicitly construct the operator-state correspondence.

- Suppressing the zero mode, we start with the vacuum state $|0\rangle$, and would like to show explicitly that it corresponds with the identity operator. The defining property of the vacuum state is $\alpha_m|0\rangle = 0$ for $m>0$, while the ground state wavefunctional is

$$\Psi_0[X_f] = \int^{X_f(\tau)}\mathcal{D}X\, e^{-S[X]}.$$

Upon acting with α_m, we get

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

Note that α_n is not a local operator; instead it corresponds to insertions over an entire timeslice of constant radius. We must evaluate the right-hand side at a larger radius, since the state only vanishes after acting with α_n.

- Since the path integral favors smooth functions, we simply assume X is smooth. But then the contour integral vanishes for $m \geq 0$, giving the desired result.
- Next, we claim that

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

In order to check this, we act with α_n,

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

We then use the OPE on the right-hand side, which gives

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

just as expected.

- Finally, the zero mode arises by insertion of the primary operator e^{ipX},

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

More generally, a complete set of local operators is given by products of e^{ipX} with terms like $\partial^n X$ and $\overline{\partial}^n \overline{X}$.

Finally, we briefly comment on the open string.

- The open string lives on the infinite strip with spatial coordinate $\sigma \in [0, \pi]$. Upon the same conformal mapping $z = e^{-iw}$, it is mapped to the upper-half plane $\operatorname{Im} z \geq 0$, with the endpoints of the string mapped to the real axis.
- We have lost translational invariance in the imaginary direction, but preserve it in the real direction, so $T_{\alpha\beta} t^\beta$ remains a conserved current, where t^α is tangent to the real axis. Letting n^α be the normal vector, Neumann boundary conditions mean that none of the current flows out of the boundary, so

$$T_{\alpha\beta} n^\alpha t^\beta = 0$$

at $\operatorname{Im} z = 0$.

- In complex coordinates, $T_{zz} = T_{\overline{z}\overline{z}}$ at $\operatorname{Im} z = 0$. This allows us to extend T_{zz} to the whole complex plane by defining
$$T_{zz}(z) = T_{\overline{z}\overline{z}}(\overline{z}).$$
Hence we may equivalently think of the stress tensor as defined on the whole plane, but only containing the holomorphic component T. This allows us to carry over all our previous results, with only one set of Virasoro generators.

- A similar trick can be used to compute the propagator for the free scalar field. The scalar field $X(z, \overline{z})$ is only defined for the upper-half plane and obeys Neumann boundary conditions. These can be accounted for by adding image charges,
$$G(z, \overline{z}, w, \overline{w}) = -\frac{\alpha'}{2} \log |z - w|^2 - \frac{\alpha'}{2} \log |z - \overline{w}|^2.$$
In particular, if z and w lie on the real axis, then the propagator is the same as the closed string, but multiplied by a factor of 2.

- The infinite past is still represented by $z = 0$, a point on the boundary. More generally, this means the state-operator map only works for local operators defined on the boundary. This ensures that theories on a strip have fewer states than those on the cylinder. For example, Neumann boundary conditions require $\partial X = \overline{\partial} X$ on the boundary, so on the strip they can only give rise to the same state. This reflects the halving of oscillator modes for the open string we saw earlier.

5 String Interactions

5.1 Motivation

Now we turn to a heuristic discussion of string interactions.

- As motivated earlier, one can describe string interactions by evaluating a series of "string diagrams", where the worldsheet has nontrivial topology, and Weyl invariance on the worldsheet can be used to simplify the diagrams. For tree-level scattering of n closed strings, it can be shown that the worldsheet can be converted to a sphere with n punctures.

- At each puncture, there must appear some local operator in the worldsheet QFT with the quantum numbers of the external string state $|\Lambda\rangle$ mapped to that point, called the vertex operator V_Λ. This "state–operator correspondence" is a common idea in conformal field theories.

- Heuristically, for each particle type there should be a local operator $W_\Lambda(\sigma, \tau)$. It must be a scalar under reparametrizations of σ and τ, and have the same Lorentz quantum numbers as Λ. We can try to build W_Λ out of X^μ and its derivatives.

- Since the tachyon is a Lorentz scalar, we can simply take $W = 1$, while for the graviton the simplest spin two operator is $W^{\mu\nu} = \partial_\alpha X^\mu \partial^\alpha X^\nu$.

- However, we must also take into account spacetime translations. Under the global symmetry $X^\mu \to X^\mu + a^\mu$, the wavefunction of an external state of momentum k^μ is multiplied by $e^{ik\cdot a}$. This can be accounted for by taking a factor of $e^{ik\cdot X}$ in the vertex operator. Finally, the vertex operator should be integrated over the worldsheet, since it may appear anywhere on it, giving

$$V_\Lambda(k) = \int d^2\sigma \sqrt{h}\, W_\Lambda(\sigma, \tau) e^{ik\cdot X}.$$

  This handles both emission and absorption; by convention k is always directed inward.

- Finally, to compute scattering amplitudes, we would expect a path integral to give

$$A(\Lambda_1, k_1, \ldots, \Lambda_M, k_M) = \kappa^{M-2} \int \mathcal{D}X\mathcal{D}h\, e^{-S} \prod_{i=1}^{M} V_{\Lambda_i}(k_i)$$

  where S is the Polyakov action and κ is a coupling constant. Actually deriving this require background in string field theory and is far beyond our scope.

- In this section, we will focus on evaluating the tree-level contribution, which for closed strings gives a sphere; accounting for higher-order contributions involves a sum over Riemann surfaces of arbitrary genus.

- Focusing on the tree-level contribution, it is most convenient to stereographically project the sphere to the plane. Since we have fixed the metric h, the path integral simplifies to

$$A = \kappa^{M-2} \int \mathcal{D}X(x, y)\, \exp\left(-\frac{1}{2\pi} \int d^2x\, \partial_\alpha X_\mu \partial^\alpha X^\mu\right) \prod_{i=1}^{M} V_{\Lambda_i}(k_i) \equiv \kappa^{M-2} \left\langle \prod_{i=1}^{M} V_{\Lambda_i}(k_i) \right\rangle$$

  which is simply a free field theory correlator that can be directly evaluated.

- However, there are further issues involving the path integral measure. As usual, we should mod out by gauge symmetries, but our final configuration has residual diffeomorphism symmetry. This is easiest to see by adopting complex coordinates in the plane. The worldsheet metric must be in the form $ds^2 = e^\phi dz d\bar{z}$, and changing coordinates to $w(z)$ where w is analytic preserves this. Infinitesimally, we have transformations $\delta z = \epsilon(z)$ where ϵ is holomorphic.

- However, we must ensure that $\delta z = \epsilon(z)$ does not have a pole at the point at infinity, which corresponds to a point in the original sphere. Equivalently, if $\tilde{z} = 1/z$, then $\delta\tilde{z}$ cannot have a pole at zero, but

$$\delta\tilde{z} = -\frac{\epsilon(z)}{z^2}$$

which implies $\epsilon(z)$ is a quadratic polynomial, $\delta z = a + bz + cz^2$. These generate a group isomorphic to $SL(2,\mathbb{C})$.

- As an explicit example, consider the scattering of M tachyons. Then

$$A = \kappa^{M-2} \int \prod_{i=1}^M d^2 z_i \left\langle \prod_{i=1}^M e^{ik_i \cdot X(z_i)} \right\rangle.$$

The free field correlator can be evaluated by completing the square, where k_i plays a role like a source in ordinary field theory, giving

$$A = \kappa^{M-2} \int \prod_{i=1}^M d^2 z_i \prod_{i<j} \exp\left(\frac{1}{2}(k_i \cdot k_j) G(z_i, z_j)\right)$$

where G is the propagator of the free field X^μ, satisfying

$$\Delta_z G(z, z') = 2\pi\delta^2(z - z').$$

- Taking the inverse Fourier transform, we find

$$G(z, z') = -2\pi \int \frac{d^2 q}{4\pi^2} \frac{e^{iq\cdot(z-z')}}{q^2} = \log(\mu|z - z'|)$$

where μ is an arbitrary constant which regulates the IR divergence. **(prove this)** Therefore we have

$$A = \kappa^{M-2} \int \prod_{i=1}^M d^2 z_i \prod_{i<j} |z_i - z_j|^{k_i \cdot k_j/2}.$$

- This integral diverges because we have not accounted for the $SL(2,\mathbb{C})$ gauge symmetry. This gauge symmetry is conventionally fixed by setting

$$z_1 = 0, \quad z_2 = 1, \quad z_3 = \infty.$$

The terms involving z_3 then multiply to $|z_3|^{-k_3^2/2} = |z_3|^{m^2/2}$ by momentum conservation, giving a constant which can be ignored. We hence have

$$A = \kappa^{M-2} \int \prod_{\ell=4}^M d^2 z_\ell \prod_{j=4}^M |z_j|^{k_1\cdot k_j/2} |1 - z_j|^{k_2\cdot k_j/2} \prod_{4\le i<j\le M} |z_i - z_j|^{k_i\cdot k_j/2}.$$

In particular, for the four-point function we have

$$A = \kappa^2 \int d^2 z_4 \, |z_4|^{k_1\cdot k_4/2} |1 - z_4|^{k_2\cdot k_4/2}.$$

- To derive these results more properly, we must ensure the gauge symmetries are not anomalous. One manifestation of $SL(2,\mathbb{C})$ gauge symmetry is that V is $SL(2,\mathbb{C})$ invariant. However, since

$$V = \int d^2z\, e^{ik\cdot X(z)}$$

  it appears that $e^{ik\cdot X}$ must be an operator of dimension two. Since X^μ is dimensionless, $e^{ik\cdot X}$ would classically have dimension zero, but it can receive an anomalous dimension. These can appear even for free field theories, as long as we work in $1+1$ dimensions. **(really?)**

- Concretely, for an operator Y of dimension p, we have

$$\langle Y(z)Y(0)\rangle \propto |z|^{-2p}.$$

  From our computations above, we already know that

$$\langle e^{ik\cdot X(z)}e^{-ik\cdot X(0)}\rangle = |z|^{-k^2/2}$$

  which indicates that the dimension is $k^2/4$. Since this must be equal to two, we have $k^2 = 8$, which gives the correct tachyon mass for the critical string. A similar computation for graviton scattering shows that the gravitons must be massless for consistency.

Next, we apply the same ideas to open strings.

- A tree-level open string diagram can be mapped onto a disc or a half-plane, with the vertex operators on the boundary. Therefore, the vertex operators take the form

$$V = \int d\tau\, \sqrt{h_{\tau\tau}}\, U(\tau), \quad U = We^{ik\cdot X}$$

  where τ is a parameter on the boundary of the worldsheet; invariance of V under conformal rescalings now require that U should have dimension one.

- For a spin zero particle we may take $W = 1$. Then it turns out that V has dimension $k^2/2$, and since the integral is only over the boundary, V must have dimension one. Then $k^2 = 2$, corresponding to the open string tachyon. For spin one, we may try $W = dX^\mu/d\tau$, which requires $k^2 = 0$.

- Mapping the worldsheet to the upper half-plane, the amplitude is

$$A(k_1, \ldots, k_M) = g^{M-2} \int dx_1 \ldots dx_M \left\langle \prod_{i=1}^{M} e^{ik_i\cdot X(x_i)} \right\rangle$$

  where g is the open string coupling constant. The residual gauge symmetries are conformal maps from the plane to itself, which preserve the real axis and are nonsingular at infinity; these take the form $\delta z = a + bz + cz^2$ where a, b, and c are real, and correspond to the group $SL(2,\mathbb{R})$.

- Note that conformal transformations can only produce cyclic permutations of the x_i, so a given string diagram fixes their cyclic order.

- Furthermore, it turns out that the endpoints of open strings may carry gauge charges. This may be motivated from the original purpose of string theory as a model for the strong interactions, where the string corresponds to a flux tube and quarks and antiquarks sit at the endpoints. (In fact, the first attempt to include fermions to describe baryons eventually led to the development of superstring theory.)

- One concrete realization of this is to consider a set of N space-filling D-branes. Then each endpoint of the open string can lie on any of the D-branes, so the state of the two endpoints is described by an $N \times N$ matrix.

- For gauge group $U(n)$, if we let the endpoints of the string transform in the fundamental and antifundamental, each vertex operator comes with a matrix $\lambda^i{}_j$ transforming in the adjoint. If the vertex operators come in the order $12 \ldots M$, we pick up a group theory factor of $\operatorname{tr}(\lambda_1 \ldots \lambda_M)$ as each antiquark is contracted with the next quark. This is called a Chan-Paton factor.

- We use the $SL(2, \mathbb{R})$ symmetry to fix

$$x_1 = 0, \quad x_{M-1} = 1, \quad x_M = \infty$$

so that the remaining x_i lie in $(0, 1)$. We will also need the Green's function $\tilde{G}(z, z')$ which satisfies the Neumann boundary conditions

$$\left. \frac{\partial G(x + iy, z')}{\partial y} \right|_{y=0} = 0.$$

This may be found by the method of images, which gives

$$\tilde{G}(z, z') = \log |z - z'| + \log |z - \overline{z}'|.$$

However, we are actually interested in the case where z and z' are both on the real axis, in which case $\tilde{G}(x, x') = 2 \log |x - x'|$, differing from G only by a factor of 2. This also accounts for the factor of 2 in the dimension of V above.

- Plugging everything in, for tachyon scattering we have the result

$$A = g^{M-2} \int_{0<x_2<\ldots<x_{M-2}<1} dx_2 \ldots dx_{M-2} \prod_{j=2}^{M-2} |x_j|^{k_1 \cdot k_j} |1-x_j|^{k_j \cdot k_{M-1}} \prod_{2 \leq \ell < m \leq M-2} |x_\ell - x_m|^{k_\ell \cdot k_m}.$$

This is the Koba–Nielsen M-particle generalization of the Veneziano amplitude. For $M = 4$ it simplifies to

$$A = g^2 \int_0^1 dx \, x^{k_1 \cdot k_2} (1 - x)^{k_2 \cdot k_3} = g^2 B \left(-\frac{s}{2} - 2, -\frac{t}{2} - 2 \right)$$

which is the Veneziano amplitude.

- It is also possible to have diagrams with both external open and closed strings. These can be evaluated by mapping the worldsheet to the upper half-plane, where the open string vertex operators are on the boundary and the closed string vertex operators in the bulk.

5.2 Vertex Operators

Next, we turn to a more detailed study of vertex operators, focusing on open strings.

- Consider a local operator at the string endpoint, $A(\tau) \equiv A(0, \tau)$. Since the string Hamiltonian is $L_0 - a$, we have
$$A(\tau) = e^{i\tau L_0} A(0) e^{-i\tau L_0}.$$

- We say $A(\tau)$ is a boundary primary operator with conformal dimension J if, under a change of variable $\tau \to \tau'$, we have
$$A'(\tau') = \left(\frac{d\tau}{d\tau'}\right)^J A(\tau).$$
This is equivalent to our definition of the conformal dimension in the previous section. **(show this)** Such operators transform "nicely" and are rather special; not every operator can be expanded as a linear combination of primary operators.

- The above condition is easier to handle at the infinitesimal level. For $\delta\tau = \epsilon(\tau)$, we have
$$\delta A(\tau) = -\epsilon \frac{dA}{d\tau} - JA \frac{d\epsilon}{d\tau}$$
by the definition above. The L_m generate transformations with $\epsilon = ie^{im\tau}$, so equivalently
$$[L_m, A(\tau)] = e^{im\tau} \left(-i \frac{d}{d\tau} + mJ\right) A(\tau).$$

- If $A(\tau)$ has an expansion in Fourier modes
$$A(\tau) = \sum_m A_m e^{-im\tau}$$
then this condition is equivalent to
$$[L_m, A_n] = (m(J-1) - n) A_{m+n}.$$

- It is straightforward to show that $X^\mu(\tau)$ has $J = 0$, as
$$X^\mu(\tau) = x^\mu + p^\mu \tau + i \sum_{n \neq 0} \frac{\alpha_n^\mu}{n} e^{-in\tau}$$
and
$$[L_m, X^\mu(\tau)] = i\alpha_m - i \sum_{n \neq 0} \alpha^\mu_{m+n} e^{-in\tau} = -i \sum_n \alpha_n^\mu e^{-i(n-m)\tau} = -ie^{im\tau} \frac{d}{d\tau} X^\mu(\tau).$$
Furthermore, the momentum operator $\dot{X}^\mu(\tau)$ has $J = 1$. However, $\ddot{X}^\mu(\tau)$ is not a boundary primary operator at all. It's easiest to show these using the expansion in Fourier modes.

- We have seen earlier that vertex operators should be primaries of dimension 1. Another way to see this makes sense is that if $A(\tau)$ has dimension 1 and $|\phi\rangle$ is a physical state, then $[L_m, A_0] = 0$, and hence $A_0|\phi\rangle$ is also a physical state. This is what we expect, as an emission or absorption should map a physical state to another.

Now we consider some examples of vertex operators.

- By the same arguments as before, we expect vertex operators to contain a factor of $e^{ik\cdot x(\tau)}$, where $x(\tau)$ is the string's center of mass position,
$$x^\mu(\tau) = x^\mu + p^\mu \tau.$$
However, this is not a local operator, so we instead consider $e^{ik\cdot X(\tau)}$. This exponential requires normal ordering, giving
$$V(k,\tau) = \,:\!e^{ik\cdot X(0,\tau)}\!:\, = \exp\left(k\cdot\sum_{n=1}^\infty \frac{\alpha_{-n}}{n}e^{in\tau}\right) e^{ik\cdot x(\tau)} \exp\left(-k\cdot\sum_{n=1}^\infty \frac{\alpha_n}{n}e^{-in\tau}\right).$$
This differs from the unordered expression by the divergent sum $\alpha' k^2 \sum 1/n$, by the BCH formula. As motivated earlier, this is the vertex operator for a tachyon.

- Naively, products of primary operators would be primary operators whose dimension is the sum of the product factors'. Indeed, this appears to be true if we naively use commutator identities above. The problem is that some products may be undefined without a subtraction or regularization scheme. In this case, we can remove the singularities by normal ordering, but typical normal-ordered products such as $:\!X^\mu(\tau)X_\mu(\tau)\!:$ aren't primaries at all.

- To compute the dimension of $V(k,\tau)$, we note that
$$[\alpha_p^\mu, e^{k\cdot\alpha_{-n}}] = p\delta_{p,n}k^\mu e^{k\cdot\alpha_{-n}}$$
which easily leads to
$$[L_m, e^{k\cdot\alpha_{-n}}] = \frac{1}{2}n\{k\cdot\alpha_{m-n}, e^{k\cdot\alpha_{-n}}\}$$
where the right-hand side contains an anticommutator, by using the commutator product rule to split apart L_m, and then using our previous identity.

- Now we need to evaluate $[L_m, V(k,\tau)]$. By expanding in a Taylor series, we have
$$[L_m, e^{ik\cdot X(0,\tau)}] = -ie^{im\tau}\frac{d}{d\tau}e^{ik\cdot X(0,\tau)}$$
which indicates that $e^{ik\cdot X(0,\tau)}$ has dimension zero, as we'd expect. When we add normal ordering, the right-hand side remains normal ordered, but the left-hand side is not. This causes the two sides to differ, introducing an extra term that yields the conformal dimension.

- Hence we can focus on the terms in $[L_m, V(k,\tau)]$ that are not normal ordered. This computation can be done by using the commutator product rule to split apart L_m, then using our previous identity. This produces m terms with creation operators on the far left, precisely
$$\frac{1}{2}\sum_{n=1}^m k\cdot\alpha_{m-n}e^{in\tau}V(k,\tau) = \left[\frac{1}{2}\sum_{n=1}^m k\cdot\alpha_{m-n}e^{in\tau}, V(k,\tau)\right] + \,:\!\frac{1}{2}\sum_{n=1}^m k\cdot\alpha_{m-n}e^{in\tau}V(k,\tau)\!:.$$
The second term contributes to $:\!(d/d\tau)e^{ik\cdot X(0,\tau)}\!:$ as in the non-normal ordered case, while the extra commutator term is
$$\left[\frac{1}{2}\sum_{n=1}^m k\cdot\alpha_{m-n}e^{in\tau}, V(k,\tau)\right] = \frac{1}{2}\sum_{n=1}^m k^2 e^{im\tau}V(k,\tau) = \frac{1}{2}mk^2e^{im\tau}V(k,\tau)$$
which yields a conformal dimension $J = k^2/2$, in accordance with our earlier computation.

- This is indeed a vertex operator when $k^2 = 2$, which is precisely the on-shell condition for the tachyon. Hence $V(k, \tau)$ is the vertex operator for the open string tachyon.

- Next, we can consider vertex operators for the photons at level one. One candidate which could have the right dimension is

$$V_\zeta(k, \tau) = \zeta \cdot \frac{dX}{d\tau} \exp(ik \cdot X), \quad k^2 = 0.$$

  This expression is already normal ordered, provided that $\zeta \cdot k = 0$, as $e^{ik\cdot X}$ is already normal ordered for $k^2 = 0$, and commuting components of $\zeta \cdot dX/d\tau$ through it gives terms proportional to $\zeta \cdot k$. Hence this operator is a primary of dimension 1. It is the vertex operator for a photon of polarization ζ.

- We may also define vertex operators for states of zero norm. For instance,

$$V(k, \tau) = -i\frac{d}{d\tau} \exp(ik \cdot X), \quad k^2 = 0$$

  is a vertex operator for the zero norm longitudinal photon state. The fact that this state is not physical is reflected in the fact that $V(k, \tau)$ is a τ derivative, so V_0 vanishes.

- Finally, we can consider vertex operators at level two. Candidates with the right dimension are

$$\zeta^{\mu\nu} \dot{X}_\mu \dot{X}_\nu :\!\exp(ik \cdot X)\!:, \quad k^2 = -1$$

  which is free of short-distance singularities if

$$k_\mu \zeta^{\mu\nu} = \eta_{\mu\nu} \zeta^{\mu\nu} = 0.$$

  This gives a total of $D(D-1)/2 - 1$ degrees of freedom, which is just enough to account for the symmetric traceless tensor of $SO(D-1)$ that appears at level two.

5.3 Bosonic Open Strings

Now we consider scattering amplitudes for open strings.

- Instead of using the path integral to evaluate a tree-level scattering amplitude, we will use a setup like that of worldline QFT, which we saw in the notes on Quantum Field Theory. Each diagram is a graph, and each edge of the graph corresponds to a worldline, which yields a propagator factor. One also writes down a factor for each vertex.

- We will take a slightly different route. Note that if we start with any tree-level open string diagram, with M external particles, we can deform it to the form of a single line on which $M - 2$ external string emissions are attached. This leads to the ansatz

$$A = g^{M-2} \langle \phi_1 | V_2(k_2) \Delta V_3(k_3) \ldots \Delta V_{M-1}(k_{M-1}) | \phi_M \rangle$$

  where Δ is the string propagator and the V_i are vertex operators. This manifestly contains poles due to on-shell propagators, as required by tree-level unitarity. However, symmetry under cyclic permutations of the external particles (i.e. duality) is not manifest.

- For scalar field theory, where $(\partial^2 + m^2)\phi = 0$, the propagator is simply $(\partial^2 + m^2)^{-1}$. For free open strings, the closest analogue of the first equation is $(L_0 - 1)|\phi\rangle = 0$, so we define the propagator to be
$$\Delta = (L_0 - 1)^{-1} = \int_0^1 z^{L_0 - 2}\, dz.$$

- For the vertex operators, it will be useful to define $z = e^{i\tau}$, so that
$$[L_m, V(k,z))] = \left(z^{m+1}\frac{d}{dz} + m z^m \right) V(k,z), \quad X^\mu(z) = x^\mu - ip^\mu \log z + i\sum_{n\neq 0} \frac{1}{n}\alpha_n^\mu z^{-n}.$$

- It is also useful to write the tachyon vertex operator as
$$V_0(k,z) \equiv \, :e^{ik\cdot X(z)}: \, = Z_0 W_0$$
where Z_0 is the zero mode operator,
$$Z_0 = e^{ik\cdot x + k\cdot p \log z} = e^{ik\cdot x} z^{k\cdot p + 1} = z^{k\cdot p - 1} e^{ik\cdot x}$$
and W_0 is the remaining factor,
$$W_0 = \exp\left(k \cdot \sum_{n=1}^\infty \frac{\alpha_n}{n} z^n \right) \exp\left(-k \cdot \sum_{n=1}^\infty \frac{\alpha_n}{n} z^{-n} \right).$$

- Similarly, the vertex operator for a massless photon is
$$V(\zeta, k, z) = \zeta \cdot \dot{X}(z) e^{ik\cdot X(z)}$$
where a dot indicates a derivative with respect to $\tau = -i\log z$.

- Next, we need to account for the external states. These can be constructed by a similar trick to that of quantum field theory. In quantum field theory, if we start with an arbitrary state with at least some overlap with the vacuum, then after evolution in imaginary time it will evolve to the vacuum state because everything else exponentially decays away. Similarly, to get, e.g. the lowest-energy state with a given conserved charge, we can start with an arbitrary state with the same charge and do the same.

- In string theory, the operators which give the desired states are precisely the vertex operators. That is, we have
$$|\Lambda, k\rangle = \lim_{\tau\to i\infty} e^{-i\tau} V_\Lambda(k,\tau)|0,0\rangle, \quad \langle \Lambda, k| = \lim_{\tau \to -i\infty} e^{i\tau}\langle 0,0| V_\Lambda(k,\tau).$$
Note that we do not have to actually Wick rotate to do this; the operator $V_\Lambda(k,\tau)$ with τ imaginary is defined in exactly the same way as $V_\Lambda(k,\tau)$ with τ real. This result is called the operator–state correspondence, and is an important principle in conformal field theory.

- Note that we are using the state $|0,0\rangle$. It is formally the vacuum of the 2D worldsheet QFT. However, it is not interpreted as a physical state; it is a tachyon with the momentum off-shell.

- We can check this explicitly for the tachyon vertex operator. By design, acting with W_0 on $|0,0\rangle$ will do nothing in the limit $\tau \to i\infty$, leaving

$$Z_0|0,0\rangle = e^{ik\cdot x} z^{k\cdot p+1}|0,0\rangle = z|0,k\rangle$$

since $e^{ik\cdot x}$ translates by k in momentum space. The factor of $e^{-i\tau}$ precisely cancels the extra z.

- Finally, the propagator can also be expressed as

$$\Delta = \int_0^\infty d\tau\, e^{-\tau(L_0-1)}$$

where the integrand generates evolution in imaginary time τ. The full amplitude is hence written in terms of infinite imaginary time evolution, with vertex operators V_1 and V_M inserted at infinity and the other V_i inserted at finite imaginary times; we integrate over these intermediate times.

- One can show that ghosts decouple as required, preserving tree-level unitarity. Another important formal task is to show that the amplitude has cyclic symmetry. Substituting in our results above and using the fact that $(L_0-1)|\phi_1\rangle = (L_0-1)|\phi_M\rangle = 0$, the amplitude becomes

$$A = g^{M-2}\int_0^1 \frac{dz_3 \dots dz_{M-1}}{z_3 \dots z_{M-1}} \langle \phi_1 | V(k_2, 1) V(k_3, z_3) \dots V(k_{M-1}, z_3 \dots z_{M-1}) | \phi_M \rangle.$$

The ordering can be made more explicit by changing variables to $y_i = z_3 z_4 \dots z_i$, giving

$$A = g^{M-2}\int_0^1 \left(\prod_{i=3}^{M-1} \theta(y_{i-1} - y_i) \frac{dy_i}{y_i} \right) \langle \phi_1 | V(k_2, y_2) \dots V(k_{M-1}, y_{M-1}) | \phi_M \rangle.$$

Now, applying the operator-state correspondence, the expectation value becomes

$$\lim_{y_1 \to \infty} \lim_{y_M \to 0} \frac{y_1}{y_M} \langle 0,0 | V(k_1, y_1) \dots V(k_M, y_M) | 0,0 \rangle.$$

- We must now evaluate an expectation value in the "unphysical" vacuum state $|0,0\rangle$. Now, string theory can be thought of as a two-dimensional field theory with an enormous symmetry group, and $SL(2,\mathbb{R})$ generated by L_{-1}, L_0, L_1 is the non-anomalous part of the Virasoro algebra. The vacuum is indeed $SL(2,\mathbb{R})$ invariant, $L_1|0,0\rangle = L_0|0,0\rangle = L_1|0,0\rangle = 0$. This is, of course, the same residual $SL(2,\mathbb{R})$ symmetry we encountered earlier.

- Infinitesimally, these $SL(2,\mathbb{R})$ transformations generate the change

$$y \to y' = y + \lambda_{-1} + \lambda_0 y + \lambda_1 y^2.$$

One might wonder why the 2×2 matrices in $SL(2,\mathbb{R})$ have a natural nonlinear action on the line; it turns out y transforms like v_1/v_2 for a vector $(v_1, v_2)^T$. The general transformation is

$$y \to y' = \frac{ay+b}{cy+d}$$

conventionally scaled so that $ad - bc = 1$. This is called a Mobius transformation.

- Mobius transformations give us the freedom to move the points at y_1 and y_M. As for the path integral, we must quotient out by the volume of $SL(2,\mathbb{R})$ to avoid overcounting. After some calculation, we find that the measure for the integral is

$$d\mu_M(y) = \delta(y_A - y_A^0)\delta(y_B - y_B^0)\delta(y_C - y_C^0)(y_A - y_B)(y_A - y_C)(y_B - y_C)\prod_{i=2}^{M}\theta(y_{i-1} - y_i)\prod_{j=1}^{M} dy_j$$

  where we have fixed three of the y_i's, namely y_A, y_B, and y_C, to specific values. This establishes cyclic symmetry in the integration variables y_i. Finally, cyclic symmetry in the product of vertex operators can also be shown, though we omit this calculation.

With all of this setup, we are now finally ready for some examples.

Example. The three-tachyon amplitude. This is

$$A = g\langle 0, -k_1|V_0(k_2, \tau)|0, k_3\rangle$$

where all momenta are directed inward and τ is an arbitrary constant which we do not need to integrate over, following the arguments above. To verify the value of τ does not matter, note that

$$A = g\langle 0, -k_1|e^{i\tau L_0}V_0(k_2, 0)e^{-i\tau L_0}|0, k_3\rangle = g\langle 0, -k_1|V_0(k_2, 0)|0, k_3\rangle$$

where we used the physical state condition $(L_0 - 1)|0, k\rangle = 0$. Plugging in the form of the tachyon vertex operator, the oscillator part W_0 does nothing, since there are no excitations to annihilate, so

$$A = g\langle 0, -k_1|Z_0|0, k_3\rangle = g\langle 0, -k_1|e^{ik_2\cdot x}|0, k_3\rangle = g\langle 0, -k_1|0, k_2 + k_3\rangle = g\delta(k_1 + k_2 + k_3).$$

The momentum-conserving delta function is generic and we will omit it below.

Example. The tachyon-tachyon-photon amplitude, where the photon has polarization ζ. Then

$$g\langle 0, -k_1|V(\zeta, k_2)|0, k_3\rangle = g\langle 0, -k_1|\zeta\cdot\dot{X}(1)V_0(k_2, 0)|0, k_3\rangle$$

where we simply set $\tau = 0$ in the vertex operator, following the discussion above. **(finish)**

6 String Compactification

6.1 Kaluza–Klein Theory

Compactifying the extra dimensions of string theory is one way to get a realistic result. In this section, we consider the simple example of one compact dimension and a closed bosonic string.

- We take the compact dimension to be
$$X^{25} \sim X^{25} + 2\pi R$$
so the background spacetime is $\mathbb{R}^{1,24} \times S^1$. This is a Kaluza–Klein compactification.

- We know that the closed string's massless fields are a traceless symmetric tensor, an antisymmetric tensor, and the dilaton. Following the original setup of Kaluza–Klein theory, we consider what happens to the symmetric tensor.

- In Einstein frame, the metric in $D = 26$ is decomposed into a $D = 25$ metric $\tilde{G}_{\mu\nu}$, a vector A_μ, and a scalar σ,
$$ds^2 = \tilde{G}_{\mu\nu} dX^\mu dX^\nu + e^{2\sigma}(dX^{25} + A_\mu dX^\mu)^2.$$
Here, Greek indices run over the noncompact dimensions $0, 1, \ldots, 24$ only.

- Diffeomorphisms affecting the noncompact dimensions lead to the expected diffeomorphism symmetry of $\tilde{G}_{\mu\nu}$ as a 25-dimensional metric, while diffeomorphisms affecting the compact direction, $\delta X^{25} = \Lambda(X^\mu)$ correspond to $\delta A_\mu = \partial_\mu \Lambda$. This indicates that A_μ is a 25-dimensional gauge field.

- One can show that the $D = 26$ Ricci scalar $\mathcal{R}^{(26)}$ is
$$\mathcal{R}^{(26)} = \mathcal{R} - 2e^{-\sigma}\nabla^2 e^\sigma - \frac{1}{4} e^{2\sigma} F_{\mu\nu}F^{\mu\nu}$$
where $\mathcal{R}$ is the $D = 25$ Ricci scalar. Therefore, the Einstein–Hilbert action is
$$S = \frac{1}{2\kappa^2}\int d^{26}X \sqrt{-\tilde{G}^{(26)}}\, \mathcal{R}^{(26)} = \frac{2\pi R}{2\kappa^2}\int d^{25}X \sqrt{-\tilde{G}}\, e^\sigma \left(\mathcal{R} - \frac{1}{4}e^{2\sigma}F_{\mu\nu}F^{\mu\nu} + \partial_\mu\sigma\partial^\mu\sigma\right)$$
where we have neglected any dependence on the compact dimension, since we only are interested in the low-energy effective action. This gives the interactions between the $D = 25$ fields.

- The $D = 25$ action is not quite of the Einstein–Hilbert form, because of the e^σ coefficient, but we may remove it by another change of frame. After this change, we see there is no potential dictating the vev of σ, which determines the physical size of the compactified dimension. In this particular case, a potential does arise as a one-loop effect.

- Such "moduli fields" are ubiquitous in compactifications of string theory and often lead to unwanted massless scalars, especially in supersymmetric theories. One must add a mechanism which gives them a potential and dynamically fixes the vev.

- We can also consider the Kaluza–Klein reduction of the other fields. The dilaton Φ simply remains a scalar, while the Kalb–Ramond field reduces to a 2-form $B_{\mu\nu}$ and a vector field $\tilde{A}_\mu = B_{\mu\,25}$. Note that in four dimensions, Kalb–Ramond fields have one degree of freedom; the corresponding particle is a massless scalar that behaves like an axion. Hence the low-energy physics consists of a metric, two $U(1)$ gauge fields, and three massless scalars.

More generally, we can consider what happens when the fields depend on the compact dimension.

- For simplicity, consider the scalar field Φ. Since it must be periodic in X^{25}, it may be expanded in a Fourier series,
$$\Phi(X^\mu, X^{25}) = \sum_n \Phi_n(X^\mu) e^{inX^{25}/R}.$$
Since it is a real scalar, $\Phi_n^* = \Phi_{-n}$.

- Ignoring the coupling to gravity for now, the kinetic terms are
$$\int d^{26}X \, \partial_\mu \Phi \partial^\mu \Phi + (\partial_{25}\Phi)^2 = 2\pi R \int d^{25}X \, \sum_n \partial_\mu \Phi_n \partial^\mu \Phi_{-n} + \frac{n^2}{R^2}|\Phi_n|^2.$$
This is an infinite set of scalar fields with masses
$$M_n^2 = \frac{n^2}{R^2}$$
where we found the $n = 0$ mode above. The heavier particles are called Kaluza–Klein (KK) modes.

- Under a diffeomorphism $X^{25} \to X^{25} + \Lambda(X^\mu)$, we have
$$\Phi_n \to \exp\left(\frac{in\Lambda}{R}\right) \Phi_n$$
which means that the n^{th} KK mode has charge n/R under the $U(1)$ gauge field A_μ. One conventionally uses the rescaled gauge field $A'_\mu = A_\mu/R$, so the charge of Φ_n is n.

6.2 Worldsheet Theory

In the previous section, we considered Kaluza–Klein theory from the perspective of spacetime fields. We now consider the worldsheet theory.

- First, the zero mode momentum in the X^{25} direction is now quantized,
$$p^{25} = \frac{n}{R}.$$
On the other hand, we can now allow more general boundary conditions,
$$X^{25}(\sigma + 2\pi) = X^{25}(\sigma) + 2\pi m R.$$
The winding number m tells us how many times the string winds around S^1.

- The mode expansion now takes the form
$$X^{25}(\sigma, \tau) = x^{25} + \frac{\alpha' n}{R}\tau + mR\sigma + \text{oscillator modes}.$$
It will be useful to introduce the quantities
$$p_L = \frac{n}{R} + \frac{mR}{\alpha'}, \quad p_R = \frac{n}{R} - \frac{mR}{\alpha'}$$

- Decomposing into left-moving and right-moving components,

$$X_L^{25}(\sigma^+) = \frac{1}{2}x^{25} + \frac{1}{2}\alpha' p_L \sigma^+ + i\sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \frac{\tilde{\alpha}_n^{25}}{n} e^{-in\sigma^+}$$

  and

$$X_R^{25}(\sigma^+) = \frac{1}{2}x^{25} + \frac{1}{2}\alpha' p_R \sigma^- + i\sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \frac{\alpha_n^{25}}{n} e^{-in\sigma^-}$$

  where $X^{25} = X_L^{25} + X_R^{25}$. The only difference versus the noncompact case is in the terms p_L and p_R. The mode expansions for all other dimensions remain unchanged.

- The $D = 25$ particle masses are given by $M^2 = -p_\mu p^\mu$, where we do not include the compact dimension. Using the usual constraint involving L_0 and $\tilde{L}_0$,

$$M^2 = p_L^2 + \frac{4}{\alpha'}(\tilde{N} - 1) = p_R^2 + \frac{4}{\alpha'}(N - 1)$$

  where N and $\tilde{N}$ are the levels in lightcone quantization, and we maintain $a = 1$.

- The "level matching" formula now says that

$$N - \tilde{N} = nm.$$

  Expanding out the mass formula gives

$$M^2 = \frac{n^2}{R^2} + \frac{m^2 R^2}{\alpha'^2} + \frac{2}{\alpha'}(N + \tilde{N} - 2).$$

  The first term is simply $(p^{25})^2$, while the second term is $(2\pi m R T)^2$, indicating the energy needed to stretch a string of tension T around the circle m times. The final term is the usual oscillator contribution.

- We may recover our initial results by looking at the massless states with $n = m = 0$, which have $N = \tilde{N} = 1$, and matching them up with the massless fields.
  - The states $\alpha_{-1}^\mu \tilde{\alpha}_{-1}^\nu |0, p\rangle$ form a rank two tensor of $SO(1, 24)$ and decompose into $G_{\mu\nu}$, $B_{\mu\nu}$, and Φ.
  - The states $\alpha_{-1}^\mu \tilde{\alpha}_{-1}^{25} |0, p\rangle$ and $\alpha_{-1}^{25} \tilde{\alpha}_{-1}^\mu |0, p\rangle$ correspond to two vectors of $SO(1, 24)$. The sum corresponds to A_μ, while the difference corresponds to $\tilde{A}_\mu$.
  - The state $\alpha_{-1}^{25} \tilde{\alpha}_{-1}^{25} |0, p\rangle$ is a scalar, corresponding to σ.

- One can show that the KK modes indeed have charge n under the A_μ field, by analyzing scattering amplitudes. Furthermore, one can show that modes with winding number m have charge m under the $\tilde{A}_\mu$ field.

If R takes on special values, then there are additional massless states.

- For example, consider states with $N = \tilde{N} = 0$. We get massless states with winding m (and hence $n = 0$) if

$$R^2 = \frac{4\alpha'}{m^2}.$$

We also get massless states with momentum n (and hence $m = 0$) if

$$R^2 = \frac{n^2\alpha'}{4}.$$

- The richest spectrum of massless states occurs when $R = \sqrt{\alpha'}$. In this case we get five families of massless states.
  - $N = \tilde{N} = 1$ with $m = n = 0$. These give the states described above.
  - $N = \tilde{N} = 0$ with $n = \pm 2$ and $m = 0$. These are KK modes of the tachyon field. They are spacetime scalars with charges $(\pm 2, 0)$ under the $U(1) \times U(1)$ gauge symmetry.
  - $N = \tilde{N} = 0$ with $n = 0$ and $m = \pm 2$. These are winding modes of the tachyon field. They are spacetime scalars with charges $(0, \pm 2)$.
  - $N = 1$ and $\tilde{N} = 0$ with $n = m = \pm 1$. These are two new spin 1 fields, $\alpha_{-1}^{\mu}|0, p\rangle$ with charges $(\pm 1, \pm 1)$.
  - $N = 0$ and $\tilde{N} = 1$ with $n = -m = \pm 1$. These are two more spin 1 fields, $\tilde{\alpha}_{-1}^{\mu}|0, p\rangle$ with charges $(\pm 1, \mp 1)$.