

NIKHEF/2007-020  
IMAFF/FM-07/02

# Introduction to String Theory\*

---

\* Based on lectures given at the Radboud Universiteit, Nijmegen.

These notes follow rather closely the course given in the fall semester of 1999. They are based to a large extent on the books by Green, Schwarz and Witten, [1], the book by Polchinski [2], the book of Lüst and Theisen [3], and the review article [5]. I refer to these sources for further references. A very useful introductory text that appeared after the original version of these notes was written is the book by Zwiebach [4].

## 1 Current problems in particle physics

Before getting into string theory, it may be good to know to which questions we hope it will provide answers.

At present, we know about four fundamental interactions. Three of them, electromagnetism, the weak and the strong interaction are very successfully described by a theory called modestly the “Standard Model”. The standard model seems to be in very good shape experimentally. All quantities that we can compute reliably, and that we can measure, are in agreement with each other. There is only one thing that we are quite sure is missing. The standard model predicts (at least) one additional particle needed for its consistency, namely the Higgs boson. If it is found, which indeed should happen within the next two decades, there is really no question left within the standard model that requires an answer.

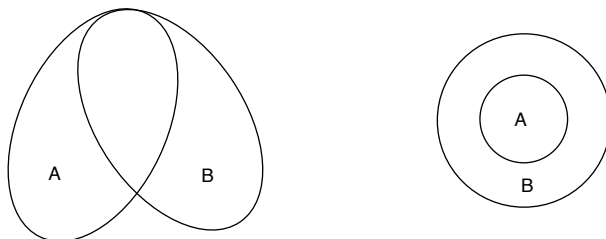
Certainly there are many questions which we may *hope* to see answered. We may wish to compute the quark and lepton masses, explain hierarchies of scales, or understand the general structure of the model, but without internal inconsistencies or conflicts with experiment we may have no other choice than to accept the model as it is. Many ideas have been proposed to address some of these questions. Examples are “Grand Unification” or “Supersymmetry”. Usually they improve the Standard Model in some ways, while making other problems much worse. Supersymmetry, for example, improves the hierarchy problem, but at the price of a drastic increase in the number of parameters. One can even imagine extensions of the Standard Model that – from our limited point of view – are not improvements at all. There may be additional particles with extremely weak interactions, that do not seem to do anything “useful”. If we are lucky we may get experimental information about what, if anything, lies beyond the Standard Model, but it seems at least as plausible that we will never know.

The fourth interaction, gravity, is in a less good shape theoretically. Classically, it is described by Einstein’s theory of General Relativity. Experimentally this theory in impressive agreement with experiment, but when one tries to quantize it one gets into serious difficulty. This is a fundamental problem, which *does* require a solution. Unlike the standard model, where the remaining problems are mainly aesthetic, quantum gravity has serious internal inconsistencies (more about this below). The strongest point in favor of string theory is that it seems to solve these problems. At the moment it is the only serious candidate for a consistent theory of quantum gravity. This does not mean that there is a proof that all problems of quantum gravity are solved, nor that it is known that no other possibilities exist.

This fact would already be enough to study String Theory seriously. But there is more. String Theory does not just contain gravity, it comes inevitably with a large number of other particles and interactions. These particles and interactions have the same features as the Standard Model. Typically there are quark and lepton-like particles, coupling to gauge bosons. Furthermore there are usually some scalars. The interactions are often *chiral*, which means that left- and right-handed particles couple in different ways. This feature is observed in the weak interactions. In fact any observed property of the Standard Model can be identified within String Theory. Among the possibilities are also Grand Unification or Supersymmetry.

It was once hoped – although for no good reasons – that string theory would uniquely determine the standard model. Meanwhile we know enough to be able to say that this is not true. In fact, there are so many possibilities that at the moment nobody is able to prove whether the Standard Model is among the possibilities. Nevertheless, if String Theory is correct as a theory of gravity, there is really no other option than that it is also a theory of all other interactions. According to our present understanding, String Theory is far more robust than the usual field-theoretic description of particles and interactions. We cannot take a String Theory description of gravity and simply add the Standard Model to it. One cannot add particles or interactions to a given String Theory, or remove particles one does not like.

This seems almost in contradiction with the fact that String Theory is far from unique. The difference between the String Theory and Field Theory description of nature is illustrated by the following picture



The first picture shows two possible string theories, the second two possible field theories. The idea is that in field theory we can extend theory A to theory B without throwing any part of theory A away. In String Theory we cannot do that. We can modify theory A to theory B by throwing part of theory A away and adding something new in its place. Suppose we know experimentally that we need everything in theory A. Then String theory would actually predict that there are no other particles or interactions. If an experiment does find something not predicted by theory A, we would be forced to give up String Theory altogether. On the other hand, if we are in the situation of the second picture, we could simply add something to our theory.

For this reason one sometimes uses the phrase “Theory of Everything” when talking about string theory. Not everybody thinks it is a good idea to use this phrase.

## 1.1 Problems of quantum gravity

One of the first problems one encounters when one attempts to quantize a field theory is that certain calculations seem to lead to infinities. For example, loop diagrams may lead to integrals of the following type

$$\int d^4p \frac{1}{(p^2 + m^2)^2} \quad (1.1)$$

where  $p^\mu$  is a four-momentum, and  $m$  some mass. It is easy to see that this integral diverges for large  $p$ .

If one thinks for a moment about what one is doing here, it seems rather silly to integrate over arbitrarily large momenta. Large momenta correspond to short distances, and by integrating to infinity, we are pretending that we understand physics to arbitrarily short distances. This is a ridiculous assumption. Perhaps space-time is discrete at extremely small distances, so that there is a maximum momentum, the inverse of the smallest distance. Perhaps something else changes at short distances. If anything changes at short distances, it may happen that our integral is actually finite, proportional to  $\log(\Lambda)$ , where  $\Lambda$  is some large, but finite, momentum scale. The trouble is that without detailed knowledge about the short distance physics we cannot compute the integral, all we can do is introduce a new parameter equal to the value of the integral. However, in quantum field theory there are integrals of this type for any loop diagram, and if we cannot compute any of them we would be forced to introduce an infinity of new parameters. To know the theory completely we would have to measure all these parameters, and that is obviously impossible.

However, some theories are nicely behaved. In such theories the unknown parameters appear only in (infinite) linear combinations such as

$$m + m_1 + m_2 + m_3 \dots \quad (1.2)$$

where  $m$  is a parameter of the original theory (for example a mass), while  $m_1 \dots$  are the unknown parameters from loop diagrams. If all physical quantities depend only on this particular combination there is no need to determine them individually. We simply give the whole sum a new name,  $m_R$ , and if we know the value of  $m_R$  we know enough. The net result of this procedure is that we end up with an equal number of parameters as we started with. If a theory has this property, we say that it is *renormalizable*. The Standard Model has this property, and for this reason we regard it as a consistent quantum theory.

But gravity does not have this property. In the case of gravity the loop corrections lead to an infinity of new parameters that cannot be absorbed into the ones we started with (there is just one, namely Newton's constant). Strictly speaking that means that any physical quantity we compute in quantum gravity depends on an infinite number of parameters, and hence the predictive power of the theory is zero. This is of course an exaggeration. If true, it would imply that we cannot even believe gravity in its classical limit, where it is so successful.

It turns out that for any reasonable value of the unknown parameters, their effect is completely negligible at observable energies.

The coupling constant of gravity is Newton's constant  $G_N$ , defined through

$$F = G_N \frac{m_1 m_2}{r^2} \quad (1.3)$$

Using  $\hbar$  and  $c$  we can construct from it a mass, the Planck mass

$$M_P = \sqrt{\frac{\hbar c}{G_N}} \approx 2 \times 10^{-8} \text{ kg} \quad (1.4)$$

In high energy physics it is customary to multiply masses with  $c^2$  and express them in units of energy, for which the GeV is used. Then we get

$$M_P = \sqrt{\frac{\hbar c^5}{G_N}} = 1.22 \times 10^{19} \text{ GeV} \quad (1.5)$$

Using other combinations of  $\hbar$  and  $c$  we can convert this to an inverse length. This length, the Planck length is  $1.6 \times 10^{-33}$  cm. Usually we use units where  $\hbar = c = 1$ , so that the distinction between energy, mass and inverse length disappears.



The gravitational potential can be obtained from graviton exchange between the two particles. The first quantum correction is due to two-graviton exchange, and is proportional to  $G_N^2$ . The ratio of the two and the one-graviton term must therefore be proportional to  $G_N$ . Since the ratio must be dimensionless, it must be of the form  $G_N E^2$ , where  $E$  is some energy of the process under consideration. Hence we see that the first correction must be proportional to  $(E/M_P)^2$ , which is extremely small as long as  $E$  is much smaller than  $10^{19}$  GeV, and as long as the unknown coefficient is of order 1. This is far beyond anything we can ever hope to achieve with accelerators.

So the problem of finding a quantum theory of gravity is a purely theoretical one. There does not appear to be any chance of observing its effects directly.

There are other problems associated with quantum mechanics and gravity. Some very interesting problems arise in the discussion of black holes. One of the most intriguing features of a black hole is the existence of a horizon. Outside observers cannot get information about what happens inside that horizon. They can allow themselves to fall into the black hole, inside the horizon, but then they cannot communicate with the outside world anymore. Objects thrown into the black hole cannot be recovered. This seems to clash

with unitarity of the quantum mechanical  $S$ -matrix, which states that the total probability is conserved. Indeed, it has been proposed that the existence of black holes would force us to give up unitarity of the  $S$ -matrix. This is called the information-loss problem. It is made more severe because of black hole evaporation. Given long enough time, a black hole radiates away all its energy through Hawking radiation until it evaporates completely. Then apparently all the information that went into it disappears forever.

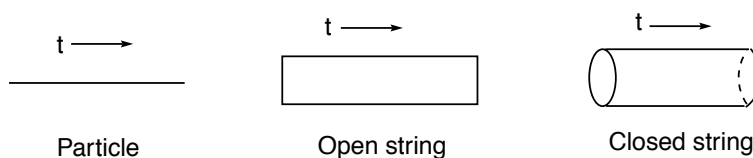
Another related problem is that of the black hole entropy. Bekenstein and Hawking have shown that a black hole has thermodynamic properties: it has black body radiation with a certain temperature, and it behaves as if it has a certain entropy, proportional to the area of the horizon. Normally one associates entropy with the logarithm of the number of states of an object, but such a notion was not available here. Presumably understanding the counting of states of a black hole in relation to its entropy will also tell us how it can store information.

String theory has made interesting contributions to both problems. First of all string theory *does* indeed change the short distance physics. It has been shown that in certain string theories the unknown or infinite corrections are in fact finite and calculable. Furthermore it has been possible to obtain a description of the states of special black holes (in special limits) so that the correct entropy could be obtained.

All of this is still very vague, so it is time to take a closer look.

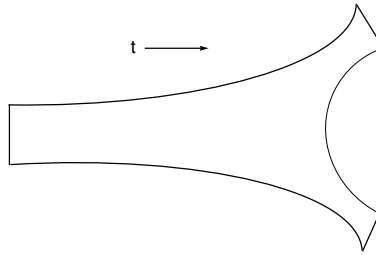
## 1.2 String diagrams

In our present description of fundamental physics the basic objects are particles – leptons, quarks, gauge bosons etc. The starting point of String Theory is the assumption that the basic objects have a one-dimensional extension, that they are like small pieces of rope. This principle allows two kinds of fundamental objects: open and closed strings. When such objects move through space they sweep out ribbons or cylinders.

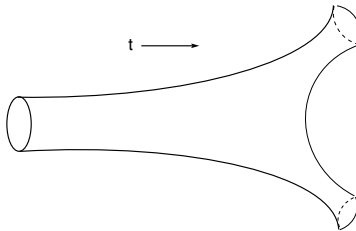


There is still one more distinction to be made: strings may or may not have a definite orientation. Depending on this we distinguish oriented and unoriented strings. All together there are at this point then four types of strings: open unoriented, open oriented, closed unoriented and closed oriented. Oriented strings must keep their orientation when they interact.

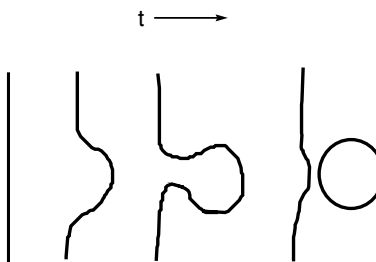
Strings interact by splitting and joining. We can represent this by plotting the strings as a function of time. For example the following picture represents an open string splitting



It is also easy to draw the splitting of closed strings



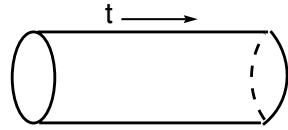
A more interesting process is that of an open string decaying into an open and a closed string.



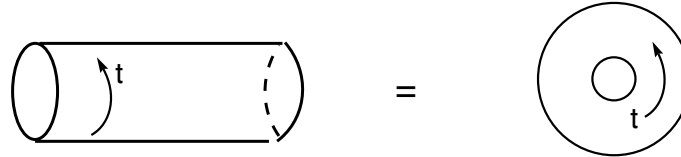
This shows that open strings necessarily imply closed strings. The other way around this is not true. It is impossible to draw a diagram where a closed string emits an open one.

In general a string diagram can be any two-dimensional surface. Such a surface can be interpreted in various ways by choosing a time coordinate on it. Then one can give a string interpretation to the interactions by slicing the surface with planes orthogonal to the time direction. For example, the previous diagram is obtained by time-slicing in a suitable way a piece of a plane to which a cylinder is attached.

On a cylinder there are two obvious ways of choosing a time coordinate



Closed string  
propagation



Open string loop

One corresponds to the propagation of a closed string, the other to an open string making a closed loop.

One of the most important features of string theory is that to each such diagram there corresponds a certain amplitude. We may try to interpret that amplitude anyway we like, but there is only one amplitude. For example, to compute the scattering amplitude for two closed strings to two closed strings one begins by drawing all valid diagrams with two initial and two final closed strings, propagating from  $-\infty$  or to  $+\infty$  respectively.

There is a precise procedure that assigns to each such diagram a number, the transition amplitude. This procedure is based on *path integrals*. These are infinite dimensional integrals over all surfaces with certain initial and final states one is interested in. The surfaces are weighted by an exponential; very roughly the time evolution from initial to final states is then given by

$$\langle \text{out} | U | \text{in} \rangle = \int_{\text{Surfaces}} e^{iS/\hbar} , \quad (1.6)$$

where  $S$  is the (classical) action. In the simplest case  $S$  is simply the area of the surface. Of course such infinite dimensional integrals require a much more detailed discussion.

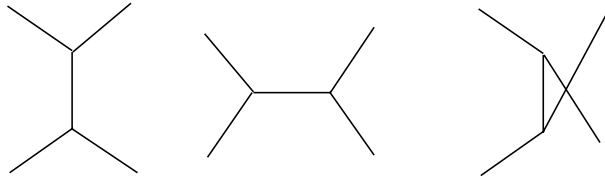
Which diagrams are allowed depends on the kind of string theory one considers. In the simplest case, oriented closed strings, the allowed diagrams would be the sphere, the torus, the double torus, etc, each with four infinite tubes attached. Two diagrams are considered equivalent if we can deform them into each other in a continuous way, without breaking anything. It is a basic result in mathematics that all such surfaces are characterized topologically by the number of handles (the *genus*), *i.e.* a torus is a sphere with one handle, etc. In the case of oriented open and closed strings one must allow the surface to have holes, corresponding to the ends of open strings. In the case of unoriented closed strings one cannot have holes, but one must allow closed surfaces that do not have a definite orientation, such as the Klein bottle (a surface has a well-defined orientation if one can consistently define a normal vector in a continuous way, on all points



on the surface). Finally, in the case of unoriented open strings one must also allow open, unoriented surfaces. The best known example is the Moebius strip.

External lines are either tubes glued to the surface itself, or strips glued to the boundaries. The latter correspond of course to external open strings.

All of this may be compared to the Feynman diagram expansion of quantum field theory. One obvious difference is that in that case one obtains a strongly increasing number of diagrams in each order. This is already clear if we look at a theory with a three-point interaction. Also in quantum field theory there is a precise calculational procedure for computing an amplitude belonging to such a diagram.



At lowest order, there are three diagrams, whereas in oriented closed string theory there is just one. At one loop order, there is still only one string diagram, but many more field theory diagrams.

But there is a more important difference between the two situations. In field theory diagrams there are special points corresponding to the splitting of a particle into two (or more) particles. This implies that we have to give additional information about the strength of the interaction at these points. In a string theory there are no special points. The surfaces are smooth, and every point is equivalent to any other point. If we time-slice a surface, there may appear to be special points where the string splits. But this depends on how we define the equal-time slices. This is not a Lorentz-invariant concept, and hence two different observers will not agree on a point on the two-dimensional surface which is special. Hence there are no special points.

One consequence is that by specifying a free string theory, we have already specified the interactions. As we will see, a string theory is specified by giving an action. The action is always some two-dimensional Lagrangian density, integrated over the surface. In the simplest cases the action is nothing but the area of the surface. The action does not know about the topology of the surface. Computing different string diagrams amounts then to computing the integrals of the same integrand over different surfaces.

The fact that string theory has no ultra-violet (short distance) divergences is also related to this fact. The divergences in quantum field theory can be traced back to the existence of point-like interactions, and can in fact be removed by “smearing out” the interactions over a finite region. But in string theory there are no point-like interactions to begin with.

String theory may nevertheless have divergences, but they are of a different kind, and can be avoided in certain cases.

## 2 Bosonic string action

The simplest string theory is the bosonic string. Although simple, it has serious defects: its spectrum contains no space-time fermions, but it does contain a particle with imaginary mass, a tachyon. It is therefore not a candidate for a theory of nature, but it is very useful as a kind of toy-model.

The description of the bosonic string is modelled on that of a relativistic point particle. We could describe the motion of such a particle by giving its position as a function of time,  $\vec{x}(t)$ . However, this does not allow a Lorentz-invariant description. Instead one could choose to specify  $X^\mu(t)$  with the condition  $X^0(t) = t$ , but this is still not Lorentz invariant. The solution is to introduce a separate variable  $\tau$  to parametrize the world line of the particle. Each point on the world line is then specified by a set of  $D$  functions  $X^\mu(\tau)$ . We work here in  $D$  space-time dimensions. We will set the speed of light,  $c$ , equal to 1.

For the action of such a relativistic point particle we may take the length of the world line, *i.e.*

$$S = -m \int d\tau \sqrt{-\frac{dX^\mu}{d\tau} \frac{dX_\mu}{d\tau}} \quad (2.1)$$

The minus sign in the square root reflects our choice of metric, namely  $(-, +, +, +)$ . For example a particle at rest at the origin can be described by  $X^0 = \text{const} \times \tau$ ,  $X^i = 0$ , which yields a positive square root argument. An important feature of this action is that the answer does not depend on how we choose to parametrize the world line, *i.e.* if we replace  $\tau$  to any (monotonic) function  $f(\tau)$ . This is important, because it is the world line itself that matters, not how we choose to parametrize it. When one considers the equations of motion it turns out that  $m$  is the mass of the particle.

The canonical momentum to  $X^\mu$  is

$$P_\mu = \frac{\delta \mathcal{L}}{\delta \dot{X}^\mu} = m \frac{\dot{X}_\mu}{\sqrt{-\dot{X}^2}}$$

The equation of motion says that this momentum is conserved as a function of  $\tau$  (*i.e.* along the world line) and furthermore we see that the relation  $P^\mu P_\mu + m^2 = 0$  (known as the “mass shell condition”) holds.

### 2.1 The Nambu-Goto action

As already mentioned above, we want to apply the same idea to String Theory. Now we generalize from lines to surfaces, which are now called world sheets. To specify the embedding of a world sheet in space-time we give  $D$  functions  $X^\mu(\sigma, \tau)$ . Now there are two variables, one to describe the direction along the string ( $\sigma$ ), and one to parametrize a time-like direction. The action is proportional to the surface-area of the world sheet, which can be written as

$$S[X] = -\frac{1}{2\pi\alpha'} \int d\sigma d\tau \sqrt{-\det h_{\alpha\beta}} \quad (2.2)$$

where  $h_{\alpha\beta}$  is a two-by-two matrix defined as

$$h_{\alpha\beta} = \partial_\alpha X^\mu \partial_\beta X_\mu, \quad (2.3)$$

where  $\partial_\alpha$  means  $\frac{\partial}{\partial\sigma^\alpha}$ , and where instead of  $\sigma$  and  $\tau$  we use variables  $\sigma^\alpha$ ,  $\alpha = 0, 1$ , with  $\sigma^0 = \tau$  and  $\sigma^1 = \sigma$ . The sign of the determinant is determined by the space-time metric. A string at rest could be given by  $X^0 = \tau$ ,  $X^i = X^i(\sigma)$ . Then  $h_{00} = -1$  and  $h_{11} = (\partial_\sigma X^i)^2$ ,  $h_{01} = h_{10} = 0$ , so that the determinant is negative. Then it will stay negative under continuous deformations (since  $\det h = 0$  implies the manifold is singular at that point). The action (2.2) is called the Nambu-Goto action.

## 2.2 The free boson action

It has the unpleasant property that it contains a square root. However, by introducing an extra set of variables  $\gamma_{\alpha\beta}(\sigma, \tau)$  we can get a different action which is classically equivalent\*

$$S[X, \gamma] = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \sqrt{-\det \gamma} \sum_{\alpha\beta} \hat{\gamma}_{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X_\mu \quad (2.4)$$

Here  $\gamma_{\alpha\beta}$  is a symmetric, non-singular two-by-two matrix that may depend on  $\sigma$  and  $\tau$ . By  $\hat{\gamma}_{\alpha\beta}$  we mean the inverse of  $\gamma_{\alpha\beta}$ . The determinant of  $\gamma_{\alpha\beta}$  is assumed to be negative here; this will be justified in a moment.

To see the classical equivalence, note that  $\gamma_{\alpha\beta}$  is not a dynamical degree of freedom: it occurs without any derivatives. Hence we can eliminate it by using the equations of motion. The latter are defined by requiring that the variation of the action with respect to  $\gamma_{\alpha\beta}$  vanishes:

$$\delta_\gamma S = 0 \quad (2.5)$$

The action depends on  $\gamma_{\alpha\beta}$  via the determinant and the inverse. The variation of the inverse is computed as follows:

$$\hat{\gamma}\gamma = \mathbf{1} \rightarrow \delta[\hat{\gamma}\gamma] = 0 \quad (2.6)$$

Therefore, using the chain rule

$$\delta[\hat{\gamma}]\gamma + \hat{\gamma}\delta\gamma = 0 \quad (2.7)$$

Multiplying with  $\hat{\gamma}$  we get then

$$\delta[\hat{\gamma}] = -\hat{\gamma}\delta[\gamma]\hat{\gamma} \quad (2.8)$$

To vary  $\sqrt{-\det \gamma}$  we proceed as follows

$$\delta[\sqrt{-\det \gamma}] = \delta \left[ \exp \left( \text{Tr} \left( \frac{1}{2} \log(-\gamma) \right) \right) \right] = \frac{1}{2} (\sqrt{-\det \gamma}) \delta[\text{Tr} \log(-\gamma)] \quad (2.9)$$

---

\* As indicated,  $\gamma_{\alpha\beta}(\sigma, \tau)$  may depend on  $\sigma$  and  $\tau$ , but to simplify the notation we drop the arguments as long as they are not essential.

The variation of the logarithm of a matrix can be computed in the same way as  $\partial_x \log(x)$ :

$$\begin{aligned}
\delta[\log(-\gamma)] &\equiv \log(-(\gamma + \delta\gamma)) - \log(-\gamma) \\
&= \log((\gamma + \delta\gamma)\hat{\gamma}) \\
&= \log(\mathbf{1} + \delta[\gamma]\hat{\gamma}) \\
&= \delta[\gamma]\hat{\gamma}
\end{aligned} \tag{2.10}$$

The final result is then

$$\delta[\sqrt{-\det \gamma}] = \frac{1}{2} \sqrt{-\det \gamma} \operatorname{Tr} \delta\gamma\hat{\gamma} \tag{2.11}$$

The variation of the action is therefore

$$\delta_\gamma S = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \sqrt{-\det \gamma} \left[ \frac{1}{2} \hat{\gamma}_{\gamma\delta} \hat{\gamma}_{\alpha\beta} h_{\alpha\beta} - \hat{\gamma}_{\alpha\gamma} \hat{\gamma}_{\delta\beta} h_{\alpha\beta} \right] (\delta\gamma)_{\gamma\delta} , \tag{2.12}$$

with implicit summation over all repeated indices. This can only vanish for arbitrary  $(\delta\gamma)_{\gamma\delta}$  if the terms between square brackets add up to zero; this yields the matrix identity

$$\hat{\gamma} h \hat{\gamma} = \frac{1}{2} \hat{\gamma} \operatorname{Tr} (\hat{\gamma} h) \tag{2.13}$$

Multiplying on both sides with  $\gamma$  gives

$$h = \frac{1}{2} \gamma \operatorname{Tr} (\hat{\gamma} h) \equiv \Lambda(\sigma, \tau) \gamma \tag{2.14}$$

This allows us to express  $\gamma$  in terms of  $h$ , up to some function  $\Lambda$ . When we substitute this into the action  $S[X, \gamma]$  the function  $\Lambda$  drops out, and we do indeed obtain the Nambu-Goto action  $S[X]$ .

It is customary to introduce upper and lower indices (analogous to the space-time Lorentz indices) also for the parameters  $\sigma^\alpha$ . Then one defines

$$\gamma^{\alpha\beta} \equiv \hat{\gamma}_{\alpha\beta} , \tag{2.15}$$

and we use this matrix to raise indices, and  $\gamma_{\alpha\beta}$  to lower them. Since these matrices are each other's inverse, this makes sense. The action reads now

$$S[X, \gamma] = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \sqrt{-\gamma} \sum_{\alpha\beta} \gamma^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X_\mu , \tag{2.16}$$

where  $\gamma$  denotes the determinant of  $\gamma_{\alpha\beta}$ . From the relation between  $\gamma$  and  $h$  we see that the natural sign of the determinant is negative. This action can be recognized as that of  $D$  free scalar fields  $X^\mu$ ,  $\mu = 1, \dots, D$  coupling to a two-dimensional metric  $\gamma_{\alpha\beta}$ .

## 2.3 Symmetries

The bosonic string action  $S[X, \gamma]$  has the following symmetries

- Poincaré invariance in  $D$  dimensions. This transformation acts only on  $X$ , not on  $\gamma_{\alpha\beta}$

$$X'^{\mu}(\sigma, \tau) = \Lambda^{\mu}_{\nu} X^{\nu}(\sigma, \tau) + a^{\mu} \quad (2.17)$$

- Reparametrization invariance in 2 dimensions. This is taken over from the reparametrization invariance of the Nambu-Goto action, and must be present for the same reason. It is also called diffeomorphism invariance, and is in fact nothing but a general coordinate transformation in two dimensions. Such a transformation is performed as follows. We can introduce a new function  $X'^{\mu}(\sigma', \tau')$  which describes the space-time position of every point on the string in terms of a new parametrization  $\sigma'(\sigma, \tau)$  and  $\tau'(\sigma, \tau)$ . By definition this new function is therefore  $X'^{\mu}(\sigma', \tau') = X^{\mu}(\sigma, \tau)$ . Now we write the action entirely in terms of the new variables  $X'$ ,  $\sigma'$  and  $\tau'$ . For most variables, this simply implies that we replace them in (2.4) by the primed variables. In particular,  $d\sigma d\tau$  is replaced by  $d\sigma' d\tau'$  and  $\partial_{\alpha}$  by  $\partial'_{\alpha}$ . Any action could be rewritten in this manner, but this is only a symmetry if we re-obtain the original action when we express the primed variables back in terms of the original ones. In this particular, changing back to the original variables introduces two potential problems: the change of integration measure when expressing  $\sigma'_{\alpha}$  in terms of  $\sigma_{\alpha}$ , and the change from  $\partial'_{\alpha}$  back to  $\partial_{\alpha}$ . Even without prior knowledge of general relativity, one realizes that in order for this to be a symmetry we have to transform  $\gamma$  as well. The complete transformation is then:

$$\begin{aligned} X'^{\mu}(\sigma', \tau') &= X^{\mu}(\sigma, \tau) \\ \gamma'_{\gamma\delta}(\sigma', \tau') &= \frac{\partial\sigma^{\alpha}}{\partial\sigma'^{\gamma}} \frac{\partial\sigma^{\beta}}{\partial\sigma'^{\delta}} \gamma_{\alpha\beta}(\sigma, \tau), \end{aligned} \quad (2.18)$$

where the new coordinates  $\sigma'$  and  $\tau'$  are functions of the old ones  $\tau$  and  $\sigma$ . Note that we are free to change  $\gamma$  as we wish. If we manage to find a transformation that gives us back the original action, then we can call the combined transformation of  $X$  and  $\gamma$  a symmetry.

- Weyl invariance. This is a symmetry not seen in the Nambu-Goto action, and is a consequence of the free function  $\Lambda$  we found above. Indeed, it is easy to see that the action is invariant under

$$\gamma'_{\alpha\beta}(\sigma, \tau) = \Lambda(\sigma, \tau) \gamma_{\alpha\beta}(\sigma, \tau) \quad (2.19)$$

without changing  $X^{\mu}$ .

The last two symmetries are redundancies of the two-dimensional theory on the world sheet. This means that the two-dimensional action has fewer variables than it seems to have. This is completely analogous to gauge symmetry in electrodynamics: the invariance

under the gauge transformation  $A_\mu \rightarrow A_\mu + \partial_\mu \Lambda$  implies that one degree of freedom of  $A_\mu$  can be removed (in addition,  $A_0$  is non-dynamical, so that the number of physical degrees of freedom of the photon in  $D$  dimensions is reduced from  $D$  to  $D - 2$ ). As is also familiar from electrodynamics, one may fix the redundancy by choosing a gauge. This has to be done in a clever way in order to make quantization of the theory manageable.

## 2.4 Conformal gauge

One way to fix the gauge in our case is to fix the two-dimensional metric  $\gamma_{\alpha\beta}$ . It is instructive to consider the system in an arbitrary number  $N$  of world sheet dimensions. The action and the first two symmetries then remain exactly the same, but the Weyl symmetry is broken:  $\gamma^{\alpha\beta}$  transforms with  $\Lambda^{-1}$ , and  $\sqrt{-\gamma}$  transforms with a factor  $\Lambda^{N/2}$ . They cancel only for  $N = 2$ . The number of components of the metric is  $\frac{1}{2}N(N + 1)$ . Using reparametrization invariance one can put  $N$  components of the metric to zero (at least locally, in a finite neighborhood of a point), reducing the number of components to  $\frac{1}{2}N(N - 1)$ . In two dimensions the extra Weyl invariance then allows us to remove one more degree of freedom, so that finally none is left! This illustrates why two dimensions is special, and why string theory is much easier than theories of membranes, whose world-sheets are  $N$ -dimensional, with  $N > 2$ .

We can use this freedom to transform the metric to the flat, Minkowski metric  $\eta_{ab} = \text{diag}(-1, 1)$  (with a  $-1$  in the  $\tau$  direction). This is called *conformal gauge*. Then the action looks even simpler\*

$$S[X] = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \eta^{ab} \partial_a X^\mu \partial_b X_\mu \quad (2.20)$$

This is the simplest action one can write down in two-dimensional field theory. It is the action of a free boson. More precisely, of  $D$  free bosons, because the index  $\mu$  on the bosons is from the two-dimensional point of view nothing other than a multiplicity. The action  $S[X, \gamma]$  we had earlier is in fact the action of a free boson coupled to two-dimensional gravity.

## 2.5 World sheet versus Space-time

The two-dimensional metric  $\gamma^{\alpha\beta}$  (the world sheet metric) should not be confused with the  $D$ -dimensional metric  $g^{\mu\nu}$ . We can make that metric manifest in the action by writing it as

$$S[X] = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \eta^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X^\nu g_{\mu\nu} \quad (2.21)$$

Here  $g_{\mu\nu}$  may depend on the space-time position, but not on the world-sheet coordinates  $\sigma^a$ , like  $\gamma$  and  $X$ . Hence from the two-dimensional point of view  $\gamma$  and  $X$  are fields, but  $g_{\mu\nu}$  is not. If we are considering a string theory embedded in flat, Minkowski space-time, we must choose  $g_{\mu\nu} = \eta_{\mu\nu}$ ,  $\eta = \text{diag}(-1, 1, \dots, 1)$ . One may also consider a string

---

\* Our convention is to use indices  $\alpha, \beta, \dots$  for world-sheet indices in arbitrary metrics, and  $a, b, \dots$  in conformal gauge.

propagating through curved space-time. This is obtained by simply substituting the metric  $g_{\mu\nu}$  of that space-time. In these lectures, however, we restrict ourselves to flat space-times. We remarked above that  $\mu$  is just a multiplicity. This is not quite true: because we embed the string in minkowski space, one of the bosons will have an action of the “wrong” sign. This has important consequences, as will be discussed later.

## 2.6 The equations of motion

Since the action  $S[X, \gamma]$  depends on two fields  $X$  and  $\gamma$  there are two equations of motion. If you are familiar with general relativity, you can immediately write down the one for the two-dimensional metric: this of course leads to the Einstein equations,  $R^{\alpha\beta} \propto T^{\alpha\beta}$ , where the left hand side is the Ricci tensor. This term is derived from the gravitational action  $\int \sqrt{-\gamma} R$ , which we do not have in our action.\* Hence the equation of motion is simply  $T^{\alpha\beta} = 0$ . If you are not familiar with general relativity you can derive this by varying  $S[X, \gamma]$  with respect to  $\gamma$ . This is exactly the calculation we did earlier. The definition of the Energy-Momentum tensor is

$$T^{\alpha\beta}(\sigma, \tau) \equiv \frac{4\pi\alpha'}{\sqrt{-\gamma}} \frac{\delta}{\delta\gamma_{\alpha\beta}} S[X, \gamma] \quad (2.22)$$

The factor  $4\pi\alpha'$  is for convenience. The computation yields

$$T^{\alpha\beta}(\sigma, \tau) = \partial^\alpha X^\mu \partial^\beta X_\mu - \frac{1}{2} \gamma^{\alpha\beta} \partial_\gamma X^\mu \partial^\gamma X_\mu \quad (2.23)$$

To compute this, compute first  $\delta_\gamma S[X, \gamma]$  as was done before. This yields an integral of the form

$$\delta_\gamma S = \int d\sigma' d\tau' \text{Tr} (F(\sigma', \tau') \delta_\gamma(\sigma', \tau')) , \quad (2.24)$$

for some matrix function  $F$ . Now we use the functional derivative

$$\frac{\delta\gamma_{\alpha\beta}(\sigma', \tau')}{\delta\gamma_{\gamma\delta}(\sigma, \tau)} = \delta_{\alpha\gamma} \delta_{\beta\delta} \delta(\sigma - \sigma') \delta(\tau - \tau') , \quad (2.25)$$

which is functional analog of the discrete variable result

$$\frac{\partial q_i}{\partial q_j} = \delta_{ij} \quad (2.26)$$

The other equation of motion is the one for  $X^\mu$ . Although it can be computed for arbitrary  $\gamma$ , for simplicity we go to conformal gauge. Then the equation is simply

$$\left[ \frac{\partial^2}{\partial\tau^2} - \frac{\partial^2}{\partial\sigma^2} \right] X^\mu = 0 \quad (2.27)$$

---

\* It can be added, and plays an interesting rôle. In two dimensions this term is however purely topological, and does not contribute to the equations of motion.

In two dimensions this equation can be solved very easily by writing it in terms of variables  $\sigma^+ = \sigma + \tau$  and  $\sigma^- = \sigma - \tau$ . The general solution is

$$X^\mu(\sigma, \tau) = X_L^\mu(\tau + \sigma) + X_R^\mu(\tau - \sigma) , \quad (2.28)$$

where  $X_L^\mu$  and  $X_R^\mu$  are arbitrary functions. A special point of  $X_L$ , for example the maximum, stays at a fixed value of  $\sigma + \tau$ . Hence as  $\tau$  increases,  $\sigma$  must decrease. Then the maximum of  $X_L$  moves to the left, and analogously the maximum of  $X_R$  moves to the right. Therefore we call these left- and right-moving components.

## 2.7 Conformal invariance

It is known from classical electrodynamics that in certain cases some symmetries are left even though the gauge was fixed. For example, if we fix the Lorentz gauge,  $\partial^\mu A_\mu = 0$ , then we still have not fixed gauge transformations  $A_\mu \rightarrow A_\mu + \partial_\mu \Lambda$ , with  $\square \Lambda = 0$ .

The same occurs here. It turns out that there are special coordinate transformations that change the metric by an overall factor:

$$\frac{\partial \sigma'^\gamma}{\partial \sigma^\alpha} \frac{\partial \sigma'^\delta}{\partial \sigma^\beta} \gamma'_{\gamma\delta}(\tau', \sigma') = \Lambda(\sigma, \tau) \gamma_{\alpha\beta}(\sigma, \tau) \quad (2.29)$$

Then one can remove  $\Lambda$  using a Weyl transformation, and the net effect is that the metric has not changed at all. Hence if we start with the conformal gauge metric  $\eta^{ab}$  we keep the same metric, and therefore we have a residual symmetry.

This problem can be formulated in any dimension. One can completely classify all transformations of the coordinates that satisfy (2.29). It should be clear that this includes in any case (Lorentz)-rotations,  $\sigma'^a = \Lambda^a_b \sigma^b$ , and translations,  $\sigma'^a = \sigma^a + \epsilon^a$ , since they do not change the metric at all. Another transformation with this property is a scale transformation,  $\sigma'^a = \lambda \sigma^a$ .

The name of this kind of a transformation derives from the fact that it preserves angles, but not lengths. Consider two vectors  $v^\alpha$  and  $w^\alpha$ , which may be in any dimension. The angle between these vectors (or its obvious generalization to Minkowski space) is

$$\frac{v^\alpha g_{\alpha\beta} w^\beta}{\sqrt{(v^\alpha g_{\alpha\beta} v^\beta)(w^\gamma g_{\gamma\delta} w^\delta)}} \quad (2.30)$$

A scale factor  $\Lambda(\sigma, \tau)$  in the metric obviously drops out in the ratio, so angles are indeed preserved. So in some sense these transformation preserve the “form” of an object, but not its size. This is why they are called *conformal transformations*.

Just as symmetries in other systems, these transformations form a group, called the *conformal group*. We have already seen that it contains the Lorentz group as a subgroup. In  $p$  space and  $q$  time dimensions, the Lorentz group is  $SO(p, q)$ . The conformal group turns out to be  $SO(p + 1, q + 1)$ , if  $p + q > 2$ .



If  $p + q = 2$ , however, the conformal group is much larger. This can easily be seen as follows. Let us try the transformation

$$\begin{aligned}\tau' &= f(\tau + \sigma) + g(\tau - \sigma) \\ \sigma' &= f(\tau + \sigma) - g(\tau - \sigma)\end{aligned}\tag{2.31}$$

where  $f$  and  $g$  are arbitrary continuous functions. The transformation of the Minkowski metric  $\eta^{ab}$  under such a transformation leads to the matrix multiplication

$$\begin{pmatrix} f' + g' & f' - g' \\ f' - g' & f' + g' \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} f' + g' & f' - g' \\ f' - g' & f' + g' \end{pmatrix} = 4f'g' \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\tag{2.32}$$

which proves that the metric transforms to a multiple of itself, and hence that this is a conformal transformation. Thus we see that we may replace  $\sigma + \tau$  and  $\sigma - \tau$  by arbitrary functions,  $f(\sigma + \tau)$  and  $g(\sigma - \tau)$ . This is a symmetry with an infinite number of generators, certainly much larger than the conformal group  $SO(2, 2)$  we would expect by extrapolating the higher dimensional result to two dimensions. This group is in fact contained in these transformations: the  $SO(2, 2)$  transformations turn out to correspond to functions  $f$  and  $g$  of the form  $f(x) = a + bx + cx^2$ .

### 3 String spectra

Consider an observer looking at a string moving through space time. The string will have a center-of-mass motion, around which it moves. The motion around the center-of-mass can be decomposed in normal modes of vibration. When the string is in one of these modes, the observer will view the energy of the mode as the mass of a particle. There is an infinite number of modes of vibration, and when we quantize the theory, each of these modes will have its own harmonic oscillator wave function.

#### 3.1 Mode expansion

##### 3.1.1 Closed Strings

Now we have to make this intuitive notion precise. Let us first consider closed strings. The world sheet of a closed string is a cylinder. On this cylinder the action is, in conformal gauge,

$$S = -\frac{1}{4\pi\alpha'} \int d^2\sigma \partial_a X^\mu \partial^a X_\mu, \quad (3.1)$$

where  $d^2\sigma \equiv d\sigma d\tau$ . The fact that the fields  $X$  live on a cylinder implies that they satisfy periodic boundary conditions:

$$X^\mu(0, \tau) = X^\mu(2\pi, \tau) \quad (3.2)$$

The point  $\sigma = 2\pi$  is the standard choice, but one may show that if we would take any other choice, the spectrum does not depend on it. Any field satisfying these boundary condition can be Fourier expanded

$$X^\mu(\sigma, \tau) = \sum_{n=-\infty}^{\infty} e^{in\sigma} f_n^\mu(\tau) \quad (3.3)$$

The classical equation of motion for  $X$  is

$$[\partial_\sigma^2 - \partial_\tau^2]X^\mu(\sigma, \tau) = 0 \quad (3.4)$$

For the Fourier modes of a classical solution this implies

$$\partial_\tau^2 f_n^\mu(\tau) = -n^2 f_n^\mu(\tau). \quad (3.5)$$

The solution is

$$f_n^\mu(\tau) = a_n^\mu e^{in\tau} + b_n^\mu e^{-in\tau}, n \neq 0 \quad (3.6)$$

and

$$f_0^\mu(\tau) = p^\mu \tau + q^\mu. \quad (3.7)$$

Putting all this together, and introducing a few convenient factors, we may write the result as

$$X^\mu(\sigma, \tau) = q^\mu + \alpha' p^\mu \tau + i\sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \left\{ \frac{1}{n} (\alpha_n^\mu e^{-in(\tau+\sigma)} + \bar{\alpha}_n^\mu e^{-in(\tau-\sigma)}) \right\}, \quad (3.8)$$

Note that  $X^\mu$  should be real. This immediately implies

$$(\alpha_n^\mu)^* = \alpha_{-n}^\mu \quad (3.9)$$

### 3.1.2 Open String boundary conditions

For open strings we choose\* the  $\sigma$  interval as  $0 \leq \sigma < \pi$ . Clearly we need a boundary condition at the two extremes. The possible boundary conditions can be determined by considering the derivation of the equations of motion, with careful attention to boundary terms. When varying the integrand of the action (the Lagrangian density),

$$\mathcal{L}(X) \equiv \partial_a X_\mu \partial^a X_\nu \eta^{\mu\nu} , \quad (3.10)$$

the variation is defined as

$$\begin{aligned} \delta_X \mathcal{L} &\equiv \mathcal{L}(X + \delta X) - \mathcal{L}(X) \\ &= \partial_a \delta(X_\mu) \partial^a X_\nu \eta^{\mu\nu} + \partial_a X_\mu \partial^a \delta(X_\nu) \eta^{\mu\nu} \\ &= 2\partial_a (\delta X_\mu) \partial^a X^\mu \end{aligned} \quad (3.11)$$

Hence

$$\delta_X S = -\frac{1}{2\pi\alpha'} \int d\tau \int_0^\pi d\sigma \partial_a (\delta X_\mu) \partial^a X^\mu \quad (3.12)$$

which after an integration by parts yields (we assume here that the open string propagates from  $\tau = -\infty$  to  $\tau = +\infty$ , and that the variations of  $X^\mu$  vanish in these limits)

$$\delta_X S = \frac{1}{2\pi\alpha'} \int d\tau \int_0^\pi d\sigma (\delta X_\mu) \partial_a \partial^a X^\mu - \frac{1}{2\pi\alpha'} \int d\tau \delta X_\mu \partial_\sigma X^\mu \Big|_{\sigma=0}^{\sigma=\pi} \quad (3.13)$$

Requiring the first term to vanish for arbitrary  $\delta X_\mu$  leads to the equations of motion,  $\partial_a \partial^a X^\nu$ . The second term must also vanish, and this happens most easily by choosing  $\partial_\sigma X^\mu = 0$  at the boundaries  $\sigma = 0$  and  $\sigma = \pi$ . This is called a Neumann boundary. A second possibility would be to assume that  $\delta X_\mu$  vanishes at the boundary, so that  $X^\mu$  is constant. Then  $X^\mu(\sigma = 0, \tau) = c^\mu$  and  $X^\mu(\sigma = \pi, \tau) = d^\mu$ , where  $c$  and  $d$  are constant vectors. Clearly this is not Poincaré invariant, and in particular translation invariance is always broken. This is called a Dirichlet boundary condition. We will come back to it later, but for now we restrict ourselves to Neumann boundaries.

### 3.1.3 Open String mode expansion

The rest of the discussion is the same as in the closed case, except that the  $\sigma$ -modes are slightly different now. We find

$$X^\mu(\sigma, \tau) = q^\mu + 2\alpha' p^\mu \tau + i\sqrt{2\alpha'} \sum_{n \neq 0} \frac{1}{n} (\alpha_n^\mu e^{-in\tau} \cos n\sigma) , \quad (3.14)$$

---

\* The size of the interval is irrelevant because of scale invariance: any other choice can be rescaled to  $\pi$ .

### 3.1.4 Open versus closed

This shows why we defined the open string with a range  $0 \leq \sigma < \pi$  and we used twice this range for the closed string. The advantage of doing it that way is that the modes come out the same way, as a sum of terms of the form  $\exp(-in(\tau \pm \sigma))$ . This is convenient, but not essential: with any other choice for the range our final result for the spectrum would be the same. This is a consequence of scale invariance of the action: any range can always be scaled back to the one above.

The dependence on  $\alpha'$  is also different in the open and the closed case. The reason for that will become clear in the next section: it ensures that the quantized modes will have the same commutation relations in both cases.

Finally, note that the open string mode expansion can be written in the form (2.28), but with  $X_L = X_R$ . Intuitively the reason for this is that in closed strings left and right-moving modes are decoupled: a mode that moves to the left will continue to do so forever. In the open string theory such a mode will bounce off the boundary and start moving to the right.

## 3.2 Quantization

The system under consideration can be treated using the methods explained in Appendix A. The Lagrangian is

$$L = -\frac{1}{4\pi\alpha'} \int_0^{\rho\pi} d\sigma \partial_a X^\mu \partial^a X_\mu, \quad (3.15)$$

so that  $S = \int d\tau L$ . Here  $\rho = 1$  for open strings and  $\rho = 2$  for closed strings.

Now it is easy to derive the canonical momentum of  $X^\mu$ :

$$\Pi_\mu = \frac{1}{2\pi\alpha'} \partial_\tau X_\mu \quad (3.16)$$

From the Poisson brackets of the classical theory we infer the correct form of the commutator in the quantum theory:

$$\begin{aligned} [X^\mu(\sigma, \tau), \Pi^\nu(\sigma', \tau)] &= i\eta^{\mu\nu} \delta(\sigma - \sigma') \\ [X^\mu(\sigma, \tau), X^\nu(\sigma', \tau)] &= 0 \\ [\Pi^\mu(\sigma, \tau), \Pi^\nu(\sigma', \tau)] &= 0. \end{aligned} \quad (3.17)$$

These relations can be expressed in terms of modes. The result is\*

$$\begin{aligned} [\alpha_k^\mu, \alpha_l^\nu] &= [\bar{\alpha}_k^\mu, \bar{\alpha}_l^\nu] = k\eta^{\mu\nu} \delta_{k+l,0} \\ [\alpha_k^\mu, \bar{\alpha}_l^\nu] &= 0 \end{aligned} \quad (3.18)$$

---

\* Note that the commutators (3.17) are at some world-sheet time  $\tau = \tau_0$ . Hence also the modes  $\alpha_k$  are the coefficients in the expansion of  $X$  at  $\tau_0$ . The classical time evolution of these modes as given (3.8) is only used to distinguish left- and right-moving modes at  $\tau_0$ , but of course the time evolution is governed by quantum mechanics, and not by the classical evolution (3.8).

and

$$[q^\mu, p^\nu] = i\eta^{\mu\nu} \quad (3.19)$$

We conclude that  $q^\mu$  and  $p^\mu$  can be interpreted as the center-of-mass coordinate and momentum, as is already clear from the classical expression.

The  $\alpha$ -operators look a lot like harmonic oscillator operators. The quantum version of (3.20) is

$$(\alpha_n^\mu)^\dagger = \alpha_{-n}^\mu \quad (3.20)$$

If we now define

$$a_n^\mu = \frac{1}{\sqrt{n}}\alpha_n^\mu, \quad (a_n^\mu)^\dagger = \frac{1}{\sqrt{n}}\alpha_{-n}^\mu \quad (3.21)$$

we obtain a harmonic oscillator algebra for each value of  $\mu$  and  $n$ .

### 3.3 Negative norm states

Note however the following. A standard harmonic oscillator commutation relation has the form

$$[a, a^\dagger] = 1 \quad (3.22)$$

One defines the vacuum by the requirement  $a|0\rangle = 0$ . If one computes the norm of the state  $a^\dagger|0\rangle$  one finds

$$\|a^\dagger|0\rangle\|^2 = \langle 0|aa^\dagger|0\rangle = \langle 0|[a, a^\dagger]|0\rangle = 1. \quad (3.23)$$

However, because of the appearance of  $\eta^{\mu\nu}$  one of the oscillators will have the commutation relation  $[a, a^\dagger] = -1$ . This leads to states of negative norm. States with negative norm often occur in quantum field theory in intermediate stages of a calculation. They are called *ghosts*. In sensible theories they should cancel. In particular, there should not exist scattering processes of physical particles in which ghosts are produced. At this stage, the string spectrum seems to contain such quantum states. If this were the end of the story, the theory would be sick.

### 3.4 Constraints

However, it is not the end of the story. So far we have completely ignored the second equation of motion,  $T_{ab} = 0$ . Written in terms of components this tensor becomes

$$T_{00} = T_{11} = \frac{1}{2}(\dot{X}^2 + X'^2) \quad (3.24)$$

$$T_{10} = T_{01} = \dot{X} \cdot X' \quad (3.25)$$

where the dot and the prime indicate  $\tau$  and  $\sigma$  derivatives, and Lorentz-indices are implicitly contracted.

We may compare this with the Hamiltonian of our system

$$H = \int_0^{\rho\pi} d\sigma (\dot{X} \cdot \Pi) - L = \frac{1}{4\pi\alpha'} \int_0^{\rho\pi} (\dot{X}^2 + X'^2) \quad (3.26)$$

Hence we see that, up to normalization,  $H$  is just the integral over  $T_{00}$ , which has to vanish. The Hamiltonian is the generator of time translations. The integral over  $T_{01}$  is proportional to  $P$ , the generator of translations in the  $\sigma$  direction. Obviously, it also has to vanish. But that is still not all, because  $T_{ab}$  has to vanish without integration. Put differently, there is an infinite number of modes of  $T_{ab}$  that has to vanish.

### 3.5 Mode expansion of the constraints

To write down these modes it is convenient to go first to different coordinates,

$$\sigma^\pm = \tau \pm \sigma \quad (3.27)$$

Then the derivatives are

$$\partial_\pm = \frac{1}{2}(\partial_\tau \pm \partial_\sigma) . \quad (3.28)$$

In these coordinates the components of the energy-momentum tensor are

$$\begin{aligned} T_{++} &= \frac{1}{2}(T_{00} + T_{01}) = \partial_+ X \cdot \partial_+ X \\ T_{--} &= \frac{1}{2}(T_{00} - T_{01}) = \partial_- X \cdot \partial_- X \end{aligned} \quad (3.29)$$

Now we define the modes as follows (for closed strings)

$$\begin{aligned} L_n &= \frac{1}{2\pi\alpha'} \int_0^{2\pi} d\sigma e^{in\sigma} T_{++} \\ \bar{L}_n &= \frac{1}{2\pi\alpha'} \int_0^{2\pi} d\sigma e^{-in\sigma} T_{--} \end{aligned} \quad (3.30)$$

Here the “bar” on the second  $L$  should not be confused with complex conjugation; it only serves to distinguish left- from right-moving modes. Under complex conjugation we have in fact

$$L_n^* = L_{-n} , \quad \bar{L}_n^* = \bar{L}_{-n} \quad (3.31)$$

In the quantum theory the  $*$  becomes a  $\dagger$ . These mode expansions are done at  $\tau = 0$ .

With a little effort, we can substitute the mode expansions for  $X^\mu$  and express the  $L$ 's in terms of the modes of  $X$ . The result is

$$L_n = \frac{1}{2} \sum_m \alpha_{n-m} \cdot \alpha_m , \quad \bar{L}_n = \frac{1}{2} \sum_m \bar{\alpha}_{n-m} \cdot \bar{\alpha}_m \quad (3.32)$$

here we have defined

$$\alpha_0^\mu = \bar{\alpha}_0^\mu = \sqrt{\frac{1}{2}\alpha' p^\mu} \quad (3.33)$$

The advantage of the  $\sigma^\pm$  coordinates is that the right- and left-moving operators  $\alpha_n$  and  $\bar{\alpha}_n$  are almost completely separated, with the exception of the zero mode  $p^\mu$ , which has no left or right orientation.

For open strings we find in a similar way

$$\begin{aligned} L_n &= \frac{1}{2\pi\alpha'} \int_0^\pi d\sigma (e^{-in\sigma} T_{--} + e^{in\sigma} T_{++}) \\ &= \frac{1}{2} \sum_m \alpha_{n-m} \cdot \alpha_m \end{aligned} \tag{3.34}$$

with  $\alpha_0^\mu = \sqrt{2\alpha'} p^\mu$ .

### 3.6 The Virasoro constraints

If we ignore the problem of ghost states for the moment, then the full Hilbert space of states would be obtained as follows. First we choose a vacuum that is annihilated by all the annihilation operators,

$$\alpha_n^\mu |0\rangle = \bar{\alpha}_n^\mu |0\rangle = 0, \quad n > 0 \tag{3.35}$$

and that furthermore has zero momentum

$$p^\mu |0\rangle = 0 \tag{3.36}$$

This zero-momentum state can be transformed into a ground state with momentum in the usual way

$$|k^\mu\rangle = e^{iq^\mu k_\mu} |0\rangle, \quad p^\mu |k^\mu\rangle = k^\mu |k^\mu\rangle, \tag{3.37}$$

where of course  $p^\mu$  is an operator and  $k^\mu$  a momentum vector. Then we obtain all excited states by action with the negative modes of the operators  $\alpha$  and  $\bar{\alpha}$ . A typical, unnormalized state the looks like

$$|n_1, \dots, n_\ell, \mu_1, \dots, \mu_\ell, m_1, \dots, m_\kappa, \nu_1, \dots, \nu_\kappa, k^\mu\rangle = \alpha_{-n_1}^{\mu_1} \dots \alpha_{-n_\ell}^{\mu_\ell} \bar{\alpha}_{-m_1}^{\nu_1} \dots \bar{\alpha}_{-m_\kappa}^{\nu_\kappa} |k^\mu\rangle.$$

For open strings the space of states looks similar, except that one doesn't have the  $\bar{\alpha}$  operators.

Now we have to see what the effect is of the energy-momentum constraints. We have just introduced the modes of the energy-momentum tensor. These modes, the operators  $L_n$  and  $\bar{L}_n$ , are called the Virasoro operators. Since the energy-momentum tensor  $T^{ab}$  classically has to vanish, one might come to the conclusion that in the quantum theory we should impose the constraint

$$L_n |\text{phys}\rangle = \bar{L}_n |\text{phys}\rangle = 0 \tag{3.38}$$

This would restrict the number of quantum states, and one might hope that the ghost states disappear.

### 3.7 Operator ordering

However, there are two problems with this idea. First of all the Virasoro operators are not yet well-defined. This is because in the quantum theory we replace the classical quantities  $\alpha$  and  $\bar{\alpha}$  by non-commuting operators. The ordering of the classical objects is irrelevant, but the ordering of the quantum operators makes a difference.

Since the only non-commuting objects are  $\alpha_k$  and  $\alpha_{-k}$ , this ordering problem only affects the operator  $L_0$ .\* The solution to this problem is to impose an ordering by hand, and to define  $L_0$  as

$$L_0 \equiv \frac{1}{2} : \sum_m \alpha_{-m} \alpha_m : \quad (3.39)$$

where the colons,  $:$ , indicate *normal ordering*. This means that all creation operators are written to the left of all the annihilation operators. Consequently the vacuum expectation value of any non-trivial normal-ordered operator  $\mathcal{O}$  is zero,  $\langle 0 | : \mathcal{O} : | 0 \rangle = 0$ , because either the state  $|0\rangle$  is destroyed by the annihilation operators acting to the right, or  $\langle 0|$  is destroyed by the creation operators action to the left. Only when an operator contains no operators  $\alpha_k, k \neq 0$  at all it may have a non-trivial vacuum expectation value. The normal-ordered Virasoro zero-mode operator may also be written as

$$L_0 = \frac{1}{2} \alpha_0^2 + \sum_{m>0} \alpha_{-m} \alpha_m . \quad (3.40)$$

It is clear that changing the order of two oscillators changes  $L_0$  by a constant. The fact that we have imposed an ordering does not mean we can ignore that constant. It only means that there is an unknown constant in the relation between  $L_0$  and the Hamiltonian:

$$\begin{aligned} H &= (L_0 - a) + (\bar{L}_0 - \bar{a}) && \text{closed strings} \\ H &= (L_0 - a) && \text{open strings .} \end{aligned} \quad (3.41)$$

This also affects the condition we must impose on the spectrum:

$$(L_0 - a) |\text{phys}\rangle = (\bar{L}_0 - \bar{a}) |\text{phys}\rangle = 0 \quad (3.42)$$

### 3.8 Commutators of constraints

But that is not the only problem. For example, consider the state

$$L_{-2} |0\rangle = \frac{1}{2} \sum_p \alpha_{-2+p} \cdot \alpha_{-p} |0\rangle \quad (3.43)$$

In the infinite sum, the only term that contributes is the one with  $p = 1$ :

$$L_{-2} |0\rangle = \frac{1}{2} \alpha_{-1} \cdot \alpha_{-1} |0\rangle \quad (3.44)$$

---

\* Here and in the following we will omit the analogous expressions for the  $\bar{\alpha}$  operators that appear in the closed string theory.



If we try to impose (3.38) we would like the norm of this state to vanish. But that is clearly nonsense:

$$\|L_{-2}|0\rangle\| = \langle 0|\frac{1}{2}\alpha_1^\mu\alpha_{1,\mu}\frac{1}{2}\alpha_{-1}^\nu\alpha_{-1,\nu}|0\rangle = \frac{1}{2}\eta_{\mu\nu}\eta^{\mu\nu} = \frac{1}{2}D \neq 0 \quad (3.45)$$

This problem is a consequence of the fact that the  $L_n$ 's have non-trivial commutation relations. It turns out to be an interesting, but rather cumbersome exercise to compute  $[L_n, L_m]$ . One may first do this classically, using Poisson brackets instead of commutators for the modes  $\alpha_n$ . Although the Poisson brackets have a non-trivial right hand side\*,

$$\begin{aligned} \{\alpha_k^\mu, \alpha_l^\nu\}_{\text{PB}} &= \{\bar{\alpha}_k^\mu, \bar{\alpha}_l^\nu\}_{\text{PB}} = ik\eta^{\mu\nu}\delta_{k+l,0} \\ \{\alpha_k^\mu, \bar{\alpha}_l^\nu\}_{\text{PB}} &= 0, \end{aligned} \quad (3.46)$$

the modes commute with each other, and there is no ordering problem. It is rather easy to compute

$$\{L_m, L_n\}_{\text{PB}} = i(m-n)L_{m+n} \quad (3.47)$$

The calculation in the quantum case is essentially identical as long as we neglect ordering. Hence we get

$$[L_m, L_n] = (m-n)L_{m+n} + \text{re-ordering contributions} \quad (3.48)$$

The reordering contributions can be computed by carefully commuting the oscillators that are in the wrong place, but there is a more clever way.

### 3.9 Computation of the central charge

Since the reordering only affects  $L_0$ , it follows that the correct form of the commutation relation can only be

$$[L_m, L_n] = (m-n)L_{m+n} + A(m)\delta_{m+n}, \quad (3.49)$$

where  $A(m)$  is a c-number, *i.e.* it commutes with the  $L_n$ 's (such a contribution is called a ‘‘central charge’’ of an algebra). Obviously  $A(-m) = -A(m)$ . Furthermore we know already that  $A(2)$  is non-zero:

$$\left. \begin{aligned} \langle 0|L_2L_{-2}|0\rangle &= \frac{1}{2}D \\ \langle 0|L_{-2}L_2|0\rangle &= 0 \end{aligned} \right\} \rightarrow \langle 0|[L_2, L_{-2}]|0\rangle = \frac{1}{2}D, \quad (3.50)$$

using  $\langle 0|L_0|0\rangle = 0$ . Hence  $A(2) = \frac{1}{2}D$ . In a similar way one finds that  $A(0) = A(1) = 0$ .

An elementary property of commutators is the Jacobi-identity

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \quad (3.51)$$

which should hold independent of what  $A$ ,  $B$  and  $C$  are. In the case it yields, for  $A, B, C = L_1, L_{-1-n}$  and  $L_n$

$$(2n+1)A(1) - (n+2)A(n) - (1-n)A(n+1) = 0 \quad (3.52)$$

---

\* Poisson brackets are denoted here as  $\{\dots\}_{\text{PB}}$ .

This is a recursion relation for the coefficients. Since  $A(1) = 0$  the equation becomes

$$A(n+1) = \left(\frac{n+2}{n-1}\right) A(n) \quad (3.53)$$

The solution is

$$A(n) = \frac{1}{12} cn(n+1)(n-1), \quad (3.54)$$

and  $c$  can be fixed by means of the special case  $n = 2$ : this gives  $c = D$ . The final result is a famous commutation relation,

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{1}{12}c(m^3 - m)\delta_{m+n}, \quad (3.55)$$

which is called the *Virasoro algebra*.

### 3.10 The Virasoro algebra

The Virasoro algebra is an example of a Lie-algebra, just as the familiar angular momentum algebra  $[J_i, J_j] = i\epsilon_{ijk}J_k$ . Strictly speaking the Virasoro algebra given above is not really a Lie-algebra. One of the properties of a Lie-algebra is a “product” that closes and satisfies the Jacobi identities. For physicists, that product is usually a commutator. However, if we regard the  $L_n$ ’s as the elements, then the product does not close unless  $c = 0$ . The solution is to add one extra element to the algebra,  $C$ , which commutes with all the other elements and whose eigenvalues are  $c$ . An element that commutes with all others but appears on the right hand side of commutators is called a *central charge*.

The Virasoro algebra has appeared here in a theory of  $D$  free bosons, with  $c = D$ . It appears in fact in all two-dimensional field theories that have conformal invariance, but with different values of  $c$ . Field theories with conformal invariance are called *conformal field theories*.

There is another way to look at  $c$ . Remember that in the classical theory  $c = 0$ , but that  $c \neq 0$  is generated by quantum effects. The classical form of the algebra is changed by quantum effects. This means that the classical symmetry is broken. When this happens, we speak of an *anomaly*. The coefficient  $c$  is often called the *conformal anomaly*.

### 3.11 Imposing the Virasoro constraints

In any case, the constant  $c \neq 0$  in the Virasoro algebra tells us once again that it is inconsistent to require  $L_n |\text{phys}\rangle = 0$  for all  $n$ . Note however that for classical/quantum correspondence we may accept a weaker condition, namely

$$\langle \text{phys}' | L_n | \text{phys} \rangle = 0 \quad (n \neq 0). \quad (3.56)$$

For this to be satisfied it is sufficient to demand that

$$L_n |\text{phys}\rangle = 0 \quad \text{for } n > 0, \quad (3.57)$$

plus the reality condition  $L_n^\dagger = L_{-n}$ . In addition there is the zero-mode condition (3.42).

## 3.12 The mass shell condition

### 3.12.1 Closed strings

Let us first consider (3.42). For closed strings, substituting  $\alpha_0^\mu = \sqrt{\frac{1}{2}}\alpha'p^\mu$ , this yields

$$\left(\frac{1}{4}\alpha'p^2 + \sum_{m>0} \alpha_{-m} \cdot \alpha_m - a\right) |\text{phys}\rangle = 0 \quad (3.58)$$

plus the same condition with bars on all relevant quantities. These two conditions can be written as

$$p^2 = -M_L^2 = -M_R^2, \quad (3.59)$$

with

$$M_L^2 = \frac{4}{\alpha'} \left( \sum_{m>0} \alpha_{-m} \cdot \alpha_m - a \right) \quad (3.60)$$

and

$$M_R^2 = \frac{4}{\alpha'} \left( \sum_{m>0} \bar{\alpha}_{-m} \cdot \bar{\alpha}_m - \bar{a} \right) \quad (3.61)$$

The condition  $p^2 = -M^2$  is of course the mass-shell condition for a relativistic particle with mass  $M$ , and D-momentum  $(\sqrt{\vec{k}^2 + M^2}, \vec{k})$ . Depending on the excited state in which we find the string, it behaves as a particle with different mass. It is essential for this interpretation that  $M^2 > 0$ . If  $M^2 < 0$ , the ‘‘particle’’ is called a *tachyon*, which is a general name for particles exceeding the speed of light. To see this note that for normal particles the relativistic energy is given by  $E = \sqrt{\vec{p}^2 + M^2} = M\gamma$ , where  $\vec{p}$  is the 3-momentum and  $\gamma = (1 - v^2)^{-\frac{1}{2}}$  (with  $c = 1$ ). We can solve this relation for  $v$ , and we get  $v^2 = \vec{p}^2/(\vec{p}^2 + M^2)$ . This is smaller than 1 for a massive particle, equal to 1 for a massless particle and larger than 1 for a tachyon, which therefore moves faster than the speed of light. In field theory the presence of tachyons in the spectrum indicates an instability of the vacuum. The standard field theory description of a (free) massive scalar particle involves a scalar field  $\phi(x)$  with a Hamiltonian

$$H = \int d^3x \left[ \left(\frac{d\phi}{dt}\right)^2 + (\nabla\phi)^2 + M^2\phi^2 \right] \quad (3.62)$$

If  $M^2 > 0$  the configuration of minimal energy is  $\phi = 0$ , but if  $M^2 < 0$  the energy is not bounded from below, and in any case the standard vacuum (corresponding to  $\phi = 0$ ) is not a minimum. It is clear that the occurrence of tachyons in the spectrum is not a good thing.

### 3.12.2 Open strings

For open strings a similar analysis gives  $p^2 = -M^2$  with

$$M^2 = \frac{1}{\alpha'} \left( \sum_{m>0} \alpha_{-m} \cdot \alpha_m - a \right) \quad (3.63)$$

In principle “ $a$ ” could be different than in the closed case, but we will see that it is in fact the same. The factor of 4 in the overall factor in comparison to closed strings may look like it was the result of some conventions we made in the past but it is not. For example, we made a different choice for the range of  $\sigma$  in both cases, but because of scale invariance that should not matter, and indeed it doesn’t. We also normalized the momenta and oscillators in a certain way, but that was simply done in order to get the correct commutation relations (3.18).

### 3.13 Unphysical state decoupling

Now we should face the problem of the unphysical, negative norm states generated by  $\alpha^\mu$  with  $\mu = 0$ . There are three main approaches to doing this

- Light-cone gauge
- Covariant operator method
- Covariant path integral method

In the first method, which will be explained in a moment, one sacrifices explicit Lorentz-invariance (temporarily), but one gets only manifestly physical, positive norm states. In the end Lorentz invariance is checked. One then discovers that this requires  $a = 1$  and  $D = 26$ .

In the second method one continues along the lines of the previous section, and now imposes the remaining Virasoro constraints. Lorentz invariance is explicit at every stage, but the absence of unphysical states leads to the requirement  $a = 1, D = 26^*$ .

Absence of such unphysical states in this formalism is guaranteed by the so-called *no-ghost* theorem.

The last method requires knowledge of path integrals and gauge fixing (the Fadeev-Popov procedure), as well as BRST-quantization. Nowadays it is considered the most attractive method, but too much machinery is required to explain it here.

### 3.14 Light cone gauge

The idea of light cone gauge is to use the residual invariance (2.31). It allows us to choose a new world sheet time parameter

$$\tau' = f(\tau + \sigma) + g(\tau - \sigma). \tag{3.64}$$

---

\* This is somewhat oversimplified. In fact, absence of unphysical states for the non-interacting string leads to  $a \geq 1, D \leq 26$ . However, if  $a \neq 1$  and  $D \neq 26$  it is not known how to make the string theory interacting. If  $D < 26$  it turns out that classical conformal invariance is broken by quantum effects (the “conformal anomaly”), and therefore it is inconsistent to go to light-cone gauge. Therefore there is no contradiction with the results obtained in light-cone gauge.

We start by choosing new coordinates in space-time

$$X^i, \quad X^\pm = \frac{1}{\sqrt{2}}(X^0 \pm X^{D-1}) \quad (3.65)$$

As we have seen before, the equations of motion tell us that we can write

$$X^\mu(\tau, \sigma) = X_L^\mu(\tau + \sigma) + X_R^\mu(\tau - \sigma) \quad (3.66)$$

This splitting in left- and right-moving modes is obvious for the oscillator modes, but the center-of-mass coordinates require a bit more attention. Explicitly we have, for closed strings

$$\begin{aligned} X_L^\mu(\tau + \sigma) &= \frac{1}{2}q^\mu + \frac{1}{2}\alpha'p^\mu(\tau + \sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \left\{ \frac{1}{n}(\alpha_n^\mu e^{-in(\tau+\sigma)}) \right\} \\ X_R^\mu(\tau - \sigma) &= \frac{1}{2}q^\mu + \frac{1}{2}\alpha'p^\mu(\tau - \sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \left\{ \frac{1}{n}(\bar{\alpha}_n^\mu e^{-in(\tau-\sigma)}) \right\} \end{aligned} \quad (3.67)$$

and for open strings

$$\begin{aligned} X_L^\mu(\tau + \sigma) &= \frac{1}{2}q^\mu + \alpha'p^\mu(\tau + \sigma) + i\sqrt{2\alpha} \sum_{n \neq 0} \left\{ \frac{1}{n}(\alpha_n^\mu e^{-in(\tau+\sigma)}) \right\} \\ X_R^\mu(\tau - \sigma) &= \frac{1}{2}q^\mu + \alpha'p^\mu(\tau - \sigma) + i\sqrt{2\alpha} \sum_{n \neq 0} \left\{ \frac{1}{n}(\alpha_n^\mu e^{-in(\tau-\sigma)}) \right\} \end{aligned} \quad (3.68)$$

Now choose in (3.64)  $f = b_L X_L^\mu + c_L$ ,  $g = b_R X_R^\mu + c_R$ . We do this in the classical theory, prior to quantization. Clearly  $f$  and  $g$ , defined in this way, are functions of  $\tau + \sigma$  resp.  $\tau - \sigma$ , so they are valid choices for a conformal transformation. Obviously we can do this for at most one choice for  $\mu$ , and we choose  $\mu = +$ . Note that the constants  $b_L, c_L$  and  $b_R, c_R$  are allowed as free parameters, because the requirement that  $f$  (or  $g$ ) depend only on  $\tau + \sigma$  (or  $\tau - \sigma$ ) does not fix them. Now we get

$$\tau' = b_L X_L^+ + b_R X_R^+ + c_L + c_R \quad (3.69)$$

In order to combine the first two terms into  $X^+$  we choose  $b_L = b_R$ . Finally we fix the values of the parameters in such a way that

$$X^+ = q^+ + \rho\alpha'p^+\tau', \quad (3.70)$$

We choose  $\rho = 1$  for closed strings and  $\rho = 2$  for open strings, so that the final outcome is that the expression for  $X^+$  looks exactly like before, except that all the oscillators are absent.

Note that these choices almost completely fix conformal invariance, since  $b_L = b_R = 1/\rho\alpha'p^+$  and  $c_L + c_R = -q^+/\rho\alpha'p^+$ . Using (2.31) we find for  $\sigma'$

$$\begin{aligned}\sigma' &= \frac{1}{\rho\alpha'p^+} (X_L^+ - X_R^+) + c_L - c_R \\ &= \sigma + c_L - c_R + \text{oscillators} ,\end{aligned}\tag{3.71}$$

in other words  $\sigma'$  is determined up to a constant shift  $c_L - c_R$ . Such a shift does not respect the open string range  $0 \leq \sigma < \pi$ , but is allowed for closed strings, where we only have a periodicity condition,  $\sigma \approx \sigma + 2\pi$  (note that the oscillator terms do not contribute to the boundary conditions). Hence for open strings we are forced to take  $c_L = c_R$ , and all residual invariance is now fixed by the choices we made. For closed strings the remaining residual invariance is now reduced to a constant,  $c_L - c_R$ .

Now we have to perform the seemingly daunting task to invert these relations, and express  $\sigma$  and  $\tau$  in terms of  $\sigma'$  and  $\tau'$ , and then substitute the result in (3.67) and (3.68). However, for one coordinate the result is already clear: we can immediately read off from (3.70) how  $X^+$  depends on the new coordinates  $\tau'$  and  $\sigma'$ . For all other components of  $X^\mu$  the result of the operation will be that all the classical oscillator coefficients  $p, q, \alpha, \bar{\alpha}$  are modified to new ones, which we may denote as  $p', q', \alpha', \bar{\alpha}'$ . The point is that (3.67) and (3.68) are already the most general expansions satisfying the boundary conditions, and after the conformal transformation, no matter how complicated, we inevitably get a new expression of the same form. So the net result is that  $X^+$  is given by (3.70), and all other components of  $X$  look the same as before, but with quantities  $\tau', \sigma', p', q', \alpha'$  and  $\bar{\alpha}'$  instead of unprimed ones. To keep the notation simple we now drop the primes on all these quantities. The benefit so far is that we have been able to completely flatten the component  $X^+$ , thus reducing the number of dynamical variables.

Now let us look at the constraints

$$\left. \begin{aligned}\dot{X}^2 + X'^2 &= 0 \\ X' \cdot \dot{X} &= 0\end{aligned}\right\} \rightarrow (\dot{X} \pm X')^2 = 0 .\tag{3.72}$$

In light cone coordinates, the inner product of two vectors  $V^\mu$  and  $W^\mu$  becomes

$$V^\mu W_\mu = V^i W^i - V^+ W^- - V^- W^+\tag{3.73}$$

Hence the constraints become

$$2(\dot{X}^- \pm X'^-)(\dot{X}^+ \pm X'^+) = (\dot{X}^i \pm X'^i)^2\tag{3.74}$$

substituting  $X^+$  we get

$$(\dot{X}^- \pm X'^-) = \frac{1}{2\alpha'\rho p^+} (\dot{X}^i \pm X'^i)^2\tag{3.75}$$

The choice of upper or lower sign gives separate constraints for left- and right-moving modes. These constraints allow us to express  $X^-$  into  $X^i$  (up to a constant, since the

constraints only involve the  $\tau$  and  $\sigma$  derivatives), so that now not only  $X^+$  but also  $X^-$  is eliminated as an independent dynamical degree of freedom.

In terms of modes, it allows us to express  $\alpha_n^-$  (and  $\bar{\alpha}_n^-$ ) in terms of a bi-linear function of  $\alpha_n^i$  (resp  $\bar{\alpha}_n^i$ ). The result is, classically, and for  $n \neq 0$ :

$$\alpha_n^- = \sqrt{\frac{2}{\alpha'} \frac{1}{\rho p^+}} \left( \frac{1}{2} \sum_m \alpha_{n-m}^i \alpha_m^i \right) \quad (3.76)$$

and similarly with bars, in the case of closed strings. This explicitly solves all modes of the two constraints with  $n \neq 0$ . Hence the modes  $\alpha_n^-$  and  $\bar{\alpha}_n^-$  are not independent degrees of freedom; they can be expressed in terms of  $\alpha_n^i$  and  $\bar{\alpha}_n^i$ . In the quantum theory this implies that we will only get canonical commutation relations for the true dynamical degrees of freedom, the modes of  $X^i$ .

The case  $n = 0$  requires a bit more attention. From (3.75) we get a formula that is quite similar to (3.76), with  $\alpha_0^- = \bar{\alpha}_0^- = \rho \sqrt{\frac{1}{2} \alpha' p^-}$ . So the first difference is that, unlike for the non-zero modes, the zero-modes  $\alpha_0^-$  and  $\bar{\alpha}_0^-$  are related for closed strings. We do not get separate equations for  $\alpha_0^-$  and  $\bar{\alpha}_0^-$ , since  $\alpha_0^-$  only appears in  $\dot{X}^-$  but not in  $X'^-$ . The second difference occurs when we quantize the theory, which up to now we did not do. But we already know that in the quantum theory for  $n = 0$  we will have an ordering problem, which we solve, as before, by normal ordering. This introduces an unknown constant  $a$ , the same constant we saw earlier when we only considered  $L_0$ . Let us first consider the constraint on  $\dot{X}^-$ , *i.e.* the sum of the left- and right constraint. After a little bit of work we find then

$$\sqrt{\frac{1}{2} \alpha' p^-} = \frac{1}{2} \sqrt{\frac{2}{\alpha'} \frac{1}{p^+}} \left( \frac{1}{2} \sum_m \alpha_{-m}^i \alpha_m^i - a + \frac{1}{2} \sum_m \bar{\alpha}_{-m}^i \bar{\alpha}_m^i - \bar{a} \right) \quad (\text{Closed strings}) \quad (3.77)$$

$$2 \sqrt{\frac{1}{2} \alpha' p^-} = \sqrt{\frac{2}{\alpha'} \frac{1}{2p^+}} \left( \frac{1}{2} \sum_m \alpha_{-m}^i \alpha_m^i - a \right) \quad (\text{Open strings}) \quad (3.78)$$

Bringing all momenta to one side (note that the terms  $\alpha_{-m}^i \alpha_m^i$  contain the contribution of the  $p^i$  components of the momenta) we find for the closed string

$$p^2 = -\frac{1}{2} (M_L^2 + M_R^2)$$

$$M_L^2 = \frac{4}{\alpha'} \left( \sum_{m>0} \alpha_{-m}^i \alpha_m^i - a \right); \quad M_R^2 = \frac{4}{\alpha'} \left( \sum_{m>0} \bar{\alpha}_{-m}^i \bar{\alpha}_m^i - \bar{a} \right) \quad (3.79)$$

and for the open string

$$p^2 = -M^2$$

$$M^2 = \frac{1}{\alpha'} \left( \sum_{m>0} \alpha_{-m}^i \alpha_m^i - a \right) \quad (3.80)$$

Did we now solve all the constraints? Clearly we did for all the non-zero modes, but for the zero-modes we have only considered so far the *sum* of the left and right constraint, and we still need to consider the difference, *i.e.*  $X'^-$ . For the open string the corresponding condition is automatically satisfied, but for the closed string it is non-trivial. It simply corresponds to (3.76) with  $n = 0$ , taking the difference between the equations for  $\alpha$  and  $\bar{\alpha}$ , and using  $\alpha_0^- = \bar{\alpha}_0^-$ . This directly leads to the requirement  $M_L = M_R$ .

Note that there is no constraint on  $q^-$ . Also  $p^+$  remains as a dynamical variable. In light cone coordinates, they will satisfy the canonical commutation relation  $[q^-, p^+] = i\eta^{-+} = -i$ . The parameter  $q^+$  would have as its canonical momentum  $p^-$ , which is fixed by the constraints. Hence there is no canonical commutator involving  $q^+$ . In fact we could have taken  $c_L = c_R = q^+ = 0$  when we defined the conformal transformation.

### 3.15 The spectrum

Let us first consider open strings. The physical space-time state corresponding to the vacuum  $|0\rangle$  in the two-dimensional theory has a mass  $M^2 = -a(\frac{1}{\alpha'})$ . At the next excitation level we find a state

$$\alpha_{-1}^i |0\rangle \tag{3.81}$$

It has mass  $M^2 = (1 - a)(\frac{1}{\alpha'})$ , and there are  $D - 2$  of them. The index  $i$  suggests that this states wants to be a space-time vector. Covariantly a space-time vector has  $D$  components. Not all these components are physical. If a  $D$ -dimensional particle is massive, we can go to its rest frame. We know then that it has  $D - 1$  physical components, corresponding to its polarizations. If a particle is massless, we can not go to its rest frame, but we can Lorentz-rotate its momentum to the form  $(E, k, 0, \dots, 0)$ , with  $k = E$ . The physical components correspond to the rotational degrees of freedom in the last  $D - 2$  components. There are thus  $D - 2$  physical degrees of freedom. These rotation groups  $SO(D - 2)$  and  $SO(D - 1)$  are called the *little groups* for massless and massive particles respectively.

This situation is well-known in electrodynamics. A photon  $A_\mu$  starts with  $D$  degrees of freedom, but one is eliminated because of the equation of motion for  $A_0$ , which contains no time derivative and hence is just a constraint ( $\partial_i E^i = 0$ , also known as Gauss' law), whereas another one is eliminated by gauge invariance. The equation of motion for a vector boson of mass  $M$  is

$$\partial_\mu F^{\mu\nu} = M^2 A^\nu, \tag{3.82}$$

which has gauge invariance  $A_\mu \rightarrow A_\mu + \partial_\mu \Lambda$  only for  $M = 0$  (whereas the non-dynamical nature of  $A_0$  holds independent of  $M$ ).

Turning the argument around, the fact that we have here  $D - 2$  components strongly suggests that this state must be a massless vector boson. This means that  $a$  must be equal to 1.

At the next excited level we find the following states

$$\alpha_{-1}^i \alpha_{-1}^j |0\rangle, \quad \alpha_{-2}^i |0\rangle \tag{3.83}$$

Note that in the first set the state with indices  $(i, j)$  is the same as the one with indices  $(j, i)$ , and we should count it only once! In total there are then  $\frac{1}{2}(D - 2)(D - 1) + (D - 2) =$



$\frac{1}{2}(D-2)(D+1)$  physical components. We expect to find (at least) a massive symmetric tensor  $T_{\mu\nu}$ .

The analysis of physical states for higher spin (massive or massless) tensor fields is somewhat more complicated than for vector bosons, but the idea is the same. The result is that the physical states for a massless particle in  $D$  dimensions are given by a representation of the “transverse group”  $SO(D-2)$ , whereas for a massive particle they are given by an  $SO(D-1)$  representation. These groups are called the “little groups” of massless and massive particles respectively. They are subgroups of the Lorentz group  $SO(D-1, 1)$  that leave the momentum vector of a particle invariant. For massive particles we can determine them most easily in the rest frame,  $P^\mu = (M, 0, \dots, 0)$ , whereas for massless particles we can go to a frame with  $P^\mu = (E, E, 0, \dots, 0)$ .

Some interesting representations of  $SO(N)$  are the singlet (dimension 1), and the vector (dimension  $N$ ). Furthermore we will need the rank-2 tensor representation. A general rank-2 tensor can be obtained by combining two vectors  $V^i$  and  $W^j$ :  $T^{ij} = V^i W^j$ . Clearly there are  $N^2$  components. But these  $N^2$  components can be split into three groups that transform into themselves:

$$T^{ij} = \left[ \frac{1}{2}(T^{ij} + T^{ji} - \frac{2}{N}\delta^{ij} \text{Tr } T) \right] + \left[ \frac{1}{2}(T^{ij} - T^{ji}) \right] + \frac{1}{N}[\delta^{ij} \text{Tr } T] . \quad (3.84)$$

The three components are called the rank-2 traceless symmetric tensor, the rank-2 anti-symmetric tensor and the singlet. Since they cannot be split further, they are called *irreducible* representations. To verify this splitting, note that a vector transforms as  $V^i = \sum_j O^{ij} V^j$ , where  $O$  is an orthogonal  $N \times N$  matrix, *i.e.*  $O^T O = O O^T = 1$ . The number of components of these tensors is respectively  $\frac{1}{2}N(N+1) - 1 = \frac{1}{2}(N+2)(N-1)$ ,  $\frac{1}{2}N(N-1)$  and 1. Later we will also need another type of representation, the *spinor* representation.

In  $D$  dimensions a symmetric tensor has  $\frac{1}{2}(D+1)(D-2)$  physical components (the dimension of the traceless symmetric tensor of  $SO(D-1)$ ), and we see that this precisely uses up all physical components. Hence there is no inconsistency with the assumption that this state is massive. We could go on like this, but we would not find any further inconsistencies.

For the closed string the reasoning is similar. The ground state  $|0\rangle$  has a mass  $M^2 = -\frac{4a}{\alpha'}$ . The first excited states are (setting  $\alpha' = 4$  for simplicity)

State	$M_L$	$M_R$
$ 0\rangle$	$-a$	$-a$
$\alpha_{-1}^i  0\rangle$	$-a$	$1-a$
$\bar{\alpha}_{-1}^i  0\rangle$	$1-a$	$-a$
$\alpha_{-1}^i \bar{\alpha}_{-1}^j  0\rangle$	$1-a$	$1-a$

We see then that the first excited state satisfying the constraint  $M_L = M_R$  is  $\alpha_{-1}^i \bar{\alpha}_{-1}^j |0\rangle$ . Note that there is no symmetry in  $i$  and  $j$  in this case, hence there are  $(D-2)^2$  states. One expects therefore at least a covariant symmetric tensor field  $g_{\mu\nu}$  and an anti-symmetric

tensor field  $B_{\mu\nu}$ . One can decompose these into physical components, assuming that they are massive, or assuming that they are massless. Just as for the vector field, there are not enough degrees of freedom to allow massive fields, and therefore they have to be massless. This again gives  $a = 1$ . It turns out that there is just one degree of freedom left over, a scalar  $\phi$ . This follows from the decomposition of the  $(D - 2)^2$  states into irreducible  $SO(D - 2)$  representations.

What are these massless fields? The symmetric tensor can be identified with the graviton. The scalar is called the *dilaton*, and in addition there is the antisymmetric tensor  $B_{\mu\nu}$ . Experimentally only the first has been observed (or more precisely, its effects have been observed in the form of gravity). There is no evidence for the other two. In four dimensions, the field  $B_{\mu\nu}$  has just one physical degree of freedom. This can be seen by computing the dimension of the anti-symmetric tensor representation of the little group. In  $D$  dimensions, and hence  $D - 2$  transverse dimensions, this is  $\frac{1}{2}(D - 2)(D - 3)$ , and for  $D = 4$  we find 1. Therefore this particle can only be a scalar under rotations. It can be shown, using the field equations, that in fact it can be transformed into a pseudo-scalar (a scalar that is odd under parity), which is often called an *axion*.

As in the case of open strings, one can continue and discover an infinite series of massive states. Only the ones with  $M_L = M_R$  are physical. If one counts the number of degrees of freedom, which quickly gets very complicated, one will always find that there are precisely enough states to fill up some set of irreducible representation of the little group  $SO(D - 1)$ .\*

### 3.16 The critical dimension

The constant  $a$ , which was just determined to be 1 can also be determined by carefully re-ordering the oscillators. For a single harmonic oscillator this gives the zero-point energy. Usually one starts with a Hamiltonian (we choose the mass equal to 1)

$$H = \frac{1}{2}(p^2 + \omega^2 q^2) , \quad (3.85)$$

with  $[q, p] = i$ . Then one introduces new operators  $a = \frac{1}{\sqrt{2\omega}}(p - i\omega q)$ ,  $a^\dagger = \frac{1}{\sqrt{2\omega}}(p + i\omega q)$ . The Hamiltonian becomes then

$$H = \frac{1}{2}\omega(aa^\dagger + a^\dagger a) = \omega[a^\dagger a + \frac{1}{2}] \quad (3.86)$$

Hence the energy of the vacuum is  $\frac{1}{2}\omega$ .

One may proceed in the same way here. The mass of a string excitation in light-cone gauge is an eigenvalue of the Hamiltonian

$$H = \frac{1}{4\pi\alpha'} \int d\sigma \sum_{i=1}^{D-2} [(\dot{X}^i)^2 + (X'^i)^2] \quad (3.87)$$

---

\* Strictly speaking that is an empty statement, because one can always get the correct count of states by allowing arbitrary numbers of singlets. The additional information one has to take into account, and which makes the statement non-trivial, is that we already know how the subgroup  $SO(D - 2)$  acts on the states, because that symmetry is explicit.

Substituting the quantum mode expansion for  $X^i$  without interchanging any operators one gets (for open strings)

$$H = \frac{1}{2\alpha'} \sum_{i=1}^{D-2} \sum_m \alpha_{-m}^i \alpha_m^i = \frac{1}{\alpha'} \left( \frac{1}{2} \sum_{i=1}^{D-2} \sum_m : \alpha_{-m}^i \alpha_m^i : + \frac{1}{2} (D-2) \sum_{n=1}^{\infty} n \right) \quad (3.88)$$

The sum diverges: it is a sum over the zero-mode energy of an infinite number of oscillators! However, if we define

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} , \quad (3.89)$$

then the sum converges for  $\text{Re } s > 1$ . This function is the Riemann zeta-function. It can be uniquely continued to  $s = -1$ , and the result is

$$\zeta(-1) = -\frac{1}{12} \quad (3.90)$$

Hence we conclude

$$a = \frac{1}{24} (D-2) , \quad (3.91)$$

and since we have already concluded that  $a = 1$ , this can only work if  $D = 26$ . This is called the *critical* dimension.

### 3.17 The Lorentz algebra\*

A better way to arrive at the conclusion that the space-time dimension must be 26 is to consider closure of the Lorentz algebra. In general, one has a set of generators  $M^{\mu\nu}$  that must satisfy the commutation relation

$$[M^{\mu\nu}, M^{\rho\sigma}] = -i\eta^{\nu\rho} M^{\mu\sigma} + i\eta^{\mu\rho} M^{\nu\sigma} + i\eta^{\nu\sigma} M^{\mu\rho} - i\eta^{\mu\sigma} M^{\nu\rho} \quad (3.92)$$

One can work out the precise form of these quantities in string theory. The result is

$$M^{\mu\nu} = q^\mu p^\nu - q^\nu p^\mu - i \sum_{n=1}^{\infty} \left\{ \frac{1}{n} (\alpha_{-n}^\mu \alpha_n^\nu - \alpha_{-n}^\nu \alpha_n^\mu) \right\} \quad (3.93)$$

(plus an  $\bar{\alpha}$  contribution for closed strings). It is not too difficult to check that this does indeed satisfy the commutation relation (3.92).

In light cone gauge,  $M^{\mu\nu}$  split into components  $M^{ij}$ ,  $M^{i-}$ ,  $M^{i+}$  and  $M^{+-}$ . The oscillators  $\alpha_n^-$  can be expressed in terms of a quadratic relation in  $\alpha^i$  (see (3.76)). This makes  $M^{i-}$  tri-linear in  $\alpha^i$ . The computation of the commutation relations of these operators  $M^{i-}$  with the others is technically of course totally different than before, but the result should be the same.

The hardest one to check is  $[M^{i-}, M^{j-}]$ , which should be zero. However, after some tedious work one finds

$$[M^{i-}, M^{j-}] = -\frac{1}{(p^+)^2} \sum_{m=1}^{\infty} \left\{ m \left( \frac{26-D}{12} \right) + \frac{1}{m} \left( \frac{D-26}{12} + 2(1-a) \right) \right\} (\alpha_{-m}^i \alpha_m^j - \alpha_{-m}^j \alpha_m^i) \quad (3.94)$$

This vanishes if and only if  $a = 1$  and  $D = 26$ .

Note that this also implies that if we choose these values, then the Lorentz algebra closes on the sets of states. In particular, all the states must then fall into Lorentz multiplets. For example, in the open string theory we found at the first excited level a symmetric tensor plus a vector of  $SO(D-2)$ . These nicely combined into a traceless symmetric tensor of  $SO(D-1)$ . Such “miracles” have to occur at any level, or else Lorentz invariance is violated. The closure of the Lorentz algebra guarantees this.

### 3.18 Unoriented strings

We have up to now discussed open and closed strings, but only oriented ones. The transformation that flips the orientation is the world-sheet parity transformation

$$\begin{aligned} \sigma' &= 2\pi - \sigma && \text{(closed strings)} \\ \sigma' &= \pi - \sigma && \text{(open strings)} \end{aligned} \quad (3.95)$$

These transformations map the world sheet into itself, but with a sign change of  $\sigma$ . Oriented strings are not invariant under this transformation, whereas unoriented strings – by definition – are invariant.

When one applies this to the mode expansions for open and closed strings one readily discovers that their effect is equivalent to the following transformation of the modes

$$\begin{aligned} \alpha_n^\mu &\rightarrow (-1)^n \alpha_n^\mu && \text{(open strings)} \\ \alpha_n^\mu &\leftrightarrow \bar{\alpha}_n^\mu && \text{(closed strings)} . \end{aligned} \quad (3.96)$$

We can implement this on the quantum theory by introducing an orientation-reversal operator  $\Omega$  with the property

$$\begin{aligned} \Omega \alpha_n^\mu \Omega^{-1} &= (-1)^n \alpha_n^\mu && \text{(open strings)} \\ \Omega \alpha_n^\mu \Omega^{-1} &= \bar{\alpha}_n^\mu && \text{(closed strings)} . \end{aligned} \quad (3.97)$$

On the states we then impose the requirement that they be invariant under the action of  $\Omega$ . Now everything is fixed if we just give a rule for the action of  $\Omega$  on the vacuum. The correct choice is  $\Omega |0\rangle = |0\rangle$  (note that choosing a  $-$ sign here removes the vacuum from the two-dimensional spectrum, and hence it removes the tachyon from the space-time spectrum. Although this looks attractive at first sight, it leads to inconsistencies). Then at the first excited level of open strings the states  $\alpha_{-1}^i |0\rangle$  is transformed into minus itself by  $\Omega$ , and therefore is not in the spectrum. At the first excited level of the closed

string the states  $\alpha_{-1}^i \bar{\alpha}_{-1}^j |0\rangle$  is transformed into the states with  $i$  and  $j$  interchanged, and hence only the symmetric part of the tensor survives. Hence we only get a graviton and a dilaton, but no anti-symmetric tensor.

This can all be summarized as follows

String Theory	Massless Spectrum
Closed, Oriented	$g_{\mu\nu}, B_{\mu\nu}, \phi$
Closed, Unoriented	$g_{\mu\nu}, \phi$
Open, Oriented	$A_\mu$
Open, Unoriented	—

One should remember that open strings cannot exist by themselves, but always require the existence of closed strings ((un)oriented open strings require (un)oriented closed strings).

### 3.19 Chan-Paton labels\*

In the case of open strings there is one additional freedom. One may attach labels to the endpoints of the string. In other words, there can be  $N$  distinct kind of endpoints, labelled by an integer  $i = 1, \dots, N$ . This goes back to work by Chan and Paton in 1969, and is related to the fact that string theory was invented to describe hadronic interactions. Open strings represented mesons, and the labels at the end of the string were associated with quarks and anti-quarks.

To take this possibility into account when we quantize an open string, we may assign a separate vacuum state to each open string with the boundary at  $\sigma = 0$  labeled by “ $a$ ”, and the boundary at  $\sigma = \pi$  labeled by “ $b$ ”. Hence we add the two labels to the vacuum state:  $|0; a, b\rangle$ .

If the open string is oriented the two labels are independent, but consistency requires them to have the same multiplicity  $N$ . Then we find  $N^2$  distinct vacuum states, and at the first excited level we find  $N^2$  vector bosons. In the free string theory we can only count them, but when we make the theory interaction we can find out more. It turns out that these  $N^2$  vector bosons must be gauge bosons of a  $U(N)$  gauge group.

If the string theory is unoriented, we have to remember that the world sheet parity transformation interchanges the two boundaries, and hence the labels  $a$  and  $b$ . Then the invariant vacuum states are the symmetric linear combinations

$$\frac{1}{\sqrt{2}}(|0; a, b\rangle + |0; b, a\rangle) \quad (3.98)$$

This is a total of  $\frac{1}{2}N(N+1)$  states. At the first excited level one has the following invariant states

$$\frac{1}{\sqrt{2}}(\alpha_{-1}^i |0; a, b\rangle - \alpha_{-1}^i |0; b, a\rangle) \quad (3.99)$$

This is anti-symmetric, so there are  $\frac{1}{2}N(N-1)$  vector bosons. They turn out to form a gauge group  $O(N)$ .

Note that the unoriented open string has gauge bosons now. Previously we were looking at the special case  $N = 1$ , and since  $O(1)$  is trivial but  $U(1)$  is not, we found no vector bosons for the unoriented open string.

There is one additional possibility if  $N$  is even. Then, instead of defining  $\Omega |0; a, b\rangle = + |0; b, a\rangle$  the definition  $\Omega |0; a, b\rangle = - |0; b, a\rangle$  is also allowed. The vacuum state is then

$$\frac{1}{\sqrt{2}}(|0; a, b\rangle - |0; b, a\rangle) , \quad (3.100)$$

which is just a single state for the minimal choice,  $N = 2$ . At the first excited level, the states are

$$\frac{1}{\sqrt{2}}(\alpha_{-1}^i |0; a, b\rangle + \alpha_{-1}^i |0; b, a\rangle) , \quad (3.101)$$

which are  $\frac{1}{2}N(N+1)$  vector bosons. It turns out that they are gauge bosons of a symplectic group,  $Sp(N)$ .

### 3.20 Discussion

So far we have looked at free strings. That is to say, we have considered world sheets that are either an infinite cylinder or an infinite strip. Interactions are taken into account by allowing world sheets with holes, corresponding to strings splitting and joining, or with semi-infinite tubes attached to the surface (external closed string states) or semi-infinite strips attached to the boundaries (external open string states). Hence what we know so far gives only limited information, but let us look at it anyway.

The main good feature that all string theories seen so far have in common is the existence of a graviton  $g_{\mu\nu}$  in the spectrum. To be precise, we see that there is a massless symmetric tensor, which is a candidate for a graviton. To see that it really is a graviton we need to look at the interactions.

It is known, however, that there is not much choice in making a massless symmetric tensor interact in a Lorentz-covariant way. The only possibility is Einstein's theory of gravity (with possibly higher curvature corrections). Hence either string theory can be made interacting in a consistent way, and then it must produce general relativity, or it is inconsistent. The former possibility turns out to be the correct one.

For example, one may compute the three-graviton interaction in string theory and compare with the analogous computation in general relativity. As expected, the form of the interaction agrees, and one gets a relation between the coupling constant of gravity – Newton's constant – and the only parameter in the string action,  $\alpha'$ .

This is somewhat misleading since we are in 26 dimensions, and not in four. Newton's law of gravity in  $D$  dimensions is

$$F = G_N \frac{m_1 m_2}{r^{D-2}} \quad (3.102)$$

which gives the famous  $\frac{1}{r^2}$  force in 4 dimensions. It follows that the dimension of  $G_N$  is dimension dependent:

$$[G_N] = [F][r]^{D-2}[m]^{-2} = [m]^{-1}[l]^{D-1}[t]^{-2} \quad (3.103)$$

If we work with the convention  $c = \hbar = 1$ , so that  $[t] = [l] = [m]^{-1}$ , we get that  $[G_N] = [l]^{D-2}$ .

The parameter dimension of  $\alpha'$  is determined as follows.  $X$  clearly has the dimension of length, and an action has the same dimension as  $\hbar$  (so that  $\exp(iS/\hbar)$  makes sense), which is 1 in our case. The dimension of the world sheet parameters  $\sigma$  and  $\tau$  is irrelevant, since it cancels. Hence  $\alpha'$  has the dimension of  $[l]^2$ .

From this analysis we learn that it is inevitable that  $G_N$  is given by

$$G_N = (\alpha')^{\frac{1}{2}(D-2)} \times \text{dimensionless constants} \quad (3.104)$$

If we manage to construct strings in four dimensions, we would expect  $G_N \propto (\alpha')$ . Conversely, this implies the  $\sqrt{\alpha'}$  is of order the Planck length, or the fundamental mass scale  $1/\sqrt{\alpha'}$  is of order the Planck mass.

We have seen that the masses of all the excited states are multiples of this parameter, and hence inevitably all excited states have masses that are of order  $10^{19}$  GeV. These particles are therefore unobservable.

The only observable string states are the massless ones (assuming we manage to get rid of the tachyon). Of course very few massless particles have been seen: the photon, (indirectly) the graviton, and also the gluons, the vector bosons of QCD. However, the standard model has a symmetry breaking mechanism, the Higgs mechanism, and we only see the broken phase. In the unbroken theory all particles (except the scalars that become the Higgs boson) are massless.

The idea is now that this unbroken theory corresponds to the massless particles in string theory. The masses of quarks, leptons and the  $W$  and  $A$  bosons are then viewed as a small perturbation. Of course when we express the masses of quarks and leptons in terms of the Planck mass, the value “zero” is indeed an excellent approximation. Unfortunately, however, the first twenty vanishing digits of their mass are far less interesting than the first non-zero digit.

The origin of the Higgs mechanism and the smallness of the scale at which it occurs remains a mystery. The smallness of the ratio  $M_W/M_{\text{Planck}} \approx 10^{-17}$  is a mystery independent of string theory, and at present string theory offers us no solution to this problem. This small parameter problem is one of several such problems in the Standard Model. There are many mysterious small ratios among the quark and lepton masses, such as the ratio of neutrino masses (expected to lie in a range of about .1 to .001 eV) and the top quark mass (175 GeV). Sometimes they are referred to as “naturalness” problems: it seems unnatural that a dimensionless parameter has a value very different from 1. Most people expect that this a sign that there is something essential missing in our understanding. However, in some cases string theory seems hint at a different type of explanation: that a huge discrete set of values is possible, and that we just observe one of them. For some of these small parameters it is plausible that if they were very different, *i.e.* more natural and hence of order 1, that no observers would exist at all. This has brought the so-called “anthropic principle” into the discussion. It is widely disliked, but already for more than two decades all we have learned about string theory seems to point in precisely that direction.

But let us forget these naturalness problems for now, and compare the massless string spectrum to what we observe:

<b>Bosonic string spectrum</b>	<b>Observed</b>	<b>Solution</b>
Tachyon	—	Fermionic string, Superstrings
Graviton	Yes	—
$B_{\mu\nu}$	No	Must acquire mass
dilaton	No	Must acquire mass
gauge symmetry ( $U(N), O(N), Sp(N)$ )	$SU(3) \times SU(2) \times U(1)$	String theory choice
$D = 26$	$D = 4$	Compactification
—	quarks, leptons	Fermionic string

The last column lists the solutions to the mismatch in the first two columns. This is also a preview of some of the subjects we still have to discuss. Here is a very brief explanation, to be made more detailed later.

- Fermionic strings are string theories which have, in addition to the bosons  $X^\mu$ , also two-dimensional fermions. It can be shown that they always contain space-time fermions in their spectrum. Hence they can give rise to quarks and leptons. Fermionic strings have a critical dimension of 10, and they may or may not have tachyons.
- Superstrings are fermionic strings with space-time supersymmetry.
- Supersymmetry is a symmetry between fermions and bosons. So far it has not been observed in nature. Supersymmetry implies absence of tachyons (but is not necessary for that).
- Compactification means that some of the space-time dimensions are rolled up into a small space, so small that it is unobservable to us.
- “Must require mass” means that by some mechanism these particles are assumed to get a mass, at some point below the Planck mass. Just as the Standard Model Higgs mechanism may give masses to some particles that one cannot derive directly from a string spectrum at the Planck scale, there may be other such mechanisms that gives masses to other particles.



## 4 Compactification

The idea of space-time compactification precedes string theory by many decades. already in the first half of the century Kaluza, Klein and also Einstein considered the idea that space-time is higher-dimensional.

### 4.1 Space-time compactification for particles

Suppose our world is actually five-dimensional, with one dimension rolled up. This means that  $x^4$  is periodic,  $x^4 \sim x^4 + 2\pi R$ . The other four dimensions remain unconstrained. We denote the five dimensions by  $m = (\mu, 4)$  with  $\mu = 0, 1, 2, 3$ .

Now take a particle in this five-dimensional world. It satisfies the mass shell condition

$$p^2 = p^\mu p_\mu + (p_4)^2 = -M^2 \quad (4.1)$$

Now  $p_4$  is not on equal footing with the space-time momenta  $p_i$ . Wave functions must respect the periodicity:

$$\Psi(x_\mu, x_4) = \Psi(x_\mu, x_4 + 2\pi R) \quad (4.2)$$

A free particle wave function,

$$\Psi = e^{ip^m x_m} \quad (4.3)$$

is only periodic if  $p_4$  is quantized:

$$p_4 = \frac{n}{R}, \quad n \in \mathbf{Z} \quad (4.4)$$

It makes then sense to put  $(p_4)^2$  on the other side of the equal sign in (4.12). Then we get

$$p^\mu p_\mu = -(p_4)^2 - M^2 \quad (4.5)$$

Hence a four-dimensional observer sees a discrete spectrum of particles of mass-squared  $M^2 + n^2/R^2$ .

### 4.2 Space-time compactification for strings

We can do the same in string theory. If we do this for closed strings there is an additional possibility. The string can wrap a couple of times around the compact dimension:

$$X^\mu(\tau, \sigma + 2\pi) = X^\mu(\tau, \sigma) + 2\pi m R^\mu \quad (4.6)$$

where  $R^\mu = R\delta^{\mu,25}$  if we compactify the 25<sup>th</sup> dimension. Now  $\mu = 0, \dots, 24$ , so that we are assuming that we are in 25 space-time dimensions. This will turn out to be the critical dimension. The rule is that the sum of compactified plus uncompactified dimensions must be 26.

If we impose this periodicity we must re-derive the string mode expansion. This is very easy, and results in

$$X^\mu(\sigma, \tau) = q^\mu + \alpha' p^\mu \tau + m R^\mu \sigma + i \sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \left\{ \frac{1}{n} (\alpha_n^\mu e^{-in(\tau+\sigma)} + \bar{\alpha}_n^\mu e^{-in(\tau-\sigma)}) \right\}, \quad (4.7)$$

It is convenient to write this in terms of left and right-moving components

$$X_L^\mu = \frac{1}{2} q^\mu + \sqrt{\frac{\alpha'}{2}} p_L^\mu (\tau + \sigma) + i \sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \left\{ \frac{1}{n} (\alpha_n^\mu e^{-in(\tau+\sigma)}) \right\}, \quad (4.8)$$

and

$$X_R^\mu = \frac{1}{2} q^\mu + \sqrt{\frac{\alpha'}{2}} p_R^\mu (\tau - \sigma) + i \sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \left\{ \frac{1}{n} (\bar{\alpha}_n^\mu e^{-in(\tau-\sigma)}) \right\}, \quad (4.9)$$

with

$$\begin{aligned} p_L^\mu &= \frac{1}{\sqrt{2}} \left( \sqrt{\alpha'} p^\mu + \frac{1}{\sqrt{\alpha'}} m R^\mu \right) \\ p_R^\mu &= \frac{1}{\sqrt{2}} \left( \sqrt{\alpha'} p^\mu - \frac{1}{\sqrt{\alpha'}} m R^\mu \right) \end{aligned} \quad (4.10)$$

### 4.3 The spectrum

Now we quantize this theory and solve the constraints. This is all nearly analogous to the uncompactified case. The result is

$$M^2 = M_L^2 = M_R^2 \quad (4.11)$$

with

$$\begin{aligned} M_L^2 &= \frac{4}{\alpha'} \left( \frac{1}{2} (p_L^2)_{\text{comp}} + \sum_{m>0} \sum_{i=1}^{23} \alpha_{-m}^i \alpha_m^i + \sum_{m>0} \alpha_{-m}^{25} \alpha_m^{25} - a \right) \\ M_R^2 &= \frac{4}{\alpha'} \left( \frac{1}{2} (p_R^2)_{\text{comp}} + \sum_{m>0} \sum_{i=1}^{23} \bar{\alpha}_{-m}^i \bar{\alpha}_m^i + \sum_{m>0} \bar{\alpha}_{-m}^{25} \bar{\alpha}_m^{25} - a \right) \end{aligned} \quad (4.12)$$

Where, as above, only the compactified components of  $p_L$  and  $p_R$  are added to the mass. When going to light-cone gauge one now uses the components 0 and 24 to define  $X^+$  and  $X^-$ , and one checks Lorentz invariance only in the uncompactified dimensions (since it is manifestly violated in the compactified ones). The computation is essentially identical, and yields the result mentioned above, namely that the total dimension (compactified+uncompactified) must be 26. The constant  $a$  comes out as 1, as for the uncompactified string.

The quantization of the momenta in the compactified directions is exactly as for particles. Hence we get

$$\begin{aligned} p_L &= \sqrt{\frac{\alpha'}{2}} \left( \frac{n}{R} + \frac{mR}{\alpha'} \right) \\ p_R &= \sqrt{\frac{\alpha'}{2}} \left( \frac{n}{R} - \frac{mR}{\alpha'} \right) \end{aligned} \tag{4.13}$$

Here  $p_L$  and  $p_R$  refer to the 25<sup>th</sup> components.

The spectrum again contains a tachyon. At the massless level we now have

$$\begin{aligned} \alpha_{-1}^i \bar{\alpha}_{-1}^i |0, 0\rangle &\longrightarrow g, B, \phi \\ \alpha_{-1}^i \bar{\alpha}_{-1}^{25} |0, 0\rangle &\longrightarrow \text{vector boson} \\ \alpha_{-1}^{25} \bar{\alpha}_{-1}^i |0, 0\rangle &\longrightarrow \text{vector boson} \\ \alpha_{-1}^{25} \bar{\alpha}_{-1}^{25} |0, 0\rangle &\longrightarrow \text{scalar} \end{aligned} \tag{4.14}$$

Here we use the notation  $|p_L, p_R\rangle$ . In addition it may happen that  $(p_L)^2$  and/or  $(p_R)^2$  is precisely equal to 2. In that case there are extra massless particles:

$$\begin{aligned} \alpha_{-1}^i |0, p_R\rangle &\longrightarrow \text{vector boson} \\ \alpha_{-1}^{25} |0, p_R\rangle &\longrightarrow \text{scalar} \\ \bar{\alpha}_{-1}^i |p_L, 0\rangle &\longrightarrow \text{vector boson} \\ \bar{\alpha}_{-1}^{25} |p_L, 0\rangle &\longrightarrow \text{scalar} \\ |p_L, p_R\rangle &\longrightarrow \text{scalar} \end{aligned} \tag{4.15}$$

It turns out that the vector bosons are in fact gauge bosons of some new gauge interaction. This is a second way to obtain gauge symmetry in addition to Chan-Paton labels.

## 4.4 T-duality

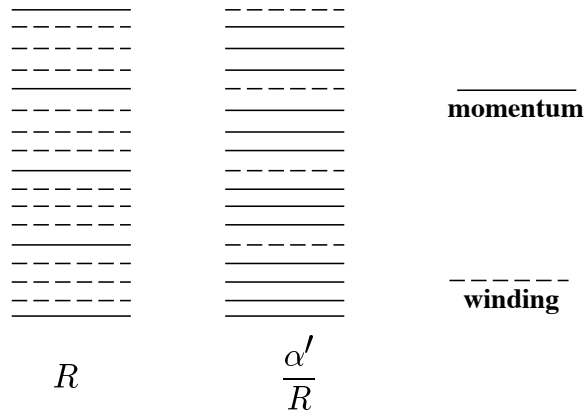
The spectrum we just found has an interesting symmetry

$$\begin{aligned} R &\longleftrightarrow \frac{\alpha'}{R} \\ n &\longleftrightarrow m \end{aligned} \tag{4.16}$$

Under this transformation  $p_L$  is unchanged and  $p_R$  changes sign, but this has no influence on the spectrum.

It turns out that this remains true even for the interacting string theory. This is remarkable, because it says that strings compactified on a very large circle are equivalent to strings compactified on a very small circle. Or, in other words, one cannot do any experiment to determine if the internal compactification manifold is large or small. This

is counter-intuitive, but there is a simple reason why this happens: the momentum and winding states are interchanged. At small  $R$ , the momentum states are very widely spaced, but it costs very little energy to wind the string around the circle. Therefore winding states have very little spacing. If we replace  $R$  by  $\alpha'/R$  it is just the other way around. Now momentum states take over the role of winding states and vice-versa. Experimentally winding and momentum states cannot be distinguished.



If you take this to the extreme limit  $R \rightarrow \infty$  it leads to the conclusion that flat space is equivalent to a point! It is however not very clear what this would mean.

This phenomenon is known as  $T$ -duality. Although we have seen it here in a very special system, it is seen in all compactifications, and is believed to be a fundamental property of string theory. It seems as if short distances (smaller than the Planck scale) are equivalent to long distances. In other words, there seems to be a fundamental shortest length scale, the Planck scale. If one probes smaller distances, nothing new is found, but one simply sees the same phenomena one has seen at larger distances already.

This is also related to the absence of short-distance singularities (“ultra-violet divergences”) in string loop calculations. What happens is that all integrals are cut off in a natural way at the Planck scale. Integrating over momenta larger than the Planck momentum would be double counting: those contributions already are taken into account in the integral over small momenta.

## 4.5 The self-dual point

We observed earlier that there is a possibility of having extra massless vector bosons. This requires  $p_L^2 = 2, p_R^2 = 0$  or vice versa. The condition  $p_R = 0$  relates  $m$  to  $n$ . Substituting this in  $p_L^2 = 2$  leads to the condition  $R = |n|\sqrt{\alpha'}$ . Then we may compute  $m$ , and we find  $m = \frac{1}{n}$ . Hence the only possibility is  $n = m = \pm 1, R = \sqrt{\alpha'}$ . Note that the  $T$ -dual of this theory is the same theory. This is called therefore called the *self dual point* on the space of compactified string theories. In the self-dual point there are precisely two states of the form  $\alpha_{-1}^i |0, p_R\rangle$  (namely  $(n, m) = (1, -1)$  or  $(-1, 1)$ ) and two of the form  $\bar{\alpha}_{-1}^i |p_L, 0\rangle$  (namely  $(n, m) = (1, 1)$  or  $(-1, -1)$ ).

One may expect these states to correspond to gauge bosons. Then what is the corresponding gauge group? If we are not in the self-dual point there are just two vector bosons, and then there is not much choice; it can only be  $U(1) \times U(1)$ . In the self-dual point there are six vector bosons, and then the only possibilities are  $U(1)^6$ ,  $SU(2) \times U(1)^3$  or  $SU(2)^2$ . The latter turns out to be the correct one, but to decide that one would have to study the interacting string.

## 4.6 Field theory interpretation\*

The appearance of massless gauge bosons and scalars in the spectrum can also be understood as follows. We have seen that before compactification the spectrum contained a graviton  $g_{\mu\nu}$  (*i.e.* a symmetric tensor), an antisymmetric tensor  $B_{\mu\nu}$  and a dilaton  $\phi$  as massless particles. Massive particles in 26 dimensions can be ignored: their momentum and winding modes will be even more massive. The tachyon *can* contribute to the massless spectrum, due to its momentum and winding modes. These are the extra, non-generic massless states found above.

The presence of the massless particles in the string spectrum implies the existence of a corresponding quantum field theory in target space. This is a rather subtle step. Up to now we have worked with a single string theory, and we have creation operators that create modes on the string. This is not the same as particle creation in quantum field theory. Formally it is the same operation if we look at the string theory as a two-dimensional quantum field theory. But that is world sheet physics. In target space (*i.e.* in 26 dimensions) a quantum field theory has similar creation operators, but they create particles. The operators  $\alpha_{-m}^i$  don't do that; they just change the vibrations of the string moving through target space. The vacuum state  $|0\rangle$  is a two-dimensional vacuum state. It corresponds to a string without any vibrations, not to an empty universe. In other words, what is missing is an operator that create strings from the target space vacuum. Attempts to construct such operators have been partly successful, and have led to *string field theory*. This is a difficult subject we do not want to go into here.

Usually one bypasses this issue, and one simply postulates that for every particle in the string spectrum there will be a multi-particle Hilbert space, and that the interactions that create or destroy such particles can be described by means of a quantum field theory.

Although it is hard to derive this quantum field theory directly from string theory, we can learn enough about it to determine it. First of all one can determine the classical equation of motion from string theory. This already determines the classical action almost completely. Secondly one can compute multi-particle scattering amplitudes (S-matrix elements) directly from string theory, and compare with a putative quantum field theory.

The presence of a massless spin-2 particle then implies that there must be a corresponding field theory. This is in fact general relativity. Likewise, a massless spin-1 particle implies a gauge theory.\* These field theories are usually called *low-energy effective field theories*. The word “low-energy” refers to the fact that we are ignoring massive string

---

\* The terminology “spin-2” and “spin-1” is strictly four-dimensional. In general it should be interpreted as “symmetric tensor” and “vector”.

excitations. It might be possible to include them in the field theory, but since there are infinitely many that is not practical. Since massive excitations are ignored, any field theory description must fail once we reach energies as big as the string scale (*i.e.* the Planck scale). The word “effective” means that it is a good substitute for the real thing, which is string theory. For example, one can have an effective theory of pion interactions, which is not as fundamental as QCD but can give good results in certain approximations.

Once we have made this step, we can try to interpret what we have found in terms of the effective low energy field theory. The relevant fields are the graviton, the anti-symmetric tensor and the dilaton. These massless fields can be decomposed in terms of 25-dimensional fields. This works as follows:

$$g_{\mu\nu} \rightarrow \begin{pmatrix} g_{MN} & A_M \\ A_M & \xi \end{pmatrix} \quad (4.17)$$

$$B_{\mu\nu} \rightarrow \begin{pmatrix} B_{MN} & B_M \\ -B_M & 0 \end{pmatrix} \quad (4.18)$$

Here  $\mu = 0 \dots 25$  and  $M = 0 \dots 24$ . We see that the graviton decomposes into the 25-dimensional graviton plus a vector boson and a scalar, whereas  $B_{\mu\nu}$  yields  $B_{MN}$  plus another vector. Hence we get two massless vector bosons and a massless scalar  $\xi$ , in agreement with what we found. The dilaton in 26 dimensions just yields the 25-dimensional dilaton. On top of this we get all the momentum and winding excitations of these particles.

In a Kaluza-Klein compactification one only has the momentum modes of string theory. They can be seen in the effective field theory by a mode-expansion in the compactified coordinates, *e.g.*

$$\xi(x^M, x^{25}) = \sum_n \xi_n(x^M) \cos(2\pi n x^{25}/R) + \sum_n \varphi_n(x^M) \sin(2\pi n x^{25}/R) \quad (4.19)$$

The fields  $\xi_n(x^M)$  and  $\varphi_n(x^M)$  ( $n \neq 0$ ) are 25-dimensional massive scalar fields, and  $\xi_0$  is a massless field. The winding modes of string theory however can never be found in this way. Here we see the limitations of effective field theory.

## 4.7 Compactification of more dimensions

The foregoing can easily be generalized to more compactified dimensions. The generalization of a circle is a torus. However, rather than trying to imagine a higher dimensional manifold, one usually prefers an algebraic description.

A circle was specified earlier by means of a single vector  $R^\mu$ . The appropriate generalization is to introduce  $N$  such vectors (linearly independent, of course) if you want to compactify  $N$  dimensions. The meaning of this set of vectors is that points in space-time that differ by any such vector are considered identical. If this identification is done in one direction, space gets rolled up into a cylinder. If we add a second direction, the cylinder gets rolled up into a torus. Beyond this we cannot even draw the resulting figure anymore, but it is not hard to describe mathematically.

Obviously, if  $x^\mu$  is the same point as (*i.e.* is identified with)  $x^\mu + 2\pi R_1^\mu$ , but also the same point as  $x^\mu + 2\pi R_2^\mu$ , and if this is true for any  $x^\mu$ , then it follows that  $x^\mu$  is identified with  $x^\mu + 2\pi R_1^\mu + 2\pi R_2^\mu$ . We see then that the set of vectors  $R_i^\mu$  must close under addition. Such a set of vectors is called a lattice.

A lattice can always be described by choosing a set of basis vectors  $\vec{e}_i, i = 1 \dots N$ , such that any vector  $\vec{\lambda}$  on the lattice  $\Lambda$  can be written as

$$\vec{\lambda} = \sum_{i=1}^N n_i \vec{e}_i \quad (4.20)$$

We say that space is compactified on (the torus corresponding to) this lattice if all points in that space that differ by lattice vectors are identified:  $x^I$  and  $x^I + 2\pi\lambda^I$  are regarded as the same point, for all  $\vec{\lambda} \in \Lambda$ . Here  $I = 1 \dots N$  labels the compactified dimensions. The lattice basis vectors  $\vec{e}_i$  have components  $e_i^I$ .

If we study plane waves on such a space, then the corresponding momenta must be such that this identification is respected. Hence

$$e^{i\vec{p}\cdot\vec{x}} = e^{i\vec{p}\cdot(\vec{x}+2\pi\vec{\lambda})} \quad \text{for all } \vec{\lambda} \in \Lambda$$

We see therefore that allowed momenta must have integral inner product with all these lattice vectors in order for wave functions to have to correct periodicity. A necessary and sufficient condition for this is

$$\vec{p} \cdot \vec{e}_i \in \mathbf{Z} \quad \text{for all } i \quad (4.21)$$

This is easily seen to imply that the momenta  $p$  must themselves lie on a lattice, the so-called *dual lattice*  $\Lambda^*$ . (This situation also occurs in solid state physics, where such a lattice is called *reciprocal lattice*; however string theorists like the word “dual”).

Now we put closed strings in this torus compactified space. As before, they can wind around the torus, which is taken into account by a term  $\vec{\lambda}\sigma$  in the string mode expansion (4.7), instead of the term  $mR^\mu\sigma$ . So for each choice of the vector  $\lambda$  we have closed strings winding in a different way around the multi-dimensional torus, and for each such choice of windings, we expand  $X^I$  in a different way.

In this general situation the mass formula stays essentially the same, but  $p_L$  and  $p_R$  are higher-dimensional vectors of the form

$$\begin{aligned} \vec{p}_L &= \sqrt{\frac{\alpha'}{2}} \left( \vec{\rho} + \frac{1}{\alpha'} \vec{\lambda} \right) \\ \vec{p}_R &= \sqrt{\frac{\alpha'}{2}} \left( \vec{\rho} - \frac{1}{\alpha'} \vec{\lambda} \right) \end{aligned} \quad (4.22)$$

with  $\vec{\lambda} \in \Lambda$  and  $\vec{\rho} \in \Lambda^*$ . The mass formula is usually written as

$$\begin{aligned}\frac{1}{4}\alpha' M_L^2 &= \frac{1}{2}p_L^2 + \mathcal{N} - 1 \\ \frac{1}{4}\alpha' M_R^2 &= \frac{1}{2}p_R^2 + \bar{\mathcal{N}} - 1 \\ M_L^2 &= M_R^2,\end{aligned}\tag{4.23}$$

where  $p_L$  and  $p_R$  are always the compactified components (for uncompactified ones the left-right splitting is not very useful anyway), and  $\mathcal{N}$  is the sum over the “number operators” of the uncompactified and compactified oscillators, *i.e.*

$$\mathcal{N} = \sum_{i=1}^{D-2} \sum_{m>0} \alpha_{-m}^i \alpha_m^i + \sum_{I=1}^N \sum_{m>0} \alpha_{-m}^I \alpha_m^I,\tag{4.24}$$

and analogously for  $\bar{\mathcal{N}}$ . Here  $D$  is the number of uncompactified dimensions. the sum  $D + N$  is restricted to the value 26. This can be established, as before, by checking the Lorentz algebra in the  $D$  uncompactified dimensions, which only depends on the number of oscillators, and not on the kind of index  $i$  or  $I$  they carry.

## 4.8 Narain lattices\*

However, there is a more general kind of torus compactification of string theory\*. Define a lattice  $\Gamma$  consisting of vectors of the form  $(\vec{p}_L, \vec{p}_R)$ . We observe that this lattice has the following properties

$$\vec{p}_L \cdot \vec{p}'_L - \vec{p}_R \cdot \vec{p}'_R \in \mathbf{Z}\tag{4.25}$$

Lattices with this property are called *Lorentzian integral*. The adjective “Lorentzian” refers to the fact that there is a minus sign: this can be reproduced by introducing a Lorentzian metric  $\text{diag}(1, \dots, 1, -1, \dots, -1)$ , with  $N$  positive and  $N$  negative entries. We then treat  $(\vec{p}_L, \vec{p}_R)$  as a  $2N$  dimensional vector in a space with this metric. The lattice vectors also satisfy

$$\vec{p}_L^2 - \vec{p}_R^2 \in 2\mathbf{Z}\tag{4.26}$$

Such a lattice is called *Lorentzian even* (actually this already implies Lorentzian integral, but not the other way around). Finally we can define a Lorentzian dual to a lattice: the Lorentzian dual  $\Gamma^*$  contains all vectors that have integral inner product with all vectors in  $\Gamma$ , where we use the Lorentzian inner product defined above.

One can easily show that the lattice  $\Gamma$  constructed above has the property that  $\Gamma = \Gamma^*$ . Such a lattice is called a Lorentzian self-dual lattice. We see thus that  $\Gamma$  is a Lorentzian even self-dual lattice. A compactification on the corresponding torus is called a *Narain*

---

\* The compactification discussed here is the general solution to the condition of “modular invariance”, which will be discussed in section 5.



*compactification*. It can be shown (and we will return to this in chapter 5) that any such compactification yields a consistent string theory.

The point is now that there exist lattices that are Lorentzian even self-dual, but are not of the form (4.22). The generalization has in fact the following form (here we denote the components of the vectors with an index  $I = 1, \dots, N$ ; the original 26-dimensional index  $\mu$  is split into  $(M, I)$ , with the uncompactified space-time index  $M$  running from 0 to  $26 - N - 1$ )

$$\begin{aligned} p_L^I &= \sqrt{\frac{\alpha'}{2}} \left( \rho^I + \frac{1}{\alpha'} (\lambda^I - B^{IJ} \lambda_J) \right) \\ p_R^I &= \sqrt{\frac{\alpha'}{2}} \left( \rho^I - \frac{1}{\alpha'} (\lambda^I + B^{IJ} \lambda_J) \right), \end{aligned} \quad (4.27)$$

where  $B^{IJ}$  is an arbitrary anti-symmetric matrix. Upper and lower indices have no special significance here.

It is interesting to count parameters. The anti-symmetric matrix  $B^{IJ}$  has  $\frac{1}{2}N(N-1)$  real parameters. The lattice  $\Lambda$  we started with is completely specified by the lengths and relative angles of the basis vectors. This defines a matrix

$$g_{ij} = \sum_I e_i^I e_j^I \quad (4.28)$$

This is a real, symmetric matrix, which has  $\frac{1}{2}N(N+1)$  parameters.

## 4.9 Background fields\*

### 4.9.1 The metric

The notation used in the previous section suggests already that  $g^{IJ}$  and  $B^{IJ}$  have something to do with the metric  $g_{\mu\nu}$  and the anti-symmetric tensor field  $B_{\mu\nu}$ . This is indeed the case. To see this, note that  $X^I$  has periodicities  $X^I \sim X^I + 2\pi\lambda^I = X^I + 2\pi e_i^I n_i$ . We can take the information about the sizes and shape of the lattice out of  $X^I$  by defining

$$X^I = \sum_i e_i^I \hat{X}^i \quad (4.29)$$

Then  $\hat{X}_i$  has standard, rectangular periodicities  $\hat{X}^i \sim X^i + 2\pi n^i$ . If we do this in the action, the following happens

$$S[X] = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \eta^{ab} \partial_a X^I \partial_b X^J \delta_{IJ} = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \eta^{ab} \partial_a \hat{X}^i \partial_b \hat{X}^j g_{ij} \quad (4.30)$$

Now all the structure is encoded in the matrix  $g_{ij}$ , and if we vary the parameters the periodicities of  $\hat{X}_i$  do not change at all. We saw (see (2.21)) that we can write down the

string action in such a way that it is valid in an arbitrary curved space. Here we see the simplest example. In general, the properties of such a space are encoded in the metric  $g_{\mu\nu}(x)$ . In the case under consideration here the metric does not depend on  $x$ , but in the compactified part of space it is a non-trivial constant matrix.

#### 4.9.2 The anti-symmetric tensor

What about  $B^{IJ}$ ? The string action written down so far does not have any room for these parameters, but it is possible to write down a string action that is valid in the presence of a non-trivial background  $B_{\mu\nu}$ . This action is

$$S[X] = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \left\{ \eta^{ab} \partial_a \hat{X}^i \partial_b \hat{X}^j g_{ij} + \epsilon^{ab} \partial_a \hat{X}^i \partial_b \hat{X}^j B_{ij} \right\} , \quad (4.31)$$

where  $\epsilon^{ab}$  is the two-dimensional Levi-Civita tensor:  $\epsilon^{10} = -\epsilon^{01} = 1$ ;  $\epsilon^{00} = \epsilon^{11} = 0$ . Because of the anti-symmetry of  $B$  we clearly need an anti-symmetric tensor here. We can write this action back into the old basis, with a diagonal space-time metric. The advantage of that basis is that it is more convenient for computing the spectrum, and indeed, that is the basis we used for previous calculations. If we define

$$B_{ij} = \sum_{IJ} e_i^I e_j^J B_{IJ} \quad (4.32)$$

we get

$$S[X] = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \left\{ \eta^{ab} \partial_a X^I \partial_b X^J \delta_{IJ} + \epsilon^{ab} \partial_a X^I \partial_b X^J B_{IJ} \right\} , \quad (4.33)$$

It may seem that this new term forces us to start from the beginning with the derivation of the equations of motion and the whole quantization procedure. But (for constant  $B_{IJ}$  at least) the effect is minimal. The extra term in the equations of motion is proportional to  $\partial_a \epsilon^{ab} \partial_b X^I$ . But this vanishes since the two derivatives commute.

The reason this term vanishes is that the extra term in the action is a total derivative:

$$\epsilon^{ab} \partial_a X^I \partial_b X^J B_{IJ} = \partial_a \left[ \epsilon^{ab} X^I \partial_b X^J B_{IJ} \right] \quad (4.34)$$

When the the action is integrated over a surface without boundaries the contribution of such a term vanishes. Therefore it also does not affect the equations of motion, which describe variations of the action.

But then one may wonder how it can have any effect at all. The reason is that the canonical momenta *do* change. We find

$$\Pi_I = \frac{1}{2\pi\alpha'} (\delta_{IJ} \partial_0 + B_{IJ} \partial_1) X^J \quad (4.35)$$

It is thus this quantity that satisfies the canonical quantization rule (3.17).

It is not difficult to see that the extra term does not affect the quantization of the oscillator modes, but there is a change in the commutators of the zero modes  $q$  and  $p$ .

The effect of the extra term is, using (4.7), that  $p^I$  is replaced by  $\pi^I = p^I + \frac{1}{\alpha'} B^{IJ} \lambda_J$  ( $\vec{\lambda}$  is the higher dimensional generalization of  $mR^\mu$  in (4.7).

The expression for the mass spectrum is still (4.23), with  $p_L^I$  expressed in terms  $p^I$  and  $\lambda^I$  exactly as before. But the canonical momentum is  $\pi^I$ , not  $p^I$ , and hence the quantization condition is that the eigenvalues of the operator  $\pi^I$  must lie on the dual lattice. Then it is more useful to express  $p_L^I$  and  $p_R^I$  in terms of the canonical momenta. This gives

$$\begin{aligned} p_L^I &= \frac{1}{\sqrt{2}} \left( \sqrt{\alpha'} p^I + \frac{1}{\sqrt{\alpha'}} \lambda^I \right) = \frac{1}{\sqrt{2}} \left( \sqrt{\alpha'} \pi^I + \frac{1}{\sqrt{\alpha'}} (\lambda^I - B^{IJ} \lambda_J) \right) \\ p_R^I &= \frac{1}{\sqrt{2}} \left( \sqrt{\alpha'} p^I - \frac{1}{\sqrt{\alpha'}} \lambda^I \right) = \frac{1}{\sqrt{2}} \left( \sqrt{\alpha'} \pi^I - \frac{1}{\sqrt{\alpha'}} (\lambda^I + B^{IJ} \lambda_J) \right) \end{aligned} \quad (4.36)$$

If we now put the eigenvalues of  $\pi$  on the dual lattice we do indeed recover (4.27).

## 4.10 Moduli\*

Combining what we saw in section 3.6 and in the previous one, we arrive now at the following picture. In a compactification to  $D = 26 - N$  dimensions, the 26-dimensional metric decomposes as follows

$$g_{\mu\nu} \rightarrow \begin{pmatrix} g_{MN} & A_M^I \\ A_M^I & \xi^{IJ} \end{pmatrix} \quad (4.37)$$

$$B_{\mu\nu} \rightarrow \begin{pmatrix} B_{MN} & B_M^I \\ -B_M^I & \zeta^{IJ} \end{pmatrix} \quad (4.38)$$

Here  $A_M^I$  and  $B_M^I$  are two sets of  $N$  vector bosons each, and  $\xi^{IJ}$  and  $\zeta^{IJ}$  are two sets of  $\frac{1}{2}N(N+1)$  and  $\frac{1}{2}N(N-1)$  scalars. All these quantities are fields in  $D = 26 - N$  dimensions, *i.e.* they are functions of the coordinate  $x^M$ . They are not functions of the compactified coordinates, however. The dependence on the compactified coordinate  $y^I$  has been mode expanded, as in the case  $N = 1$  discussed above.

We observe now that the parameters of the compactification can be viewed as *vacuum expectation values* of the massless scalars in  $D = 26 - N$  dimensions.

In general, a classical field theory is quantized by expanding around a certain classical background:  $\phi = \phi_{\text{cl}} + \phi_{\text{q}}$ . The quantum field is expanded in modes, and the vacuum state is chosen in such a way that  $\langle 0 | \phi_{\text{q}} | 0 \rangle = 0$ , so that  $\langle 0 | \phi | 0 \rangle = \phi_{\text{cl}}$ . Then one calls  $\phi_{\text{cl}}$  the vacuum expectation value of  $\phi$ . Also the term *background field* is often used for  $\phi_{\text{cl}}$ .

If we want a theory to respect certain symmetries, then we must require that at least the classical values of the fields,  $\phi_{\text{cl}}$ , respect it. For example translation invariance requires that  $\phi_{\text{cl}}$  is constant. Lorentz-invariance requires that a vector cannot have a vacuum expectation value, but a scalar can (whereas for a symmetric tensor the v.e.v. must be proportional to  $\eta_{\mu\nu}$ ).

A scalar field can get a vacuum expectation value without breaking Poincaré invariance. This is precisely what is occurring here. We may expand the scalar fields in  $D$  dimensions as follows

$$\begin{aligned}\xi^{IJ} &= e_i^I e_j^J g_{ij} + \xi_{\text{qu}}^{IJ} \\ \zeta^{IJ} &= B^{IJ} + \zeta_{\text{qu}}^{IJ}\end{aligned}\tag{4.39}$$

We see now that all parameters of the torus compactification can be interpreted as vacuum expectation values of scalar fields.

This is not limited to torus compactification. It is often expressed by saying that *string theory does not have free parameters*. All freedom is due to changing the vacuum. There is no parameter one fixes “by hand”, like the quark masses or the standard model couplings. In more realistic compactified string theories, the quark and lepton masses (or rather the Yukawa couplings that determine them) are indeed determined by vacuum expectation values of similar scalar fields. The same would be true for all 28\* parameters of the Standard Model.

Note that in the Standard Model these parameters are input. They are not determined by anything we know; we can measure them and express physical quantities in terms of them, but Quantum Field Theory gives us no clue about their origin. If we are able to get the Standard Model from string theory, *all* its parameters are functions of vacuum expectation values of scalar fields. If the parameters are related to v.e.v.’s of fields, there is a hope that they can be determined dynamically. Maybe one set of parameters is energetically favorable, or at least a local minimum of some potential. If a parameter is in the Lagrangian one starts with, there is no hope to determine it from the theory itself.

If one studies interacting strings, a novel quantity appears that looks like a parameter, namely the coupling constant that determines the value of the three-string vertex. But it turns out that even this string coupling constant is not a parameter, but related to the v.e.v. of the dilaton. One might object that  $\alpha'$  is a parameter, but that is not true. It is a quantity with a dimension, and changing it just changes all dimensionful quantities in the same way. If we make *all* the masses in our universe twice as heavy, and also change Newton’s constant accordingly, this can never be observed. Only dimensionless ratios are true parameters.

The continuous parameters that relate different compactified strings are often called *moduli*. The name comes from mathematics, where it is used for parameters that describe deformations of a manifold. The space formed by the moduli is called *moduli space*. For circle compactifications in field theory it is just a half-line  $R \geq 0$ , corresponding to the value of the radius  $R$ . In string theory it gets a more interesting structure due to T-duality, which identifies two parts of the half-line. But even if one compactifies just one dimension in string theory there is already more structure.

---

\* Assuming three right-handed neutrinos

## 4.11 Vacuum selection\*

What determines the vacuum? In scalar field theories one often has a potential, *i.e.*

$$\mathcal{L} = \frac{1}{2} [-\partial_\mu\varphi(\vec{x}, t)\partial^\mu\varphi(\vec{x}, t) - m^2\varphi(\vec{x}, t)^2] - V(\varphi) . \quad (4.40)$$

For constant fields  $\varphi$  only the mass term and the potential are relevant, and the Hamiltonian density for constant fields is

$$\mathcal{H} = \frac{1}{2} [m^2\varphi(\vec{x}, t)^2] + V(\varphi) \quad (4.41)$$

The Hamiltonian is the space-integral of this quantity. Since for constant  $\varphi$  the integrand is independent of  $x$ , the integral is infinite or at best zero, but if we cut off the infinity by putting the system in a box, we can compare different choice of the v.e.v. of  $\varphi$  by comparing the resulting energy. Presumably the true vacuum is then the one of minimal energy.

For a typical potential the equation of motion only allow a finite number of constant solutions. Those solutions are simply the minimum of potential plus mass-term.

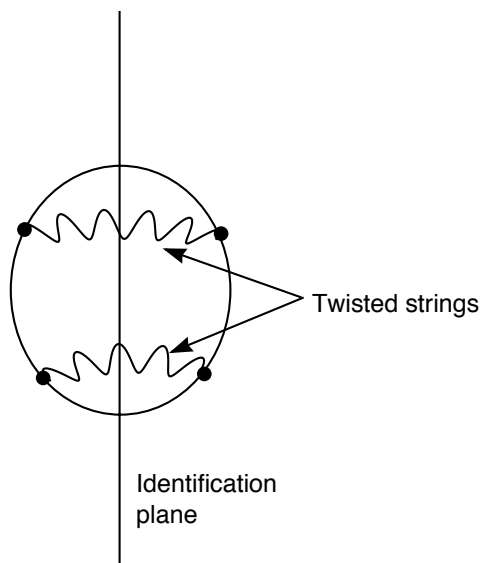
However, there is a problem with this point of view. We have a tachyon  $T$ , which is a scalar with  $m^2 < 0$ . This means that the point  $\langle T \rangle = 0$  is an instable point in the tachyon potential. If we knew the full potential  $V(T)$  we could try to determine the correct minimum, but we do not know it. The tachyon couples to all other particles, and in its presence it makes no sense to try to discuss the minima of the moduli potentials.

This will all be different in superstrings. There the moduli are in fact flat directions in the potential (and this can be shown also in the presence of interactions), and there is no tachyon. Hence we get an enormous vacuum degeneracy: there is a huge set of possible string compactifications, and no principle that selects one over any other. The only hope is then that some “non-perturbative” effect will lift the degeneracy. This has become known as the problem of moduli stabilization: to show that the flat directions are lifted by corrections to the lowest order approximation, and that a non-trivial potential with local minima is generated. Everything we know suggests that if indeed this is possible, then the number of such local minima will be huge. Numbers as large as  $10^{500}$  have been given as estimates in certain cases. This would imply as many choices for the scalar field vacuum expectation values that determine the Standard Model parameters.

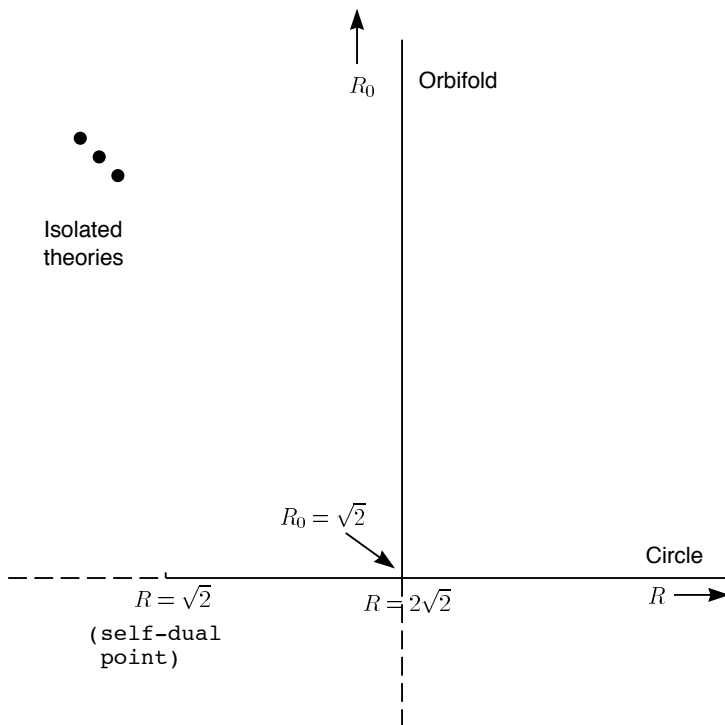
## 4.12 Orbifolds\*

There are many other compactifications one might consider. A very popular one is the *orbifold*. We will just sketch the procedure for the one-dimensional case. We start with the compactification on a circle with radius  $R$ , and we declare that all points related by a reflection are identified. This imposes first of all a condition on the closed string states. Only string states that are invariant under  $X \rightarrow -X$  can live in such a world. It is easy to see that this eliminates, for instance, all states created by an odd number of oscillators.

In addition one now has to allow closed strings that begin and end at identified points. This gives new states in the Hilbert space. This part of the Hilbert space is called the *twisted sector*.



Since this can be done for any circle radius, one obtains a second infinite half-line of solutions, with again a T-duality relating small  $R$  to large  $R$ . For one compactified dimension (relevant for 25-dimensional strings) the full moduli space is known, and there are two surprises. They can be seen in the following picture (here  $N = R\sqrt{\frac{2}{\alpha'}}$ ).



The first is that the orbifold line and the circle line meet each other. This means that an orbifold theory (in fact the orbifold of the self-dual circle theory) is in fact equivalent to a circle theory at a different radius. This is a non-trivial fact that can only be seen in string theory (indeed, orbifolds are singular as manifolds, and a field theory compactification on such a manifold is not even well-defined). It is again a kind of duality: two seemingly different theories are in fact equivalent. The second surprise is the existence of three isolated theories that have no moduli at all. These were discovered by P. Ginsparg.

## 5 Fermions on the world sheet

### 5.1 Generalizations of the bosonic string

When we started with the bosonic string, we had a nice intuitive action, which was just the surface area of the world sheet. This Nambu-Goto action was then shown to be classically equivalent to a different action, which was the action of  $D$  free bosons. We saw that this action had Poincaré invariance in target space, and reparametrization invariance and Weyl invariance on the world sheet.

The basic idea underlying all other string construction methods is to modify this action to some other two-dimensional action, but in such a way that the essential symmetries are preserved. Poincaré invariance is clearly essential, but only in the  $D$  dimensions we are interested in (which ultimately will be  $D = 4$ ). Weyl invariance and reparametrization invariance (and conformal invariance, a subset of these symmetries) clearly played an important rôle in the discussion of the foregoing sections, although the light-cone gauge derivation of the spectrum does not really show very clearly that these symmetries are *essential*. But they are, in the sense that all important results – and in particular the appearance of gravity – remain true provide we make sure that these symmetries are present in the theory.

Often one does not write explicitly reparametrization invariant theories but instead one works directly in conformal gauge, and demands that the theory has conformal invariance.

An rather trivial example is the compactified string. The world sheet action is\*

$$S[X, \gamma] = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau \eta^{ab} [\partial_a X^\mu \partial_b X_\mu + \partial_a X^I \partial_b X_I] \quad (5.1)$$

Since  $X^I$  lives on a torus and  $X^M$  does not, 26-dimensional Poincaré invariance is broken to  $D$ -dimensional Poincaré invariance. But the theory still has conformal invariance (and it can be written in a reparametrization invariant way by re-introducing a world-sheet metric  $\gamma_{\alpha\beta}$ ).

In this example the compactified part of the theory (often called the *internal theory*) is still written in terms of free bosons  $X^I$ , but this is not necessary. Just as in four dimensions, there are other field theories than those made out of free bosons. In general any two-dimensional field theory with conformal invariance is acceptable. This kind of theory is called a *conformal field theory*, usually abbreviated as CFT. Such theories exist also in more than two dimensions, but in particular in two dimensions a lot is known, and many such theories can be constructed. Hence in general one has

$$S[X, \gamma] = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau [\eta^{ab} \partial_a X^\mu \partial_b X_\mu + \text{some conformal field theory}] \quad (5.2)$$

---

\* From now on we use indices  $\mu$  and  $\nu$  for  $D$ -dimensional space-time; There will be no need anymore for an index that covers compactified and uncompactified components simultaneously.



## 5.2 The rôle of the conformal anomaly\*

There is one important restriction on this internal CFT, which is the generalization of the restriction for free bosons  $X^\mu$  to 26 dimensions. To each conformal field theory belongs a number, the *conformal anomaly*. It can be computed by means of a two-dimensional one-loop quantum correction, which will not be explained here.

Each conformal field theory has an energy-momentum tensor that can be expanded in modes  $L_n$  and  $\bar{L}_n$ , just as we did for the bosonic string. Classically these objects satisfy

$$\{L_m, L_n\}_{\text{PB}} = i(m-n)L_{m+n} . \quad (5.3)$$

It can be shown that this relation is equivalent to having classical conformal invariance. One can work out the commutator of two such operators in the quantum theory, and in general the result is

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}(m^3 - m)\delta_{m+n} . \quad (5.4)$$

The first term is what one expects on the basis of classical/quantum correspondence. We have seen before that if one allows an extra term in the algebra, the central charge term, it must have the form shown in (5.4)\*. The extra term has no classical correspondence, and can be interpreted as a violation of the classical symmetry by quantum effects. Such an effect is generally called an *anomaly*. Since the symmetry affected by it is the conformal symmetry,  $c$  is often called the conformal anomaly. It may also be thought of as counting the number of degrees of freedom of a theory (although it is not necessarily integer). Indeed, for a free boson theory with  $N$  bosons the conformal anomaly is  $N$ . The consistency requirement is now that the total conformal anomaly must be 26, *i.e.*

$$D + c_{\text{int}} = 26 , \quad (5.5)$$

where  $c_{\text{int}}$  is the conformal anomaly of the internal CFT.

The best way to understand this, although this falls outside the scope of these lectures, is as follows. In a path integral quantization of string theory one has to perform a gauge fixing procedure. This introduces extra fields, the so-called Fadeev-Popov ghosts. It turns out that these fields also contribute to the conformal anomaly, and their contribution is  $-26$ . Hence if we get  $+26$  from the other fields in the theory, the space-time coordinate bosons  $X^\mu$  and the internal CFT, then the total conformal anomaly is zero. Then exact conformal invariance holds for the quantum theory. This is the condition one needs.

Note that it is not very clear that theories constructed in this way can be thought of as space-time compactifications of a 26-dimensional theory. In general, we do not have bosons  $X^I$  as ‘remnants’ of 26-dimensional bosons. In many cases<sup>†</sup> the resulting theories can be viewed as compactifications on non-flat manifolds (a torus is flat), but this is not essential. The nice world-sheet surface interpretation of the string action will also be more and more lost.

---

\* More accurately: one requires the vacuum to satisfy  $L_n|0\rangle = 0$  for  $n \geq -1$ . Then in particular  $L_{-1}|0\rangle = 0$  so that  $A(1) = 0$ . <sup>†</sup> In fact this has not been studied extensively for bosonic strings, but the remark is based on experience with fermionic strings.

### 5.3 Internal free fermions

In the spirit of the foregoing discussion one may, as a first attempt to build a new theory, consider fermionic theories. We do this directly in conformal gauge (coupling free fermions to non-trivial metric  $\gamma_{ab}$  is possible, but it is a technical complication that is not going to give any additional insight). The action is

$$S = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau [\partial_a X^\mu \partial^a X_\mu - i\bar{\psi}^A \rho^a \partial_a \psi^A] \quad (5.6)$$

with an implicit summation over  $A$ . Here  $\rho^a$  are the two-dimensional Dirac matrices, the equivalent of  $\gamma^\mu$  in four dimensions. These matrices satisfy

$$\{\rho^a, \rho^b\} = -2\eta^{ab}\mathbf{1} \quad (5.7)$$

In  $D = 2n$  dimensions, the Dirac matrices are  $2^n \times 2^n$  matrices. Hence in two dimensions we get  $2 \times 2$  matrices, and it is easy to show that the following set satisfies the anti-commutation relation

$$\rho^0 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \rho^1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad (5.8)$$

Analogous to four dimensions,

$$\bar{\psi} = \psi^\dagger \rho^0 . \quad (5.9)$$

The representation for the Dirac matrices chosen here is the so-called *Majorana representation*. In this representation the Dirac operator  $i\rho^a \partial_a$  is real, and hence we may consistently impose the condition that  $\psi^A$  is real. This is called a *Majorana fermion*. Since it is real, we have in fact

$$\bar{\psi} = \psi^T \rho^0 . \quad (5.10)$$

The Dirac equation is

$$i\rho^a \partial_a \psi^A = 0 \quad (5.11)$$

Writing this in matrix notation we get

$$\begin{pmatrix} 0 & \partial_0 - \partial_1 \\ -\partial_0 - \partial_1 & 0 \end{pmatrix} \begin{pmatrix} \psi_-^A \\ \psi_+^A \end{pmatrix} = 0 \quad (5.12)$$

This leads to two decoupled equations

$$(\partial_0 - \partial_1)\psi_+^A = 0 , \quad (\partial_0 + \partial_1)\psi_-^A = 0 \quad (5.13)$$

which are satisfied if  $\psi_+^A$  depends only on  $\tau + \sigma$  and  $\psi_-^A$  depends only on  $\tau - \sigma$ . Hence just like the bosons, the free fermions split into left and right-moving components.

## 5.4 Quantization

The quantization of the fermionic action works similar to that of bosons, the main difference being that instead of commutators we get anti-commutators. The canonical momentum of a fermion is

$$\pi(\psi_k^A) = -\frac{i}{2\pi\alpha'}\psi_k^A \quad (5.14)$$

The canonical commutation relation yields

$$\{\psi_k^A(\sigma), \psi_l^B(\sigma')\} = 2\pi\alpha'\delta^{AB}\delta_{kl}\delta(\sigma - \sigma') . \quad (5.15)$$

---

There is a perhaps unexpected factor of two in the canonical momentum. This is related to the fact that  $\psi$  and  $\bar{\psi}$  are actually the same variable. The standard Lagrangian density for complex (Dirac) fermions is  $\mathcal{L} = -\bar{\psi}\not{\partial}\psi$ , and  $\psi$  and  $\bar{\psi}$  are treated as independent variables. This leads to a quantization condition  $\{\bar{\psi}, \psi\} = i\delta(\sigma - \sigma')$ . The standard Lagrangian for Majorana fermions differs from that of Dirac fermions by an additional factor 1/2, and  $\psi$  and  $\bar{\psi}$  are treated as related variables. Despite these differences, one gets same quantization condition (string actions differ from the “standard” field theory actions by a factor  $1/2\pi\alpha'$ , which appears in the quantization condition in the obvious way). A proper treatment of the quantization of free fermions requires the Dirac constraint formalism, which produces the correct factor of two; see *e.g.* [3] for details. Furthermore there are several sign choices to be made in the quantization procedure because fermionic variables must be treated as anti-commuting rather than commuting, even classically. These are just subtleties in the quantization of free fermions in any dimension, and that have nothing to do with string theory. Therefore this will not be discussed in more detail here.

---

Now we want to expand the fermionic fields in modes. To do this we need to know their boundary conditions. For closed strings the possible boundary conditions follow from the requirement that all physical quantities are periodic under  $\sigma \rightarrow \sigma + 2\pi$ . Physical quantities for fermions are always bilinears, and hence the fermions themselves can be either periodic or anti-periodic. This then leads to two distinct kinds of mode expansions

### Periodic fermions

$$\begin{aligned} \psi_-^A(\sigma, \tau) &= \sqrt{\alpha'} \sum_{n \in \mathbf{Z}} \bar{d}_n^A e^{-in(\tau - \sigma)} \\ \psi_+^A(\sigma, \tau) &= \sqrt{\alpha'} \sum_{n \in \mathbf{Z}} d_n^A e^{-in(\tau + \sigma)} \end{aligned} \quad (5.16)$$

### Anti-periodic fermions

$$\begin{aligned} \psi_-^A(\sigma, \tau) &= \sqrt{\alpha'} \sum_{r \in \mathbf{Z} + \frac{1}{2}} \bar{b}_r^A e^{-ir(\tau - \sigma)} \\ \psi_+^A(\sigma, \tau) &= \sqrt{\alpha'} \sum_{r \in \mathbf{Z} + \frac{1}{2}} b_r^A e^{-ir(\tau + \sigma)} \end{aligned} \quad (5.17)$$

For historical reasons such fermions are called Ramond fermions and Neveu-Schwarz fermions respectively. The notation  $b$  and  $d$  for their modes is also traditional. As usual, the “bar” distinguishes left- and right-moving modes, and is not meant to indicate complex conjugation.

It is now straightforward to substitute these mode expansions in the canonical commutation relation (5.15) and obtain the commutation relations for the modes:

$$\begin{aligned}\{b_r^A, b_s^B\} &= \delta^{AB} \delta_{r+s} \\ \{d_n^A, d_m^B\} &= \delta^{AB} \delta_{n+m}\end{aligned}\tag{5.18}$$

and similarly for the “barred” components. Furthermore barred and unbarred components commute (or anti-commute, if one prefers).

For open strings the discussion is slightly different. As we have seen for the bosonic string, the boundary condition arise from the surface term that one gets in a careful derivation of the equations of motion:

$$\begin{aligned}\delta_{\psi^A} S(\psi) &= \frac{i}{4\pi\alpha'} \int d\sigma d\tau [(\delta\bar{\psi}^A)\rho^a \partial_a \psi^A + (\bar{\psi}^A)\rho^a \partial_a \delta\psi^A] \\ &= \frac{i}{4\pi\alpha'} \int d\sigma d\tau [2(\delta\bar{\psi}^A)\rho^a \partial_a \psi^A + \partial_a \{(\bar{\psi}^A)\rho^a \delta\psi^A\}]\end{aligned}\tag{5.19}$$

If we require this variation to vanish for general variations  $\delta\psi^A$ , then the first term requires the equations of motion to be satisfied, *i.e.*  $\rho^a \partial_a \psi^A = 0$ , while the second term is a derivative, so it only contributes at the boundary. Since there are boundaries only in the  $\sigma$  direction, we get the requirement

$$\bar{\psi}^A \rho^1 \delta\psi^A = (\psi^A)^T \rho^0 \rho^1 \delta\psi^A = 0\tag{5.20}$$

at the boundary. This yields

$$\psi_+^A \delta\psi_+^A - \psi_-^A \delta\psi_-^A = 0\tag{5.21}$$

Note that we may consider variations for any value of  $A$  independently, so that this is a condition for every index  $A$  separately. It can be satisfied by relating  $\psi_+$  and  $\psi_-$  at the boundary (so that also their variations are related in the same way), *i.e.*  $\psi_+^A = \pm\psi_-^A$ ,  $\delta\psi_+^A = \pm\delta\psi_-^A$ . This would appear to allow in total four possibilities, namely the choices  $\pm$  at  $\sigma = 0$  and  $\sigma = \pi$ . But actually there are just two, since the overall relative sign between the  $+$  and  $-$  components of  $\psi$  is purely conventional. Hence we may fix  $\psi_+^A(\sigma=0) = \psi_-^A(\sigma=0)$  and then we have just two choices left. This yields exactly the same mode expansions is found above, except that one must replace  $\bar{b}$  by  $b$  and  $\bar{d}$  by  $d$ . Then “periodic” refers to the boundary condition  $\psi_+^A(\sigma=\pi) = \psi_-^A(\sigma=\pi)$  and anti-periodic to  $\psi_+^A(\sigma=\pi) = -\psi_-^A(\sigma=\pi)$ .

The quantization condition then comes out exactly the same as in the closed case.

## 5.5 The energy-momentum tensor

The energy-momentum tensor was previously defined as

$$T^{ab}(\sigma, \tau) \equiv \frac{4\pi\alpha'}{\sqrt{-\gamma}} \frac{\delta}{\delta\gamma_{ab}} S[\gamma] \quad (5.22)$$

where  $S$  is an action depending on the two-dimensional metric  $\gamma$  and other fields. This definition is only usable if we have the action in a form where  $\gamma$  is explicit. In the case of the bosonic string we started with the action in such a form, but then we went to conformal gauge, and the explicit dependence on  $\gamma$  disappeared. In the case of free fermions we started directly in conformal gauge. An action involving  $\gamma$  (*i.e.* an action that is reparametrization or general coordinate invariant) can be written down but for fermions this leads to some extra complications.

## 5.6 Noether currents

There is a second method for obtaining the energy momentum tensor, the Noether method. In general this works as follows. Suppose we have an action that is invariant under some (infinitesimal) symmetry:

$$S[\phi] = S[\phi + \epsilon\delta_S\phi] \quad (5.23)$$

where  $\phi$  stands for the set of fields,  $\epsilon$  is a constant parameter, and  $\delta_S\phi$  is a variation of  $\phi$  that corresponds to some symmetry. For example translation invariance corresponds to  $\delta_S\phi = n^a\partial_a\phi$ , where  $n^a$  is the direction of the translation. Now suppose we make  $\epsilon$  depend on the space-time point  $x$  (we work here in arbitrary dimensions, but we will apply the results to fields on a string world-sheet). Then the action need not be invariant, but since it must be invariant when  $\epsilon$  is constant, it must be true that

$$\delta S = \int d^D x J^\mu \partial_\mu \epsilon(x) , \quad (5.24)$$

for some  $J^\mu$  (there may be higher derivatives as well, but by partial integration one can always get this form). On the other hand, when  $\phi$  is a classical solution, *any* variation of the action around it should vanish, and this should in particular be true for the variation  $\epsilon(x)\delta\phi$ . Hence

$$\int d^D x J^\mu \partial_\mu \epsilon(x) = 0 \quad (5.25)$$

or, after integrating by parts

$$\int d^D x \epsilon(x) \partial_\mu J^\mu = 0 . \quad (5.26)$$

This must be true for arbitrary functions  $\epsilon(x)$ , and that implies that the *current*  $J^\mu$  is conserved,

$$\partial_\mu J^\mu = 0 \quad (5.27)$$

Note that this is true only under the assumption that the set of fields is a classical solution. Therefore we expect that to verify  $\partial_\mu J^\mu = 0$  one has to use the equations of motion. In this way one associates with every continuous symmetry a conserved current.

## 5.7 The Noether current of translation invariance

Let us now apply this principle to translation invariance. It turns out that the current associated to translation invariance in the direction  $n^\nu$  is  $J^\mu \propto n_\nu T^{\mu\nu}$ . We will not prove this in general here, but it can easily be verified for the bosonic string action.

Suppose we consider the translation  $\sigma^a \rightarrow \sigma^a + \epsilon n^a$ . Then  $X \rightarrow X + \epsilon n^a \partial_a X$  (for simplicity we leave out the target space index  $\mu$  on  $X$ ). The action then transforms as follows

$$S = -\frac{1}{4\pi\alpha'} \int d^2\sigma \eta^{ab} \partial_a X \partial_b X \rightarrow S - \frac{1}{2\pi\alpha'} \int d^2\sigma \eta^{ab} \partial_a [\epsilon n^c (\partial_c X)] \partial_b X + \mathcal{O}(\epsilon^2) \quad (5.28)$$

Hence

$$\begin{aligned} \delta S &= -\frac{1}{2\pi\alpha'} \int d^2\sigma \eta^{ab} \partial_a [\epsilon(\sigma) n^c (\partial_c X)] \partial_b X \\ &= -\frac{1}{2\pi\alpha'} \int d^2\sigma \eta^{ab} [(\partial_a \epsilon(\sigma)) n^c (\partial_c X) \partial_b X + \epsilon(\sigma) n^c (\partial_a \partial_c X) \partial_b X] \end{aligned} \quad (5.29)$$

The last term can be written as

$$\begin{aligned} \int d^2\sigma \eta^{ab} \epsilon(\sigma) n^c (\partial_a \partial_c X) \partial_b X &= - \int d^2\sigma \eta^{ab} \partial_c \epsilon(\sigma) n^c (\partial_a X) \partial_b X \\ &\quad - \int d^2\sigma \eta^{ab} \epsilon(\sigma) n^c (\partial_a X) \partial_c \partial_b X \\ &= - \int d^2\sigma \eta^{ab} \partial_c \epsilon(\sigma) n^c (\partial_a X) \partial_b X \\ &\quad - \int d^2\sigma \eta^{ab} \epsilon(\sigma) n^c (\partial_b X) \partial_c \partial_a X \end{aligned} \quad (5.30)$$

Hence

$$\int d^2\sigma \eta^{ab} \epsilon(\sigma) n^c (\partial_a \partial_c X) \partial_b X = -\frac{1}{2} \int d^2\sigma \eta^{ab} \partial_c \epsilon(\sigma) n^c (\partial_a X) \partial_b X \quad (5.31)$$

The variation of  $S$  is therefore

$$\begin{aligned} \delta S &= -\frac{1}{2\pi\alpha'} n^c \int d^2\sigma \left[ \eta^{ab} \partial_a \epsilon(\sigma) (\partial_c X) \partial_b X - \frac{1}{2} \eta^{ab} \partial_c \epsilon(\sigma) \partial_a X \partial_b X \right] \\ &= -\frac{1}{2\pi\alpha'} n_c \int d^2\sigma \left[ \partial_a \epsilon(\sigma) (\partial^c X) \partial^a X - \frac{1}{2} \eta^{ab} \eta^{ac} \partial_a \epsilon(\sigma) \partial_a X \partial_b X \right] \\ &= -\frac{1}{2\pi\alpha'} n_c \int d^2\sigma \partial_a \epsilon(\sigma) T^{ac}, \end{aligned} \quad (5.32)$$

where in the last step we substituted (2.23). The last line will now be taken as the definition of the energy-momentum tensor in the general case.

Applying this to the free fermion term in the action we get (dropping the index  $A$  for convenience)

$$\delta\psi = \epsilon n^c \partial_c \psi \quad (5.33)$$

so that

$$-2\pi\alpha'\delta S[\psi] = -\frac{1}{2}in^c \int d^2\sigma [\epsilon(\partial_c\bar{\psi})\rho^a\partial_a\psi + \bar{\psi}\rho^a\partial_a(\epsilon\partial_c\psi)] \quad (5.34)$$

If we use the equation of motion,  $\rho^a\partial_a\psi = 0$ , this reduces to

$$-\frac{1}{2}in^c \int d^2\sigma\partial_a\epsilon(\bar{\psi}\rho^a\partial_c\psi) \quad (5.35)$$

from which we read off

$$T_{ac}(\psi) = -\frac{1}{2}i\bar{\psi}\rho_a\partial_c\psi \quad (5.36)$$

Note that this tensor is not symmetric, and hence not equal to the energy momentum tensor one would get from a variation with respect to  $\gamma_{ac}$ . This is due to the fact that in general a current  $J_\mu$  does not follow uniquely from the Noether procedure, but only up to terms of the form  $\partial_\mu A_{\mu\nu}$ , where  $A_{\mu\nu}$  is an anti-symmetric tensor. One can indeed add such a term to  $T^{ac}$  to symmetrize it (this is called an *improvement term*; see exercise 5.6 in [6] for more details). In two dimensions the tensor turns out to be both symmetric and traceless provided one uses the equations of motion. To see the former, note that (with  $\rho^0\rho_0 = -1$ )

$$\begin{aligned} T_{01} &= -\frac{1}{2}i\bar{\psi}\rho_0\partial_1\psi \\ &= \frac{1}{2}i(\psi_-\partial_1\psi_- + \psi_+\partial_1\psi_+) \end{aligned} \quad (5.37)$$

while

$$\begin{aligned} T_{10} &= -\frac{1}{2}i\bar{\psi}\rho_1\partial_0\psi \\ &= -\frac{1}{2}i(\psi_-\partial_0\psi_- - \psi_+\partial_0\psi_+) \end{aligned} \quad (5.38)$$

Now use  $\partial_0\psi_- = -\partial_1\psi_-$  and  $\partial_0\psi_+ = \partial_1\psi_+$  to show that  $T_{01} = T_{10}$ . For the  $++$  and  $--$  components of the tensor we find

$$T_{++} = \frac{1}{2}(T_{00} + T_{01}) = -\frac{i}{4}\bar{\psi}\rho_0(\partial_0 + \partial_1)\psi = \frac{i}{4}\psi^T(\partial_0 + \partial_1)\psi \quad (5.39)$$

and analogously for  $T_{--}$ . Obviously  $\psi_+$  appears only in  $T_{++}$  and  $\psi_-$  only in  $T_{--}$ .

## 5.8 The mass spectrum

It is now easy to compute the mass spectrum. As we did before we expand Virasoro generators in modes, defined exactly as before.\* Now we get

$$L_m = \frac{1}{2} \sum_n \alpha_{-n}^i \alpha_{m+n}^i + \frac{1}{2} \sum_r (r + \frac{1}{2}m) b_{-r}^A b_{m+r}^A \quad (5.40)$$

---

\* To get the result directly in this form one should use the symmetrized form of the energy-momentum tensor. Note that the terms proportional to  $m$  are actually zero due to the anti-commutators.

in the Neveu-Schwarz case, and

$$L_m = \frac{1}{2} \sum_n \alpha_{-n}^i \alpha_{m+n}^i + \frac{1}{2} \sum_n \left(n + \frac{1}{2}m\right) d_{-n}^A d_{m+n}^A \quad (5.41)$$

in the Ramond case. The space-time part has been written in light-cone gauge. As before, there are ordering ambiguities. They only affect  $L_0$ . In the free boson case we had argued that for each boson separately they resulted in  $(L_0)_{\text{cl}}(X) = (L_0)_{\text{qu}}(X) - \frac{1}{24}$ , where  $(L_0)_{\text{cl}}$  is the expression derived directly from the action, and  $(L_0)_{\text{qu}}$  is the normal ordered expression. For 24 transverse boson  $X^i$  this produced exactly the required shift  $-a = -1$ .

## 5.9 The Neveu-Schwarz ground state

We can do the same for fermions. For a single Neveu-Schwarz fermion the “naive” classical contribution to  $L_0$  is

$$(L_0)_{\text{cl}}(\psi) = \frac{1}{2} \sum_{r=-\infty}^{\infty} r b_{-r} b_r \quad (5.42)$$

Just as in the bosonic case, we define the vacuum by

$$b_r |0\rangle = 0 \text{ for } r > 0 \quad (5.43)$$

so that the negative modes create states. Normal ordering is defined as putting all creation operators on the left of all the annihilation operators. We see then that in the infinite sum all the terms with  $r < 0$  are in the wrong order, and we use the anti-commutator to re-order them. This yields

$$r b_{-r} b_r = -r b_r b_{-r} + r = s b_{-s} b_s - s, \quad (5.44)$$

with  $s = -r$ . Hence for each positive, half-integer value  $s$  we get a contribution  $-s$ . They can be added as follows

$$\begin{aligned} \frac{1}{2} \sum_{s \in \mathbf{Z} + \frac{1}{2}; s > 0} (-s) &= -\frac{1}{4} \sum_{n=0}^{\infty} (2n+1) \\ &= -\frac{1}{4} \left[ \sum_{m=1}^{\infty} m - \sum_{n=1}^{\infty} 2n \right] \\ &= -\frac{1}{4} \left[ \sum_{m=1}^{\infty} m - 2 \sum_{n=1}^{\infty} n \right] \\ &= \frac{1}{4} \sum_{m=1}^{\infty} m = -\frac{1}{48} \end{aligned} \quad (5.45)$$

Hence we find that the contribution to  $a$  of a Neveu-Schwarz fermion is precisely half that of a boson. As was the case for bosons, there are better ways to arrive at this result, but this would require more discussion.



## 5.10 The Ramond ground state

For Ramond fermions the discussion is slightly more complicated. This is due to the presence of a zero-mode fermion  $d_0$ . Should we impose  $d_0^A |0\rangle = 0$ ? If one attempts to do that one would find a contradiction with the anti-commutator

$$\{d_0^A, d_0^B\} = \delta^{AB} \quad (5.46)$$

In fact there is another problem. One may easily show that

$$[L_0, d_m] = -m d_m, \quad (5.47)$$

so that acting with  $d_m$  increase the  $L_0$  eigenvalue of a state by  $-m$  (this holds also for the other operators  $\alpha_m$  and  $b_r$ ). Hence acting with  $d_0$  does not change the  $L_0$  eigenvalue, or in other words, the energy of a state. Hence the state  $d_0^A |0\rangle$  is degenerate with the vacuum.

This leads one to consider as the ground state of the Ramond fermions a set of states  $|a\rangle_R$  with the property

$$d_0^A |a\rangle_R = \sum_b M_{ab}^A |b\rangle_R \quad (5.48)$$

Here  $M_{ab}^A$  is a set of matrices (as many as there are fermions) which must satisfy

$$\{M^A, M^B\} = \delta^{AB} \quad (5.49)$$

Such a relation is called a *Clifford algebra*. In general one prefers to write

$$M^A = \frac{1}{\sqrt{2}} \gamma^A, \quad (5.50)$$

and then a solution is obtained by taking for  $\gamma^A$  the (Euclidean) Dirac matrices, which satisfy

$$\{\gamma^A, \gamma^B\} = 2\delta^{AB} \quad (5.51)$$

For example, for a single fermion the solution is  $\gamma^1 = 1$ , for two fermions one has

$$\gamma^1 = \sigma^1, \quad \gamma^2 = \sigma^2, \quad (5.52)$$

for three fermions one gets  $\gamma^i = \sigma^i$ , for four fermions one gets the four-dimensional Euclidean Dirac matrices (which are  $4 \times 4$  matrices) etc. In general, in  $2n + 1$  and  $2n$  dimension one gets  $2^n \times 2^n$  matrices, and the ground states  $|a\rangle_R$  has  $2^n$  components.

The energy of the ground state (which yields the mass in target space) can be evaluated as we did for the Neveu-Schwarz fermions, and one easily finds that the shift in  $L_0$  is in fact  $-\frac{1}{2} \sum_{n=1}^{\infty} n = +\frac{1}{24}$  per fermion, the opposite of free boson.

## 5.11 Excited states

In principle each fermion  $\psi^A$  could be periodic or anti-periodic, with separate choices for each label  $A$ . We will see later that if we introduce interactions, there will be constraints on this. Here we will just consider the two extremes: that all are Neveu-Schwarz, or that all are Ramond.

### 5.11.1 Neveu-Schwarz

The higher excited states in the spectrum are as follows. In a model with NS (Neveu-Schwarz) fermions the target space state corresponding to the vacuum  $|0\rangle$  has mass  $-a_{\text{NS}}$ , where  $a_{\text{NS}}$  is the shift in the zero-point of  $L_0$  for the Neveu-Schwarz case. The lowest states are summarized in the following table. Here we use units  $4/\alpha'$  for closed strings and  $1/\alpha'$  for open strings, and we consider only the left-moving sector of the closed string. As before, left and right sectors must be combined, imposing the condition  $M_L^2 = M_R^2$ .

State	$M^2$
$ 0\rangle$	$-a_{\text{NS}}$
$b_{-\frac{1}{2}}^A  0\rangle$	$\frac{1}{2} - a_{\text{NS}}$
$\alpha_{-1}^i  0\rangle, b_{-\frac{1}{2}}^A b_{-\frac{1}{2}}^B  0\rangle$	$1 - a_{\text{NS}}$

The first excited state is  $b_{-\frac{1}{2}}^A |0\rangle$ , and it has mass  $\frac{1}{2} - a_{\text{NS}}$  (here “mass” is an abbreviation for “mass-squared”). Next we get a state  $\alpha_{-1}^i |0\rangle$  plus the states  $|A, B\rangle = b_{-\frac{1}{2}}^A b_{-\frac{1}{2}}^B |0\rangle$ , all with mass  $1 - a_{\text{NS}}$ . Note that  $|A, B\rangle$  must be antisymmetric in  $A$  and  $B$ . The value of  $a$  is  $\frac{D-2}{24} + \frac{N}{48}$ , where  $D$  is the number of space-time dimensions and  $N$  the number of fermions ( $A = 1, \dots, N$ ).

### 5.11.2 Ramond

In a Ramond model (*i.e.* all  $N$  fermions have periodic boundary conditions) the ground state is  $|a\rangle_R$  with mass  $-a_R$ , with  $a_R = \frac{D-2}{24} - \frac{N}{24}$ . The first excited states are the states  $\alpha_{-1}^i |a\rangle_R$  and  $d_{-1}^A |a\rangle_R$  with mass  $1 - a_R$ . Here we have normal ordered the operators  $d$  and  $\alpha$  with respect to the states  $|a\rangle_R$ , using

$$\alpha_m |a\rangle_R = d_m |a\rangle_R = 0 \text{ for } m > 0. \quad (5.53)$$

### 5.11.3 Lorentz multiplets

The naive Lorentz symmetry argument is not of much help if  $N > 1$ . The argument was that the state  $\alpha_{-1}^i |0\rangle$  had only  $D - 2$  components, and hence has to be a massless vector. This then determines  $a$ , and then also  $D$ . In the present case in the NS model there are additional states at the second excited level; this is also true in the R model, and furthermore the ground states  $|a\rangle_R$  are themselves degenerate. As before, the correct approach is to consider the Lorentz algebra. This yields  $D + \frac{1}{2}N = 26$  in both cases. This is a consequence of the fact that each fermion contributes  $\frac{1}{2}$  to the conformal anomaly, as can easily be verified using the explicit form of the Virasoro generators, Eqns. (5.40) and (5.41). In general, if we use instead of compactified free bosons or free fermions some other two-dimensional theory with conformal symmetry and a conformal anomaly  $c$ , the dimension is determined as  $D + c = 26$ . If we substitute the dimension in the formulas for  $a_{\text{NS}}$  and  $a_R$  we obtained earlier from the naive zero-mode summation we get  $a_{\text{NS}} = 1$  and

$a_R = 1 - \frac{N}{16}$ . Note that if we could build a Ramond model with sixteen or more fermions, the ground state would have non-negative mass-squared, *i.e.* would not be tachyonic. However, we will see that this is not possible: Purely Ramond or purely Neveu-Schwarz string theories are not consistent. Inconsistencies arise as soon as one studies interactions.

## 5.12 Bosonization\*

It turns out that what we have constructed so far is nothing but a compactified bosonic string. If one compactifies a bosonic string on a suitably chosen lattice, the resulting spectrum is exactly the one we got above, with a precise relation between R and NS sectors. The number of bosons must be twice the number of fermions. The proof of this statement is related to *bosonization*: in two dimensions two fermions  $\psi^1$  and  $\psi^2$  are related to a boson  $\phi$ . The relation is  $\psi^1 = e^{i\phi}$  and  $\psi^2 = e^{-i\phi}$ . This is however beyond the scope of these lectures.

It should be noted that even though we started with fermions in two dimensions, the space-time spectrum contains only Lorentz tensors, and no spinors. Hence we only get *bosons* in the target space spectrum.

## 5.13 Fermionic strings

One can obtain a much more interesting result by making a small modification in the foregoing discussion. Instead of giving the fermions an “internal” index ‘ $A$ ’ we give them a Lorentz index  $\mu$ . If we only have the bosons  $X^\mu$  and the fermions  $\psi^\mu$  it is easy to derive some interesting consequences. The action is of course

$$S = -\frac{1}{4\pi\alpha'} \int d\sigma d\tau [\partial_a X^\mu \partial^a X^\nu - i\bar{\psi}^\mu \rho^a \partial_a \psi^\nu] \eta_{\mu\nu} \quad (5.54)$$

The canonical commutation relations now become

$$\{d_n^\mu, d_m^\nu\} = \eta^{\mu\nu} \delta_{n+m}, \quad \{b_r^\mu, b_s^\nu\} = \eta^{\mu\nu} \delta_{r+s} \quad (5.55)$$

Now we see immediately that there is a problem: the components  $d_n^0$  and  $b_r^0$  have anti-commutators of the wrong sign, and will lead to ghost states in the spectrum. In the case of the bosonic string this was solved by imposing the  $T^{ab} = 0$  constraints. Those constraints were explicitly solve by going to light-cone gauge, using conformal invariance. This won't help us with  $\psi^\mu$ . Since we have already used up conformal invariance to reduce  $X^\mu$  to  $X^i$ , we cannot achieve the same reduction for  $\psi^\mu$ . The way out turns out to be an additional symmetry called super-conformal invariance. We will come back to it later, but let us assume that it does what we expect, and fixes two components of  $\psi^\mu$ , reducing it to  $\psi^i$ .

## 5.14 The critical dimension

The following holds for open strings, or the left- or right-moving modes of the closed string (up to a factor  $4/\alpha'$  for closed strings or  $1/\alpha'$  for open strings). Consider the first excited

state in the NS sector,  $b_{-\frac{1}{2}}^\mu |0\rangle$ , which is a vector boson. In light cone gauge it would become  $b_{-\frac{1}{2}}^i |0\rangle$ , which has only  $D - 2$  components, and is the only state at its mass level. By the arguments used before, it must be massless. This requires  $a_{\text{NS}} = \frac{1}{2}$ . On the other hand,  $a_{\text{NS}} = \frac{D-2}{24} + \frac{N_{\text{tr}}}{48}$ . Here  $N_{\text{tr}}$  is the number of “transverse” fermions, *i.e.* the fermions that actually appear in the mass formula, *i.e.*  $\psi^i$ ,  $i = 1, \dots, D - 2$ . Hence  $N_{\text{tr}} = D - 2$ , and from  $a_{\text{NS}} = \frac{1}{2}$  one easily deduces that  $D = 10$ . If we consider the state  $\alpha_{-1}^i |0\rangle$  that gave rise to a massless vector boson in the bosonic string, then we find that now there is no problem making it massive. It is accompanied by the set of states  $b_{-\frac{1}{2}}^i b_{-\frac{1}{2}}^j |0\rangle$ , which is an anti-symmetric tensor. The total number of components at this level is

$$\frac{1}{2}(D-2)(D-3) + (D-2) = \frac{1}{2}(D-2)(D-1), \quad (5.56)$$

precisely the number of components of an asymmetric tensor of  $SO(D-1)$ . Hence at this level we will have a massive anti-symmetric tensor.

We conclude therefore that the fermionic string considered here has a “critical” dimension  $D = 10$ . The number of world-sheet fermions and bosons is equal, namely  $N = D = 10$ . Note that this does not satisfy the relation  $D + \frac{1}{2}N = 26$  mentioned in the previous section.

## 5.15 The Lorentz algebra\*

This can be understood as follows. We have argued that the relation  $D + \frac{1}{2}N = 26$  follows from closure of the Lorentz-algebra in light-cone gauge. Indeed, in the bosonic string that will always lead to relation (5.5):  $D + c_{\text{int}} = 26$ . However, the set of Lorentz generators of the bosonic string, as shown in (3.93), acts only on the bosonic degrees of freedom  $X^\mu$ . This is also true if we introduce compactified coordinates, or internal fermions  $\psi^A$ . Since the Lorentz-rotations do not act on these internal degrees of freedom, the form of (3.93) stays exactly the same. If one goes to light-cone gauge the generators  $M^{i-}$  do depend on the internal degrees of freedom because we eliminate  $\alpha_n^-$  by solving the condition  $T^{ab} = 0$ , and  $T^{ab}$  contains all degrees of freedom. However, the generators  $M^{\mu\nu}$  still act only on  $X^\mu$ .

Clearly we get new Lorentz generators if we give the fermions a space-time index  $\mu$ . These generators act on  $\psi^\mu$  as well as  $X^\mu$ . The derivation is easy: the generators are just the Noether currents associated with the global transformation  $X^\mu \rightarrow \Lambda^\mu_\nu X^\nu$ ,  $\psi^\mu \rightarrow \Lambda^\mu_\nu \psi^\nu$ , where  $\Lambda^\mu_\nu$  is a Lorentz transformation. The result – written in terms of modes – consist of the terms we already had, (3.93), plus the following terms:

$$\begin{aligned} \text{NS:} \quad & -i \sum_{r=\frac{1}{2}}^{\infty} (b_{-r}^\mu b_r^\nu - b_{-r}^\nu b_r^\mu) \\ \text{R:} \quad & -\frac{i}{2} [d_0^\mu, d_0^\nu] - i \sum_{n=1}^{\infty} (d_{-n}^\mu d_n^\nu - d_{-n}^\nu d_n^\mu) \end{aligned} \quad (5.57)$$

This of course changes the discussion of the Lorentz drastically, and it is no surprise that the result is different. Indeed, in the NS-sector one finds  $a_{\text{NS}} = \frac{1}{2}$  and  $D = 10$ , and in the R-sector  $a_{\text{R}} = 0$ ,  $D = 10$ .

## 5.16 The Ramond ground state

We now arrive at a very interesting conclusion regarding the Ramond ground state. Consider the first term in the expression for the Lorentz-generator. We have seen before that  $d_0^A$  acts on the Ramond ground states as a Dirac  $\gamma$  matrix. This does not change if we replace the index  $A$  by a space-time index  $\mu$ :

$$d_0^\mu |a\rangle_R = \frac{1}{\sqrt{2}} \gamma_{ab}^\mu |b\rangle_R \quad (5.58)$$

Consider now the action of the Lorentz transformation (5.57) on these states. Infinitesimal Lorentz transformations are generated by acting with the operator  $\omega_{\mu\nu} M^{\mu\nu}$  on a state, where  $\omega$  is a set of small parameters. We see that under Lorentz transformations  $|a\rangle_R$  transforms as

$$|a\rangle_R \rightarrow |a\rangle_R - \frac{i}{4} [\gamma^\mu, \gamma^\nu]_{ab} \omega_{\mu\nu} |b\rangle_R . \quad (5.59)$$

The matrix  $-\frac{i}{4} [\gamma^\mu, \gamma^\nu]$  (often called  $\Sigma^{\mu\nu}$ ) is the infinitesimal Lorentz generator in the spinor representation. Apparently the Ramond ground states rotate like spinors. The corresponding particles must thus be space-time spinors and – assuming the spin-statistics relation holds – they must be fermions. Then all their excited states are fermions as well, since they are obtained from the ground states by action with space-time bosonic operators  $d^\mu$  and  $\alpha^\mu$ .

## 5.17 World sheet supersymmetry\*

An important missing point in the foregoing is the condition that allows us to fix the light-cone gauge for  $\psi^\mu$ . This condition originates from an extra symmetry of the action. One may verify that it is invariant under the transformation

$$\begin{aligned} \delta X^\mu &= \bar{\epsilon} \psi^\mu \\ \delta \psi^\mu &= -i \rho^a \partial_a X^\mu \epsilon , \end{aligned} \quad (5.60)$$

where  $\epsilon$  is a spinor that anti-commutes with itself and with  $\psi^\mu$ .

A symmetry like this one, which mixes bosons and fermions, is called a *supersymmetry*. It is not a symmetry of nature: for example, there are no bosonic partners of the quarks and leptons. There is a lot of speculation that supersymmetry may be an approximate symmetry of nature, and that the “superpartners” of the standard model particles will soon be discovered, but that is irrelevant for our purposes. The supersymmetry we consider here is a *world-sheet* symmetry, and its presence does not imply that the *target space* spectrum we will get is supersymmetric.

By considering  $\sigma$ -dependent parameters  $\epsilon$  we may work out the Noether current of space-time supersymmetry. The result is

$$\delta S = \frac{1}{\pi\alpha'} \int d^2\sigma \partial_a \bar{\epsilon} J^a \quad (5.61)$$

with

$$J^a = \frac{1}{2} \rho^b \rho^a \psi^\mu \partial_b X_\mu \quad (5.62)$$

It is easy to verify that this current is indeed conserved. To do this both the free fermion and the free boson equations of motion are needed.

The current  $J^a$  is called the supercurrent.

## 5.18 World sheet supergravity\*

Up to now the presence of world-sheet supersymmetry was a mere coincidence, but in fact it turns out to be crucial. However, it is not completely trivial how to maintain it when we try to write the action in a manifestly reparametrization invariant way. Then we would introduce a metric  $\gamma_{\alpha\beta}$ , which means introducing a new two-dimensional field in addition to  $X^\mu$  and  $\psi^\mu$ . Supersymmetry acts on this field. It cannot act trivially, because it transforms bosons into fermions. Hence we are obliged to introduce an additional field, the superpartner of the world-sheet metric  $\gamma_{\alpha\beta}$ .

Instead of  $\gamma_{\alpha\beta}$  a more convenient set of fields is the so-called *zwei-bein*  $e_\alpha^a$ ,

$$\gamma_{\alpha\beta} = \eta_{ab} e_\alpha^a e_\beta^b \quad (5.63)$$

The supersymmetry transformation of the zweibein is

$$\delta e_\alpha^a = -2i\bar{\epsilon}\rho^a \chi_\alpha, \quad (5.64)$$

where  $\chi_\alpha$  is a new field. It is a spinor with an extra two-dimensional vector index. If we construct such a field in four dimensions, we get a spin-3/2 fields (spin-1/2 times spin-1 yields spin-3/2), but in two dimensions the notion of spin is rather trivial (the little groups are  $SO(1)$  for massive and  $SO(0)$  for massless particles). In any case, this field is called a *gravitino*. It turns out that if we want to write down an action that is reparametrization invariant and has supersymmetry, we are forced to write down a *supergravity* theory. The complete action is

$$S = -\frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{-\gamma} \times \left\{ \gamma^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X_\mu - i\bar{\psi}^\mu \rho^\alpha \partial_\alpha \psi_\mu + 2\bar{\chi}_\alpha \rho^\beta \rho^\alpha \psi^\mu \partial_\beta X_\mu + \frac{1}{2} \bar{\psi}_\mu \psi^\mu \bar{\chi}^\alpha \rho^\beta \rho^\alpha \chi_\beta \right\}. \quad (5.65)$$

For more details about this action and how it is obtained see [1]. The first two terms are just the generally covariant form of the conformal gauge action. The third term couples

the gravitino to the supercurrent. The last term is a four-fermi interaction required by supersymmetry. Note that there are no kinetic terms for the gravitino. In more than two dimensions there would be such a term, but it vanishes in two dimensions. This action is not just supersymmetric, it is in fact invariant under *local* supersymmetry, which means that the transformations can be  $\sigma$ -dependent.

### 5.19 The constraints\*

The correct quantization procedure for the fermionic string is to start with the supergravity action, derive the equations of motion, and only then go to conformal gauge, if one wishes. This yields one additional equation of motion, namely the one for the gravitino. In conformal gauge this equation is simply

$$J_a = 0 . \tag{5.66}$$

(“conformal gauge” means a little more than for the bosonic string. It turns out that one can choose  $e_\alpha^a = \delta_\alpha^a$ , and one can use the extra fermionic symmetries to set  $\chi_\alpha = 0$ .) The other three equations of motion are the usual ones for  $X^\mu$  and  $\psi^\mu$ , plus the condition  $T_{ab} = 0$ .

An obvious guess is now that the new condition is the constraint we were looking for earlier. This is indeed correct. In light-cone coordinates  $\psi^\mu$  splits into  $\psi^i, \psi^+$  and  $\psi^-$ . Using conformal symmetry and a supersymmetry transformation that survives the conformal gauge choice, we may bring  $X^+$  to the same form as before, and we can set  $\psi^+$  to zero. Then the conditions  $J_a = T_{ab} = 0$  can be solved, yielding an expression  $X^-$  and  $\psi^-$  in terms of the other components:

$$\begin{aligned} \partial_+ X^- &= \frac{1}{p^+} T_{++}^\perp \\ \psi^-(\sigma^+) &= \frac{2}{p^+} J_+^\perp \end{aligned} \tag{5.67}$$

where  $T^\perp$  and  $J^\perp$  are the the energy momentum tensor and the supercurrent, but with the summation indices  $\mu = 0, \dots, D-1$  replaced by  $i = 1, \dots, D-2$ . In the case of closed strings there are analogous relations for the right-moving components  $\partial_- X^-$  and  $\psi^-(\sigma^-)$ , whereas in the open string left- and right-moving components of all fields are the same.

These relations express  $\alpha_n^-$  and  $b_r^-$  (or  $d_n^-$ ) in terms of all the other oscillators. To get the full spectrum it therefore sufficient to act with all transverse oscillators, as we already did above, anticipating this result.

### 5.20 The conformal anomaly\*

In the case of the bosonic string the critical dimension could be obtained by cancelling the conformal anomaly of the fields  $X^\mu$  against the contribution of the ghost of superconformal

invariance. This yields  $D - 26 = 0$ . How does this change if we consider fermionic strings? It turns out that the extra local supersymmetry we got requires additional Fadeev-Popov ghosts. These ghosts also contribute to the conformal anomaly, and their contribution is  $+11$ . For  $D$  bosons and  $D$  fermions we get then the equation  $D + \frac{1}{2}D - 26 + 11 = 0$ , which gives  $D = 10$ .



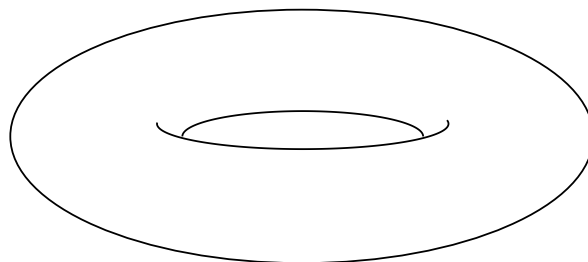
## 6 One loop diagrams

### 6.1 Neveu-Schwarz or Ramond?

The discussion so far give the impression that we can make a choice: that one can have Neveu-Schwarz or Ramond strings. If this were true the result would be undesirable, because we would get either a theory with only bosons, or a theory with only fermions. However, just as one cannot make a theory with only open strings, one cannot make a theory with only NS or R strings. As soon as one takes interactions into account, both are needed. This is seem most elegantly if one considers the string one-loop diagram. This is what we will compute now. To make life easy we consider a one loop diagram without external lines. Furthermore we consider a closed string diagram, because that will in any case occur in any string theory, open or closed.

### 6.2 The torus

A closed string loop diagram is of course a torus, as shown here



What we mean by computing the diagram is to compute a number, an amplitude, that corresponds to the diagram. Since there are no external lines, this amplitude is not the amplitude for some scattering process, but it turns out that it can be interpreted as a contribution to the energy of the vacuum in target space. Experimentally, this number would be measurable because it contributes to the cosmological constant, which affects the evolution of the universe. Although this is a very interesting subject, it is not our main interest now. We will not be interested in the result of the calculation, but in its internal consistency. This will impose strong constraints on the theory, which determine precisely the relation between R and NS.

### 6.3 The Polyakov path integral\*

The procedure one uses for doing these calculations is the path-integral. Conceptually this is rather easy to understand, but the details are messy, and require advanced mathematical

methods. Below the main steps are sketched. Details can be found, for example, in [1],[3] and [2]. Schematically, the calculation goes as follows.

$$A = \int \mathcal{D}[\text{fields}] \mathcal{D}[\gamma] \mathcal{D}[\chi] e^{-S_E[\text{fields}, \gamma, \chi]} \quad (6.1)$$

Here  $S_E$  is the Euclidean action, obtained from the action in two-dimensional Minkowski space by defining  $\sigma_0 = -i\sigma_2$  (see appendix A). Furthermore “fields” stands for all fields in the theory, such as  $X^\mu$  and  $\psi^\mu$  and any internal degrees of freedom. One integrates over all values of these fields in all points on the torus, and also over all values of the two-dimensional metric  $\gamma$  and the gravitino. Hence all  $X^\mu(\sigma_a)$  are treated as integration variables, labelled by  $\mu, \sigma_0$  and  $\sigma_1$ . Obviously this is an infinite-dimensional integral.\*

The analytic continuation to imaginary time has the advantage of making the path integral better behaved, because in Minkowski space the weight factor is  $\exp(iS/\hbar)$ . In the classical limit,  $\hbar \rightarrow 0$ , this integral is dominated by the extrema of the action, the classical solutions. In the Euclidean formulation it is clear that fluctuations around the classical solutions, which increase the action, are suppressed by the negative exponential.

However, there may be fluctuations that do not change the action. These are usually related to symmetries. Let us first consider the integration over the two-dimensional metric  $\gamma$ . Then the symmetries we have to worry about are reparametrization and Weyl invariance. If we take the definition we gave above, integrating these symmetries would simply give rise to an infinite factor. Therefore we change the definition:

$$A_{\text{fix}} = \frac{A}{\text{Vol}(\text{Diff})\text{Vol}(\text{Weyl})\text{Vol}(\text{Super})} \quad (6.2)$$

where  $\text{Vol}(\text{Diff})$  denotes the total integration over the reparametrizations (diffeomorphisms),

$\text{Vol}(\text{Diff})$  the integration over the Weyl symmetry and  $\text{Vol}(\text{Super})$  the integration over the local supersymmetries. Note that we are not yet considering the space-time Poincaré symmetries of  $X^\mu$ , because we are discussing the integral over the two-dimensional metric, which is not affected by these symmetries.

Of course  $A_{\text{fix}}$  is a ratio of two infinite quantities, which requires some care. This is done by means of the Fadeev-Popov procedure, which will not be described in detail here. Essentially one writes the integral over  $\gamma$  and  $\chi$  as an integral over fluctuations of these fields around the conformal gauge, and the latter integrals then cancel against the “Vol” factors in the denominator. In doing these manipulations one has to make a change of integration variables for  $\gamma$  and  $\chi$ , and this results in some Jacobian factors for the change of measure. The result is thus

$$A_{\text{fix}} = \int d\Omega \int \mathcal{D}[\text{fields}] \text{Jac}(\gamma) \text{Jac}(\chi) e^{-S_E[\text{fields}, \eta, 0]}, \quad (6.3)$$

---

\* An additional complication is that fact that some fields are fermionic. The the corresponding path integral variables must be anti-commuting, which requires a special formalism which will not be discussed here.

where  $\text{Jac}(\gamma)$  and  $\text{Jac}(\chi)$  are the Jacobians. The integral over  $\Omega$  is over all two-dimensional surfaces that are really inequivalent: surfaces that cannot be transformed into each other by coordinate and Weyl-transformations. We will return to this later.

To take care of the Jacobians one introduces extra fields, the Fadeev-Popov ghost  $b, c, \beta$  and  $\rho$ , such that\*

$$\text{Jac}(\gamma)\text{Jac}(\chi) = \int \mathcal{D}b\mathcal{D}c\mathcal{D}\beta\mathcal{D}\rho e^{-S_E(b,c,\beta,\rho)}, \quad (6.4)$$

for some suitable ghost action. It turns out that this action is Gaussian, so that the integral can be done explicitly. A Gaussian path integral is just a generalization of the well-known result for a finite-dimensional integral

$$\int dx_1 \dots dx_n e^{-x_i A_{ij} x_j} = \pi^{n/2} [\det A]^{-\frac{1}{2}}. \quad (6.5)$$

If we consider uncompactified strings, the action of  $X$  and  $\psi$  is Gaussian as well, and can also be done.

In the special case of the torus it turns out that the ghost contribution precisely cancels against two degrees of freedom of  $X^\mu$  and  $\psi^\mu$ , so that we are finally left with

$$A_{\text{fix}} = \int d\Omega \int \mathcal{D}[X]\mathcal{D}[\psi] e^{-S_E[X^i, \psi^i, \eta, 0]}, \quad (6.6)$$

with the action in conformal gauge. This cancellation only works in 26 dimensions for the bosonic string and 10 dimensions for the superstring.

## 6.4 Riemann surfaces\*

The foregoing discussion (with the exception of the last paragraph) was not limited to the torus, but holds for any surface. Before continuing with the torus, let us make a few general remarks about string perturbation theory.

The full perturbative expression for any amplitude has the form

$$\langle \text{out} | U | \text{in} \rangle = \sum_{\text{topologies}} \int d\Omega \int \mathcal{D}[\text{fields}] \int \mathcal{D}b\mathcal{D}c\mathcal{D}\beta\mathcal{D}\rho e^{-S_E(\text{fields}, b, c, \beta, \rho)} \quad (6.7)$$

Here the sum is over all surfaces that cannot be deformed continuously into each other, and the integral is over all parameters of each surface that do not correspond to coordinate changes and Weyl transformations.

Since string world sheets are two-dimensional we are in the lucky situation that we can make use of powerful mathematical results regarding such surfaces, which are called *Riemann surfaces*. The parameters of such a surface are called the *moduli*, and for each topology the moduli are known.

---

\* In the literature  $\rho$  is usually called  $\gamma$ , but unfortunately that character is already in use in this context.

The topology of Riemann surfaces is encoded in three integers: the number of handles, also known as the *genus*  $g$ , the number of holes  $b$  and the number of “crosscaps”  $c$  (a crosscap is a hole of which opposite sides are glued together with opposite orientation.). For example, a sphere has  $g = b = c = 0$ ; a torus is a sphere with one handle added to it, so  $g = 1, b = c = 0$ ; a disk is a sphere with one hole cut out, so  $g = c = 0$  and  $b = 1$ .

Which surfaces are allowed depends on the kind of string theory one considers. For example, the presence of a hole means that strings can have endpoints. Hence surfaces with holes appear only for open string theories. A crosscap reverses orientation, and appears only for unoriented strings. In general one has

Closed, Oriented	$b = c = 0$
Closed, Unoriented	$b = 0$
Open, Oriented	$c = 0$
Open, Unoriented	no restrictions

where “open” means, as usual, “open and closed”.

The amplitudes of the closed oriented string are sums over the genus  $g$ , and hence the contributing surfaces are the sphere, the torus, the double torus, etc. This is reminiscent of a loop expansion in field theory, and hence one might expect that there is an expansion parameter, a coupling constant, in terms of which higher genus surfaces are suppressed.

## 6.5 The rôle of the dilaton\*

In the discussion of bosonic string compactification we have seen that we can describe string propagation in non-trivial gravitational backgrounds  $g_{\mu\nu}$  and anti-symmetric tensor backgrounds  $B_{\mu\nu}$ . There is one other massless field in the uncompactified closed string, namely the dilaton. It seems natural to look for a modification of the action that involves the dilaton  $\phi$ . There are in fact several natural candidates, for example

$$S(X, \gamma, \phi) \propto \int d\sigma d\tau \sqrt{\gamma} \phi \tag{6.8}$$

or

$$S(X, \gamma, \phi) \propto \int d\sigma d\tau \sqrt{\gamma} X^\mu X_\mu \phi \tag{6.9}$$

to be added to the action we already have. The problem with both proposal is that they both break Weyl invariance. Weyl-invariance is a transformation of the two-dimensional fields  $X$  and  $\gamma$ . From the two-dimensional point of view the space-time background fields  $g, B$  and  $\phi$  are not fields, but more like coupling constants. They are not supposed to transform under reparametrizations and Weyl transformations.

If Weyl invariance is broken, we loose the possibility to go to the conformal gauge, and later to light-cone gauge. This turns out to have fatal consequences. However, there

is one possibility that turns out to be Weyl invariant:

$$S(X, \gamma, \phi) \propto \int d\sigma d\tau \sqrt{\gamma} R(\gamma) \phi, \quad (6.10)$$

where  $R$  is the two-dimensional curvature scalar. In four dimensions, and without the factor  $\phi$ , this is the Einstein action for gravity. In two dimensions, however, gravity is trivial: the equations of motion obtained by varying  $\int d\sigma d\tau \sqrt{\gamma} R(\gamma)$  with respect to  $\gamma$  vanish identically.

This means that the gravitational action (or (6.10) with constant  $\phi$ ) does not change if we change  $\gamma$  in a continuous way. Hence the action is the same for a sphere and an egg-shaped surface, that can be transformed into each other by continuously varying the metric. However, for two surfaces that cannot be continuously transformed into each other, such as a sphere and a torus, the action may be different. This is indeed the case. It turns out that the integral of  $\sqrt{\gamma} R(\gamma)$  over a manifold is a *topological invariant*. The quantity

$$\chi = \frac{1}{4\pi} \int d\sigma d\tau \sqrt{\gamma} R(\gamma) \quad (6.11)$$

is called the *Euler number* of the surface. It is an integer that does not depend on the metric, but only on the topology, and is given by

$$\chi = 2 - 2g - b - c. \quad (6.12)$$

For example a sphere has Euler number two, a disk (a sphere with one hole) has Euler number 1, and a torus (a sphere with one handle) has Euler number 0.

Unlike  $g$  and  $B$ , a background value for  $\phi$  has no immediate effect on the classical or quantum treatment of the free boson  $X$ . However, the action (6.10) plays an important rôle. Suppose we write the dilaton as  $\phi = c + \tilde{\phi}$ . Here  $c$  is some constant, and  $\tilde{\phi}$  represents the fluctuations of the dilaton field around that constant value. Then the action can be written as  $S_E = \tilde{S}_E + c\chi$  (here we have chosen a convenient normalization for the dilaton action, and we define  $\tilde{S}$  as the  $c$ -independent part of the action). The result is

$$\langle \text{out} | U | \text{in} \rangle = \sum_{\text{topologies}} e^{-c\chi} \int d\Omega \int \mathcal{D}[\text{fields}] \int \mathcal{D}b \mathcal{D}c \mathcal{D}\beta \mathcal{D}\rho e^{-\tilde{S}_E(\text{fields}, b, c, \beta, \rho)} \quad (6.13)$$

We see that  $c$  appears as a topology-dependent weight factor for the terms in the sum. It is customary to define  $e^c = \kappa$ , which is called the string coupling constant. The motivation for this terminology is that each time we add a handle to the surface, we get an extra factor  $\kappa^2$  in the contribution to the amplitude. It is as if there is one factor  $\kappa$  for each emission of a closed string.

The string coupling constant is thus related to the classical value (vacuum expectation value) of a field, the dilaton. In this sense it is not a “parameter” of the theory. We would speak of a parameter if its value had to be defined *a priori*. But since  $\kappa$  is related to a vacuum expectation value of a field, it is possible that some dynamical mechanism (for example the minimization of a potential) determines its value. However, as long as this mechanism is unknown there is not much practical difference between the two descriptions.

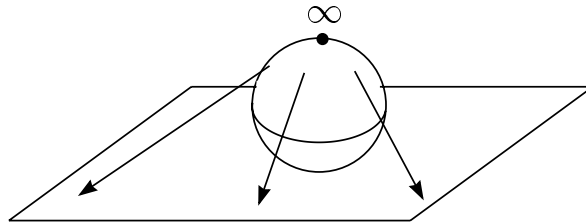
## 6.6 Moduli\*

The parameters of the surfaces integrated over by the  $d\Omega$  integral are called the *moduli* of the surface. They should not be confused with the moduli of a compactification surface: the latter is a space-time surface, whereas here we talk about world-sheets. Mathematically, however, they are the same.

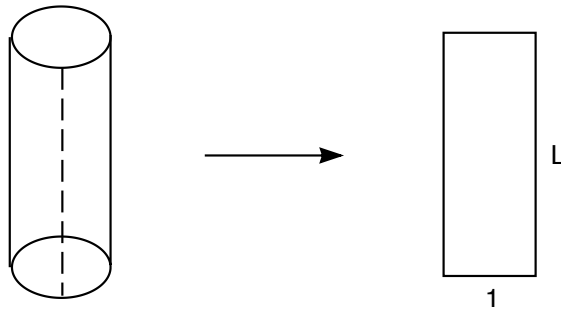
The number of moduli of a surface of genus  $g$  (with  $b = c = 0$ ) is  $2g$ . This implies that a sphere has no moduli. Conformally, all spheres are equivalent.

In general, any manifold can be covered with coordinate patches. On each patch one can go to conformal gauge, so that the metric is  $\eta_{ab}$ . If the manifold has moduli, they are related to the different ways of sewing the patches together.

On a sphere, one can choose a patch that covers the entire sphere except for one point. This patch can be mapped to the two-dimensional plane. Infinity in this plane corresponds to the missing point on the sphere. In this way all spheres are mapped to the same plane, and hence they are all equivalent.

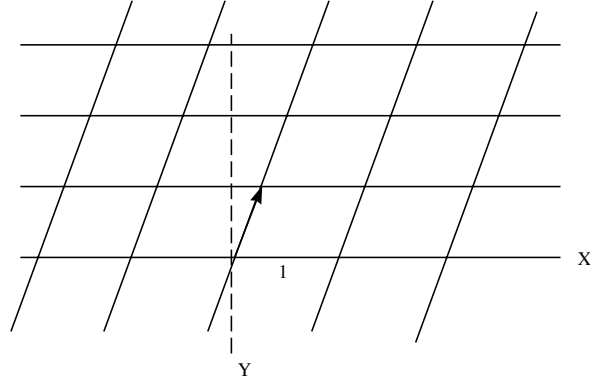


A torus is made by first making a cylinder, and then gluing the ends of the cylinder together. A cylinder has just one modulus. One can map it to a rectangle (by cutting it open), and by scale invariance one can make sure that one side of the rectangle has length one. The length of the other side,  $L$ , can then not be changed anymore, and corresponds to a modulus parameter.



To see that there really is a free parameter one may walk on the surface starting in one corner at a 45 degree angle between the two coordinate axes. Since we are in a metric  $\eta_{ab}$  this angle is unambiguously defined. If  $L = 1$  we end in another corner, if  $L \neq 1$  we end on one of the sides. No coordinate or Weyl transformation can change that.

The gluing of the two ends of the cylinder introduces a second parameter. This parameter can be understood as a rotation of the top of the cylinder before gluing it to the bottom. Usually one represents a torus as a two-dimensional plane modulo a lattice



One of the axes is chosen along the  $x$ -axis, and the length of a basic  $x$ -interval is chosen equal to 1. Then the lattice is specified by a two-dimensional vector as indicated in the picture. The projection of this vector along the  $y$ -axis corresponds to the length of the cylinder, the projection along the  $x$ -axis to the rotation before gluing.

## 6.7 Complex coordinates

When we go to Euclidean space the factors  $e^{-in(\tau\pm\sigma)} \equiv e^{-in(\sigma_0\pm\sigma_1)}$  appearing in mode expansions become  $e^{-n(\sigma_2\pm i\sigma_1)}$ . This suggests that it may be more natural to introduce complex coordinates

$$z = \sigma_2 + i\sigma_1, \quad \bar{z} = \sigma_2 - i\sigma_1 \quad (6.14)$$

This is convenient for other reasons as well. For example, Riemann surfaces are most easily described as complex surfaces (often called “complex curves”, since they have just one complex coordinate). The origin of the simplifications is Cauchy’s theorem, which is a very powerful tool for evaluating contour integrals. Indeed, these techniques are used abundantly in string theory. Of course Cauchy’s theorem is special to two real, or one complex dimension. This is another reason why one can get a lot further with strings than with higher dimensional objects.

In terms of complex coordinates the torus is described by the complex plane modulo a lattice, and the lattice is described by a point in the complex plane. This point is usually called  $\tau^*$ .

The integral over the moduli, previously denote  $\int d\Omega$ , becomes in the case of the torus

$$\int \frac{d^2\tau}{(\text{Im } \tau)^2}, \quad (6.15)$$

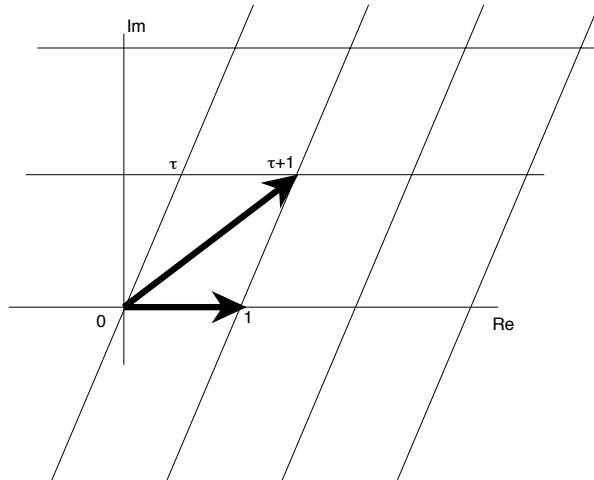
with  $d^2\tau = d \text{Re } \tau d \text{Im } \tau$ . The denominator is a measure factor, that gives the correct weight to different values of  $\tau$ . It emerges from a careful analysis of the path integral. We will not derive it here, but we will see later some justification for its presence.

---

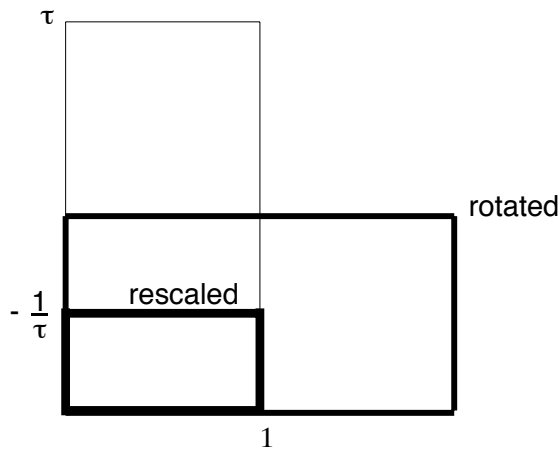
\* This should not be confused with the world-sheet time coordinate  $\tau$  used before. Unfortunately  $\tau$  is used for two different quantities in the string literature. From now on we will use Euclidean coordinates  $\sigma^1$  and  $\sigma^2$ , so that no confusion is possible.

## 6.8 Modular transformations

Up to now we had defined the torus in terms of a lattice. This lattice was defined by two basis vectors, corresponding to the points “1” and “ $\tau$ ” in the complex plane. However, the same lattice – and hence the same torus – can be described just as well by choosing different basis vectors. For example the choice “1”, “ $\tau + 1$ ” clearly describes the same lattice



This can be generalized further. One should keep in mind that the torus was defined by rotating one basis vector along the real axis in the complex plane, and scaling it to 1. The choice of this basis vector is free; we can also choose the direction “ $\tau$ ”. This has the effect of interchanging the two basis vectors. This rotation, combined with a rescaling the new basis vector along the real axis, has the effect of replacing  $\tau$  by  $-\frac{1}{\tau}$ . This is most easily illustrated by taking  $\tau$  purely imaginary:





The set of such transformations of the torus forms a group, called the *modular group*. We have identified two elements of that group, namely

$$\begin{aligned} T : \quad \tau &\rightarrow \tau + 1 \\ S : \quad \tau &\rightarrow -\frac{1}{\tau} \end{aligned} \tag{6.16}$$

It turns out that these two transformations generate the entire group. The most general modular transformation has the form

$$\tau \rightarrow \frac{a\tau + b}{c\tau + d}, \quad a, b, c, d \in \mathbf{Z}; \quad ad - bc = 1. \tag{6.17}$$

This group is isomorphic to  $SL_2(\mathbf{Z})/\mathbf{Z}_2$ . The group  $SL_2$  can be defined by the set of  $2 \times 2$  matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{6.18}$$

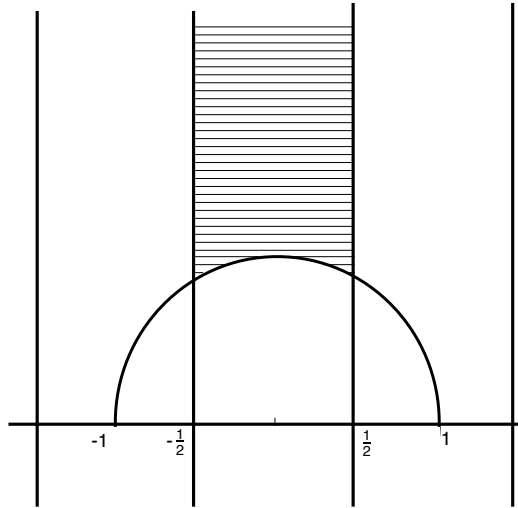
with determinant 1. The group  $SL_2$  contains the element  $-\mathbf{1}$ . In the modular transformation this is indistinguishable from the identity, and for this reason the modular group is actually isomorphic to  $SL_2(\mathbf{Z})/\mathbf{Z}_2$  rather than  $SL_2(\mathbf{Z})$ . One may check that the modular transformations satisfy

$$(ST)^3 = S^2 = 1. \tag{6.19}$$

We see thus that not all parameters  $\tau$  give different tori.

## 6.9 Integration over moduli

This has implications for the integral over moduli. The integral over  $\tau$  is not over the full positive upper half plane, but should be restricted to a region that covers the set of distinct tori just once. An example of such a region is shown below



The entire upper half plane is covered with an infinite number of such regions of different shapes and sizes. For example, the lower part of the strip  $-\frac{1}{2} \leq \text{Re } \tau < \frac{1}{2}$  contains an infinite number of identified regions. This can be seen as follows. The map  $\tau \rightarrow -1/\tau$  takes the outside of the circle to the inside. Suppose we take the dashed region and shift it by  $n$  units, using  $\tau \rightarrow \tau + 1$   $n$  times. This gives an infinity of regions. Now apply  $\tau \rightarrow -1/\tau$ . This maps any such region to the inside of the circle, and in fact within the range  $-\frac{1}{2} \leq \text{Re } \tau < \frac{1}{2}$ . Hence the lower part of the strip  $-\frac{1}{2} \leq \text{Re } \tau < \frac{1}{2}$  contains an infinite number of copies of the “standard” region, the dashed area in the picture.

When we integrate over each such region we get the same answer. At least we *should* get the same answer, since otherwise the result depends on the choice we make among the equivalent regions, and this would clearly be absurd. The requirement that we always get the same answer is called *modular invariance*. Although it is formulated here as a property of the *integral*, it turns out that the only way for it to hold in general (for example also for diagrams with external lines) is that in fact the *integrand* is invariant.

It turns out that the measure (6.15) is already modular invariant by itself. One may verify that

$$\text{Im} \left( \frac{a\tau + b}{c\tau + d} \right) = |c\tau + d|^{-2} \text{Im } \tau \quad (6.20)$$

and that

$$d^2\tau' = |c\tau + d|^{-4} d^2\tau, \quad \tau' = \frac{a\tau + b}{c\tau + d}, \quad (6.21)$$

so that the factors precisely cancel. In other words, the integration measure used in (6.15) is modular invariant.

It follows that the integrand, obtained from the path integral, must also be modular invariant.

## 6.10 Computing the path integral

Now we would like to compute the path integral for the physical degrees of freedom of the bosonic string. For the bosons  $X^i$  this can be done as follows. Since the integral is Gaussian, it gives simply a determinant,  $[\det(-\square)]^{-1/2}$ . But this is not a useful expression.

Instead we will make use of a standard formula for path integrals in quantum mechanics of one variable:

$$\int_{\text{PBC}} \mathcal{D}q e^{-S_E(q)} = \text{Tr } e^{-\beta H} \quad (6.22)$$

Here  $q(t)$  is a coordinate, and one considers a time interval  $0 \leq t \leq \beta$ . The abbreviation “PBC” means periodic boundary conditions. Concretely it means that  $q(0) = q(\beta)$ . On the left-hand side one is integrating over all paths that return to their starting point after a Euclidean “time” interval  $\beta$ .

This result can be straightforwardly generalized to theories with many dynamical variables  $q_k$ . Here we have infinitely many such variables. The quantities  $X^i(\sigma^1)$ , for each

$i$  and for each value of  $\sigma^1$  should all be viewed as dynamical variables  $q_k$ . But this is not a serious difficulty.

In the lattice description of the torus we regard the real axis as the  $\sigma^1$  direction, and the imaginary axis as the Euclidean time direction. If  $\text{Re } \tau = 0$  we would find for the path integral, as a straightforward generalization of (6.22)\*

$$\int \mathcal{D}X e^{-S_E(X)} = \text{Tr } e^{-2\pi \text{Im } \tau H} . \quad (6.23)$$

The only small subtlety here is the factor  $2\pi$ . It appears because the torus as depicted earlier has a periodicity 1 rather than  $2\pi$  along the  $\sigma^1$  direction. To make contact with earlier conventions we had to scale up the entire lattice by a factor of  $2\pi$ , so that we get  $2\pi\tau$  instead of  $\tau$ .

What happens if  $\text{Re } \tau \neq 0$ ? In that case we have to twist the torus before gluing it together again, and the periodic boundary conditions in the Euclidean time direction are defined including such a shift. The operator performing such a shift in the  $\sigma^1$  direction is the momentum operator  $P$ . A shift by an amount  $2\pi \text{Re } \tau$  is achieved by the operator

$$e^{iP(2\pi \text{Re } \tau)} \quad (6.24)$$

The correct result is obtained by inserting this shift operator in the trace, so that we get

$$\int \mathcal{D}X e^{-S_E(X)} = \text{Tr } e^{-2\pi \text{Im } \tau H} e^{iP(2\pi \text{Re } \tau)} . \quad (6.25)$$

The operators  $H$  and  $P$  are the time and space translation operators on the cylinder, and they can be derived from the energy-momentum tensor. We have already seen that  $H = (L_0 - 1) + (\bar{L}_0 - 1)$ . The expression for the momentum operator is

$$P = \frac{1}{2\pi\alpha'} \int_0^{2\pi} d\sigma^1 T_{01} = L_0 - \bar{L}_0 \quad (6.26)$$

Putting everything together we find then

$$\begin{aligned} \int \mathcal{D}X e^{-S_E(X)} &= \text{Tr } e^{-2\pi \text{Im } \tau [(L_0-1) - (\bar{L}_0-1)] + 2\pi i \text{Re } \tau (L_0 - \bar{L}_0)} \\ &= \text{Tr } e^{2\pi i \tau (L_0-1)} e^{-2\pi i \bar{\tau} (\bar{L}_0-1)} \end{aligned} \quad (6.27)$$

---

\* To be precise: an infinite factor has been dropped on both sides of the equation. On the left hand side the integral over constant  $X$  gives infinity; on the right hand the trace is proportional to the normalization of the vacuum  $\langle 0|0 \rangle = \delta(0)$ . This is a momentum conservation  $\delta$ -function which is trivial because there are no external momenta. This factor is due to the translation invariance of the action; the rest of Poincaré invariance gives a finite integral over the rotations of  $X^i$ , which requires no special treatment.

## 6.11 The bosonic string partition function

How do we compute this trace? Consider first of all only the left-moving part of the string. A trace of the kind we are trying to compute can be expanded as

$$\text{Tr } e^{2\pi i\tau(L_0-1)} = \sum_n d_n q^n \quad (6.28)$$

where

$$q \equiv e^{2\pi i\tau} \quad (6.29)$$

and  $d_n$  is the number of states with  $L_0 - 1$  eigenvalue  $n$ . Such a function is usually called a *partition function*. So now it is simply a matter of counting states at each level. This is a simple combinatorial exercise. Remember that the Hilbert space of physical states is obtained by taking the vacuum  $|0\rangle$  and acting with the oscillators  $\alpha_{-n}^i$ ,  $n > 0$ . Suppose we had just one oscillator  $\alpha_{-n}$ . The set of states is then

State	Mass
$ 0\rangle$	-1
$\alpha_{-n} 0\rangle$	$n - 1$
$\alpha_{-n}^2 0\rangle$	$2n - 1$
$\alpha_{-n}^3 0\rangle$	$3n - 1$
...	...

If these were all the states, the trace would yield

$$q^{-1} + q^{n-1} + q^{2n-1} + q^{3n-1} + \dots = q^{-1} \frac{1}{1 - q^n} \quad (6.30)$$

Now we give the oscillator an index  $i$ ,  $1 \leq i \leq D - 2$  ( $D$  will of course have to be equal to 26, but we leave it as a parameter for the moment). All these oscillators act independently on the states we already have, and this means that any additional oscillator has a multiplicative effect on the partition function of the states that were already taken into account. Hence the full contribution of a set of oscillators  $\alpha_{-n}^i$  is

$$q^{-1} \left( \frac{1}{1 - q^n} \right)^{D-2} \quad (6.31)$$

By the same logic, oscillators with a different mode index  $n$  also have a multiplicative effect, so that we get

$$q^{-1} \prod_{n=1}^{\infty} \left( \frac{1}{1 - q^n} \right)^{D-2} \quad (6.32)$$

This is the complete partition function of the left-moving sector. It is multiplied by a similar factor from the right-moving sector, with  $q$  replaced by  $\bar{q}$ . Often the result is

expressed in terms of the *Dedekind  $\eta$ -function*

$$\eta(\tau) = e^{\pi i \tau / 12} \prod_{n=1}^{\infty} (1 - e^{2\pi i n \tau}) . \quad (6.33)$$

This function – with precisely this factor  $e^{(1/24)2\pi i \tau}$  – turns out to have nice properties under modular transformations. We will see them in a moment. Comparing this with the string partition function we see that precisely for  $D = 26$  the result can be expressed completely in terms of  $\eta$ -functions:

$$P_{\text{osc.}}(\tau, \bar{\tau}) = [\eta(\tau)\eta(-\bar{\tau})]^{-24} . \quad (6.34)$$

note that  $\eta(-\bar{\tau}) = (\eta(\tau))^*$ .

But we are not quite finished. We have forgotten that our Hilbert space also contains momentum excitations of the vacuum,  $|p^i\rangle$ . These excitations work completely independently of the oscillators, and again give a multiplicative factor. Since the momenta are continuous, the trace now becomes an integral

$$P_{\text{mom.}}(\tau, \bar{\tau}) \propto \int \frac{d^{D-2}p}{(2\pi)^{D-2}} e^{-2\pi \text{Im } \tau (\frac{1}{2}\alpha' \vec{p}^2)} \quad (6.35)$$

This is a Gaussian integral, which is easy to do. It yields

$$P_{\text{mom.}}(\tau, \bar{\tau}) = [2\pi\sqrt{\alpha' \text{Im } \tau}]^{2-D} \quad (6.36)$$

Putting everything together we find for the bosonic string one-loop path integral (up to normalization)

$$\int \mathcal{D}[\text{all fields...}]_{\text{one-loop}} e^{-S_E(X)} \propto \int \frac{d^2\tau}{(\text{Im } \tau)^2} \frac{1}{(\text{Im } \tau)^{12}} \eta(\tau)^{-24} \eta(\bar{\tau})^{-24} \quad (6.37)$$

## 6.12 Modular invariance of the bosonic string

To check if indeed the bosonic string satisfies the requirement of modular invariance we need the transformation of the  $\eta$  function. This can be found in many books. The  $\tau \rightarrow \tau + 1$  transformation is trivial:

$$\eta(\tau + 1) = e^{i\pi/12} \eta(\tau) \quad (6.38)$$

For the other basic transformation one finds

$$\eta(-1/\tau) = \sqrt{-i\tau} \eta(\tau) \quad (6.39)$$

Hence the combination that appears in the string partition function transforms as

$$|\eta(\tau + 1)|^2 = |\eta(\tau)|^2 ; \quad |\eta(-1/\tau)|^2 = |\tau| |\eta(\tau)|^2 \quad (6.40)$$

The factor  $|\tau|$  in the second transformation cancel against the transformation factor of the momentum contribution,  $[\sqrt{\text{Im } \tau}]^{-1}$  for each dimension. Since these two basic transformations generate the entire modular group, it is clear that the bosonic string is modular invariant.

Note that  $D = 26$  is essential. If  $D \neq 26$  we are left with factors  $q\bar{q}$  to some power if we try to express the partition function in terms of  $\eta$ -functions. These factors are not invariant under  $\tau \rightarrow -1/\tau$ . We discover here a second reason why the critical dimension of the string must be 26.

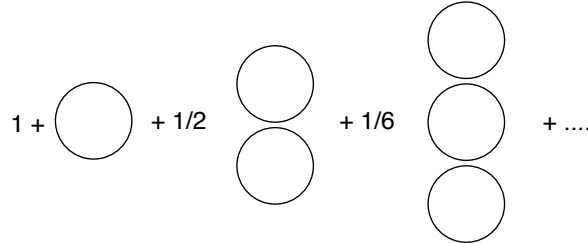
### 6.13 Strings versus particles\*

We have seen that the string spectrum consists of an infinity of modes, which an observer interprets as particles. One may therefore think that a string loop diagram is in some way related to the sum of particle loop diagrams. To see if that is true, we will now do the same computation for a single particle, and at the end sum over all particles.

We follow here the discussion in [1], part II, page 55. The string loop diagrams are supposed to contribute to the “cosmological constant”  $\Lambda$ . For a massive real free boson field theory the relevant expression is

$$e^\Lambda = \int \mathcal{D}\phi \exp \left( - \int d^D x \frac{1}{2} [\partial_\mu \phi \partial^\mu \phi + m^2 \phi^2] \right) = \frac{1}{\det \sqrt{-\square + m^2}} \quad (6.41)$$

The reason we compare the string calculation with the logarithm of the field theory path integral is that the latter generates not only the one loop diagram, but also all disconnected diagrams composed by putting loops together. The string diagram we have computed is however strictly one-loop. Pictorially the field theory path integral is



The factors  $\frac{1}{N!}$  are combinatorial factors, and the whole diagrammatic expansion exponentiates. Hence by taking the logarithm we pick out precisely the one-loop contribution.

The result is equal to

$$\begin{aligned} \Lambda &= -\frac{1}{2} \log \det (-\square + m^2) \\ &= -\frac{1}{2} \text{Tr} \log[-\square + m^2] \\ &= -\frac{1}{2} \int \frac{d^D p}{(2\pi)^D} \langle p | \log(-\square + m^2) | p \rangle \\ &= -\frac{1}{2} \int \frac{d^D p}{(2\pi)^D} \log(p^2 + m^2) \end{aligned} \quad (6.42)$$

There is another way of seeing that this is the right expression. Up to irrelevant factors,  $(p^2 + m^2)^{-1}$  is the propagator of a massive particle. When computing a Feynman diagram one associates with every internal line such a propagator. If one differentiates a propagator with respect to  $m^2$  it splits into two propagators:

$$\frac{d}{dm^2} \frac{1}{p^2 + m^2} = -\frac{1}{(p^2 + m^2)^2} \quad (6.43)$$

The expression

$$\int \frac{d^D p}{(2\pi)^D} \frac{1}{(p^2 + m^2)^N} \quad (6.44)$$

is the loop integral for a one loop diagram with  $N$  external lines, each with zero momentum and trivial coupling. Diagrammatically such a diagram is an  $N$ -sided polygon. By differentiating with respect to  $m^2$  we go from an  $N$ -sided to an  $N + 1$ -sided polygon. Our interest is in a one-loop diagram with zero external lines, which is what we computed in string theory. By inverting the recursion, we should be able to compute that by integrating the  $N = 1$  diagram with respect to  $m^2$ . This of course precisely gives the logarithm.

To evaluate (6.42) we use the following integral representation

$$\log(X) = -\int_0^\infty \frac{dt}{t} e^{-tX} \quad (6.45)$$

This can be “derived” by differentiating w.r.t.  $X$  on both sides. The divergence at  $t = 0$  is independent of  $X$  and drops out when one differentiates. So in fact the correct, regulated expression has a constant term  $+\infty$  on the left-hand side:

$$\log(X) = -\lim_{\epsilon \rightarrow 0} \left[ \int_\epsilon^\infty \frac{dt}{t} e^{-tX} - \int_\epsilon^\infty \frac{dt}{t} e^{-t} \right] \quad (6.46)$$

We drop the  $m$ -independent, but infinite constant. The rest of the calculation goes as follows

$$\begin{aligned} \Lambda &= -\frac{1}{2} \int \frac{d^D p}{(2\pi)^D} \log(p^2 + m^2) \\ &= \frac{1}{2} \int_0^\infty \frac{dt}{t} \int \frac{d^D p}{(2\pi)^D} e^{-t(p^2 + m^2)} \\ &= \frac{1}{2} \int_0^\infty \frac{dt}{t} e^{-tm^2} \int \frac{d^D p}{(2\pi)^D} e^{-tp^2} \\ &= \frac{1}{2} \int_0^\infty \frac{dt}{t} e^{-tm^2} \frac{1}{(4\pi t)^{D/2}} \end{aligned} \quad (6.47)$$

As remarked above, in this calculation an  $m$ -independent infinite term was discarded, proportional to the volume of momentum space. Apart from that, the integral has now an even stronger divergence at  $t = 0$ . The latter divergence can be recognized as the

usual UV divergence in quantum field theory. This can be seen as follows. As seen above, differentiating (6.42)  $N$  times with respect to  $m^2$  gives (6.44), which is the basic integral one encounters when one computes a one-loop diagram with  $N$  external lines. For  $N = D/2$  this diagram diverges logarithmically for large  $p$ , and for  $N < D/2$  there is a power divergence. These are called an ultra-violet divergences. On the other hand, if we differentiate the result of the integration  $N$ -times w.r.t.  $m^2$  we get

$$\sim \int \frac{dt}{t} t^{N-D/2} e^{-tm^2} \quad (6.48)$$

which is also logarithmically divergent for  $N = D/2$  and power divergent for  $N < D/2$ . We conclude that the divergence at  $t = 0$  must be interpreted as the field theory ultraviolet divergence. The first,  $m$ -independent divergence has a different interpretation. It amounts to an infinite shift in the vacuum energy, and can be attributed to the contribution of the ground state energies of the infinitely many oscillators. This is the target space analog of the infinite sum over “ $n$ ” we encountered in string theory. Dropping this term amounts to setting the energy of the vacuum to zero.

If we treat string theory as an infinite set of field theories, we expect to get\*

$$\Lambda_{\text{string}} = - \sum_m \log \sqrt{\det (p^2 + m^2)} \quad (6.49)$$

where the sum is over all physical states in the theory. Using the integral representation of a single particle we get (for the uncompactified bosonic string)

$$\Lambda_{\text{string}} = \frac{1}{2} \int_0^\infty \frac{dt}{t} \left( \frac{1}{4\pi t} \right)^{13} \sum_m e^{-tm^2} \quad (6.50)$$

The sum can be evaluated in terms of the partition function (remember that  $q = e^{2\pi i\tau}$ )

$$\eta^{-24}(\tau)\eta^{-24}(-\bar{\tau}) = (q\bar{q})^{-1} \prod_n (1 - q^n)^{-24} \prod_n (1 - \bar{q}^n)^{-24} \quad (6.51)$$

If we replace  $q$  by  $e^{-m^2}$  this is almost the sum we are looking for, except that the partition function contains more than just physical states. It also includes states with  $m_L \neq m_R$ , and we have seen before that such states are not physical. We can project out these “unmatched” states by integrating over  $\text{Re } \tau$ . For the imaginary part of the partition function we get

$$e^{4\pi \text{Im } \tau} [1 + \dots] \quad (6.52)$$

The tachyon mass is  $m^2 = -4/\alpha'$ . From the leading term we read off

$$(4/\alpha')t = 4\pi \text{Im } \tau \quad (6.53)$$

---

\* We may ignore the spins of the particles here. We simple sum over all different spin states, and treat the particles as scalars.



Hence

$$\sum_m e^{-tm^2} = \int_{-\frac{1}{2}}^{\frac{1}{2}} d(\text{Re}\tau) \eta^{-24}(\tau) \eta^{-24}(-\bar{\tau}) \quad (6.54)$$

with  $\text{Im } \tau = t/(\pi\alpha')$ . This results in the following expression for  $\Lambda$

$$\Lambda_{\text{string}} = \frac{1}{2} \int_0^\infty \frac{dt}{t} \left( \frac{1}{4\pi t} \right)^{13} \int_{-\frac{1}{2}}^{\frac{1}{2}} d(\text{Re}\tau) \eta^{-24}(\tau) \eta^{-24}(-\bar{\tau}) \quad (6.55)$$

Changing integration variables to  $\text{Im } \tau$  we find

$$\Lambda_{\text{string}} = \frac{1}{2} \left( \frac{1}{4\pi^2\alpha'} \right)^{13} \int \frac{d^2\tau}{(\text{Im } \tau)^2} (\text{Im } \tau)^{-12} \eta^{-24}(\tau) \eta^{-24}(-\bar{\tau}) \quad (6.56)$$

Apart from overall factors, which are not of interest here, we see that we have obtained the string loop diagram, with one important difference: instead of integrating over a fundamental domain, as shown in the figure in section (5.9), we are integrating over the entire strip  $-\frac{1}{2} < \text{Re } \tau \leq \frac{1}{2}$ . But this strip includes the true string contribution an infinite number of times. Hence if the string contribution is finite, the field theory sum is necessarily infinite.

Furthermore we see that if we take the standard integration region in string theory, as shown in the figure, the dangerous  $\text{Im } \tau \rightarrow 0$  region is cut out. This is how string theory avoids the field theory divergence. It includes a natural, built in cut-off that is required by its internal consistency.

Nevertheless the bosonic string loop integral is not finite. This is entirely due to the tachyon. Indeed, for a negative mass squared  $m^2 = -A$  we get an integral of the form

$$\int_0^\infty \frac{dt}{t} \left( \frac{1}{4\pi t} \right)^{13} e^{At}, \quad (6.57)$$

which diverges badly for large  $t$ . Massless particles ( $A = 0$ ) do not give a divergence for large  $t$  and massive particles are also completely safe. So if we get rid of the tachyon, we may expect a finite result.

Note that the tachyon divergence comes from the opposite side of the integration region as the UV-divergence. This is usually called an infrared divergence. The origin of such divergences is rather different, and are usually associated with the definition of the vacuum.

## 6.14 Partition function for compactified strings

It is instructive to consider the compactified string partition function. It can be obtained rather easily from the uncompactified string partition function. The only difference is that we do not have continuous components  $p^I$  for the compactified momenta. Since their contribution is absent, we have to drop a factor  $(\sqrt{\text{Im } \tau})^{-N}$  from the bosonic string partition function. However, there are also additional contributions to the mass due to

the “momenta”  $\vec{p}_L$  and  $\vec{p}_R$ , see (4.12). Since these momenta can be assigned to any state, completely independently from the oscillators, this leads to a multiplicative factor in the partition function. This factor is

$$P_\Gamma(\tau) = \sum_{(\vec{p}_L, \vec{p}_R) \in \Gamma} e^{i\pi\tau\vec{p}_L^2} e^{-i\pi\tau\vec{p}_R^2}, \quad (6.58)$$

where we are assuming that the momenta  $(\vec{p}_L, \vec{p}_R)$  lie on a lattice  $\Gamma$ . The compactified string partition function is thus given by

$$P_{\text{comp.}} = P_{\text{uncomp.}} (\sqrt{\text{Im } \tau})^N P_\Gamma \quad (6.59)$$

Since  $P_{\text{uncomp.}}$  is modular invariant, we have to check modular invariance of the remaining factors.

To check  $\tau \rightarrow \tau + 1$  is, as always, easy. We see that this requires that  $\vec{p}_L^2 - \vec{p}_R^2$  is an even integer. We recognize this is the requirement that  $\Gamma$  should be a Lorentzian even lattice.

To derive the other transformation rule requires a trick, the *Poisson re-summation formula*. This works as follows. Suppose we have a lattice  $\Lambda$ , and a continuous function  $f(\vec{w})$ . Consider the quantity

$$\sum_{\vec{w} \in \Lambda} f(\vec{w}) \quad (6.60)$$

The Poisson formula re-writes this sum as a sum over the dual lattice  $\Lambda^*$ . First define

$$F(\vec{z}) = \sum_{\vec{w} \in \Lambda} f(\vec{w} + \vec{z}) \quad (6.61)$$

Obviously  $F(\vec{z})$  is periodic in  $\vec{z}$  with the periodicity of  $\Lambda$ . Hence we may Fourier expand it. The Fourier components of functions with periodicity  $\Lambda$  lie on the dual lattice  $\Lambda^*$

$$F(\vec{z}) = \sum_{\vec{v} \in \Lambda^*} e^{-2\pi i \vec{z} \cdot \vec{v}} F^*(\vec{v}) \quad (6.62)$$

The coefficient  $F^*$  are determined using an inverse Fourier transform:

$$F^*(\vec{v}) = \frac{1}{\text{vol}(\Lambda)} \int_{\text{unit cell of } \Lambda} d^N y e^{2\pi i \vec{y} \cdot \vec{v}} F(\vec{y}), \quad (6.63)$$

where  $\text{vol}(V)$  denotes the volume of the unit cell, *i.e.* the result of the integral of 1 over the unit cell. Now we substitute  $F(y)$ :

$$F^*(\vec{v}) = \frac{1}{\text{vol}(\Lambda)} \int_{\text{unit cell of } \Lambda} d^N y e^{2\pi i \vec{y} \cdot \vec{v}} \sum_{\vec{w} \in \Lambda} f(\vec{w} + \vec{y}), \quad (6.64)$$

and we observe that the sum over all cells combines with the integral over the unit cell to an integral over all of  $\mathbf{R}^N$ , where  $N$  is the dimension of  $\Lambda$ . Hence we may write

$$\begin{aligned} F^*(\vec{v}) &= \frac{1}{\text{vol}(\Lambda)} \int_{\text{unit cell of } \Lambda} d^N y \sum_{\vec{w} \in \Lambda} e^{2\pi i(\vec{w}+\vec{y})\cdot\vec{v}} f(\vec{w} + \vec{y}) \\ &= \frac{1}{\text{vol}(\Lambda)} \int_{\mathbf{R}^N} d^N x e^{2\pi i\vec{x}\cdot\vec{v}} f(\vec{x}) \end{aligned} \quad (6.65)$$

Applying this to  $F(0)$  in (6.61) yields on the one hand

$$F(0) = \sum_{\vec{w} \in \Lambda} f(\vec{w}) \quad (6.66)$$

but on the other hand, according to (6.62),  $F(0)$  can also be written as

$$F(0) = \sum_{\vec{v} \in \Lambda^*} F^*(\vec{v}) \quad (6.67)$$

Using (6.65) and taking out the volume factor, we get the relation

$$\sum_{\vec{w} \in \Lambda} f(\vec{w}) = \frac{1}{\text{vol}(\Lambda)} \sum_{\vec{v} \in \Lambda^*} f^*(\vec{v}) \quad (6.68)$$

with

$$f^*(\vec{v}) = \int_{\mathbf{R}^N} d^N x e^{2\pi i\vec{x}\cdot\vec{v}} f(\vec{x}) \quad (6.69)$$

For our purpose we apply this to the transformed term in the lattice partition function

$$f(\vec{p}_L, \vec{p}_R) = e^{-i\pi\vec{p}_L^2/\tau} e^{i\pi\vec{p}_R^2/\bar{\tau}} . \quad (6.70)$$

It turns out that we should regard  $(\vec{p}_L, \vec{p}_R)$  as a vector on a Lorentzian lattice; only then do we get a simple answer. The foregoing discussion holds also for such lattices. The volume of the unit cell is defined as

$$\text{vol}(\Lambda) = \sqrt{|\det g|} , \quad g_{ij} = \vec{e}_i \cdot \vec{e}_j , \quad (6.71)$$

where  $\vec{e}_i$  is a set of basis vectors. This definition generalizes to the Lorentzian case by simply replacing the inner product by a Lorentzian one. The same is done with all other inner products in the foregoing paragraph. The Poisson re-summation formula allows us to write the lattice sum of (6.70) as a sum over the Lorentzian dual lattice. The summand is

$$f^*(\vec{v}_L, \vec{v}_R) = \int_{\mathbf{R}^{2N}} d^{2N} x e^{2\pi i(\vec{x}_L \cdot \vec{v}_L - \vec{x}_R \cdot \vec{v}_R)} f(\vec{x}_L, \vec{x}_R) . \quad (6.72)$$

As long as  $\text{Im } \tau > 0$  this is a well-defined Gaussian integral, which is easy to do, and yields

$$f^*(\vec{v}_L, \vec{v}_R) = |\tau|^N e^{i\pi\tau\vec{v}_L^2 - i\pi\tau\vec{v}_R^2} \quad (6.73)$$

Hence

$$P_{\Gamma}\left(-\frac{1}{\tau}\right) = \frac{1}{\text{vol}(\Gamma)} |\tau|^N \sum_{(\vec{v}_L, \vec{v}_R) \in \Gamma^*} e^{i\pi\tau\vec{v}_L^2} e^{-i\pi\bar{\tau}\vec{v}_R^2}, \quad (6.74)$$

where  $\Gamma^*$  is the Lorentzian dual of  $\Gamma$ . We observe that the factor  $|\tau|^N$  cancels against the contribution from  $(\text{Im } \tau)^{N/2}$ . The only way to get modular invariance is clearly  $\Gamma = \Gamma^*$  and  $\text{vol}(\Gamma) = 1$ . The latter condition follows in fact from the self-duality, since  $\text{vol}(\Gamma)$  is always the inverse of  $\text{vol}(\Gamma^*)$ .

The conclusion is thus that the torus compactified string is modular invariant if and only the momenta  $(\vec{p}_L, \vec{p}_R)$  lie on a Lorentzian even, self-dual lattice.

## 7 Superstrings

Now we will consider the fermionic string theory discussed in chapter 4. Remember that its critical dimension was 10, and that in light-cone gauge the spectrum was built out of 8 transverse bosons and 8 transverse fermions. The contribution of the bosons to the partition function does not differ from what we have seen in the previous section. It is

$$(\operatorname{Im} \tau)^{-4} (\eta(\tau)\eta(-\bar{\tau}))^{-8} \quad (7.1)$$

Here we have split the coefficient  $a$  in  $L_0 - a$  into two parts, one associated with the bosons and another associated with the fermions. For the bosonic contribution we make a choice per boson as in the case of the bosonic string, and the rest will have to be supplied by the fermions. Only in this way can we express the bosonic partition in terms of  $\eta$ -functions. Now we compute the fermionic contributions.

### 7.1 Free fermion partition functions

### 7.2 Neveu-Schwarz states

Consider first the Neveu-Schwarz sector. We have seen that the modes satisfy the following anti-commutator

$$\{b_r, b_s\} = \delta_{r+s,0} \quad (7.2)$$

and that the vacuum state is annihilated by the negative modes of  $b_r$ .

Let us first see what a single left-moving Neveu-Schwarz field contributes. At the first few levels, we find the following states, denoting the  $L_0$  eigenvalue by  $h$ :

$$\begin{array}{ll} h = 0 & |0\rangle \\ h = \frac{1}{2} & b_{-\frac{1}{2}} |0\rangle \\ h = 1 & \text{none} \\ h = \frac{3}{2} & b_{-\frac{3}{2}} |0\rangle \\ h = 2 & b_{-\frac{3}{2}} b_{-\frac{1}{2}} |0\rangle \\ h = \frac{5}{2} & b_{-\frac{5}{2}} |0\rangle \\ h = 3 & b_{-\frac{5}{2}} b_{-\frac{1}{2}} |0\rangle \\ h = \frac{7}{2} & b_{-\frac{7}{2}} |0\rangle \\ h = 4 & b_{-\frac{7}{2}} b_{-\frac{1}{2}} |0\rangle, b_{-\frac{5}{2}} b_{-\frac{3}{2}} |0\rangle \end{array} \quad (7.3)$$

Note that fermionic oscillators must satisfy the Pauli exclusion principle, so that for example  $b_{-1/2} b_{-1/2}$  is zero. For this reason there is no state at level  $h = 1$ , and we have to go to  $h = 4$  to find more than one state.

It is straightforward to compute the partition function. Each oscillator  $b_{-r}$  can act once or zero times on the ground state. If there were just one oscillator  $b_{-r}$  there would just be two states,  $|0\rangle$  and  $b_{-r}|0\rangle$  with  $h = 0$  and  $h = r$ . The single oscillator contribution is thus  $\text{Tr } q^{L_0} = 1 + q^r$ , where  $q = \exp 2\pi i\tau$ . All oscillators with different modes act independently, and it is easy to see that each contributes via additional factors of this form. Furthermore we have to take into account the shift in the ground state energy. For a free boson it is natural to assign  $-1/24$  to each boson, so that for 24 transverse bosons we get precisely the required  $-1$ . This gives the factor  $q^{1/24}$  in the  $\eta$ -function. For fermionic strings we have seen that the critical dimension is 10, so that there are 8 transverse dimensions. To shift in the ground state energy should be  $-\frac{1}{2}$ , to which the bosons contribute  $-1/3$ . Hence the fermions must contribute  $-1/6$ , or  $-1/48$  per transverse fermion.

The result is thus, for each fermion,

$$P_{\text{NS}}(q) = \text{Tr } q^{L_0 - \frac{1}{48}} = q^{-\frac{1}{48}} \prod_{r=1/2}^{\infty} (1 + q^r) . \quad (7.4)$$

### 7.3 Ramond states

In the Ramond sector the fermionic oscillators are integer moded. The ground state is in this case not a single state, but a set of dimension  $2^{N/2}$ , where  $N$  is the number of fermions. The ground state forms in fact a spinor representation of  $SO(N)$ . For the shift in the vacuum energy we have seen before that it is the same as for a free boson. Hence we get for the contribution to the partition function for a single fermion

$$P_{\text{R}}(q) = \text{Tr } q^{L_0 - \frac{1}{24}} = \sqrt{2} q^{\frac{1}{24}} \prod_{n=1}^{\infty} (1 + q^n) , \quad (7.5)$$

by exactly the same arguments as used above. The normalization looks a bit strange, since one would expect coefficients in partition functions to “count states”. However, for  $N$  fermions we get  $P_{\text{R}}(q)^N$ , and there is then no problem for even  $N$ . If we include left- and right-movers, we get  $P_{\text{R}}(q)^N P_{\text{R}}(\bar{q})^N$ , and then all coefficients are integers for *any*  $N$ .

### 7.4 Boundary conditions

We now have two kinds of fermionic partition functions, NS and R. We have seen before that Neveu-Schwarz fermions appear when one choose anti-periodic boundary conditions around the cylinder, while Ramond fermions appear for periodic boundary conditions.

On the torus this corresponds to boundary conditions along the real axis. One might expect then that it should also be possible to change the boundary conditions along the imaginary axis, since the choice of the axes is merely a matter of convention.

Indeed, this is possible. But first there is something else to be said about the path integral for fermions. A bosonic path integral is an integral over real (or complex)

numbers. For example, in the case of one coordinate  $q(t)$  we integrate over commuting numbers  $q$ , not over operators. In the case of fermions, that would not be correct, because even “classically” fermionic degrees of freedom must anti-commute. This can be taken into account by defining integrals over “anti-commuting c-numbers”, also known as *Grassmann variables*. Using this formalism one can define the path integral, and derive the relation with a trace over a Hilbert space, as we did for bosonic variables. The result is almost the same, except for one small detail:

$$\int_{\text{ABC}} \mathcal{D}\psi e^{-S_E(\psi)} = \text{Tr} e^{-\beta H} \quad (7.6)$$

Here “ABC” stands for anti-periodic boundary conditions. These are the partition functions we have just computed.

We may also consider periodic boundary conditions for fermions, but then we get a different answer. The answer can be obtained by flipping the sign of each fermion before gluing the top of the cylinder to the bottom, thus changing the boundary condition from anti-periodic to periodic. This can be achieved by inserting an operator into the trace. This operator must have eigenvalue  $-1$  for states made with an odd number of fermions, and  $+1$  for states made with an even number. To construct it we introduce the fermion number operator  $F$ .

$$F = \frac{1}{2\pi\alpha'} \int_0^{2\pi} \psi^\dagger \psi . \quad (7.7)$$

In terms of modes,

$$F = \sum_r : b_r b_{-r} ; F = \sum_r : d_r d_{-r} \quad (7.8)$$

Now we insert into the trace the operator  $(-1)^F$ . This operator has the property

$$(-1)^F b = -b(-1)^F ; \quad (-1)^F d = -d(-1)^F , \quad (7.9)$$

for any fermionic oscillator  $b_n$  or  $d_n$ . Then

$$\int_{\text{PBC}} \mathcal{D}\psi e^{-S_E(\psi)} = \text{Tr} (-1)^F e^{-\beta H} \quad (7.10)$$

To compute such a trace we must know how  $(-1)^F$  acts on the ground states. In the NS sector it is natural to require that the vacuum has zero fermion number.

$$F |0\rangle = 0 ; \quad (-1)^F |0\rangle = |0\rangle \quad (7.11)$$

In the Ramond sector it is a bit more subtle. We have seen that  $d_0$  transforms the set of ground states into themselves. On the other hand  $d_0$  changes the  $(-1)^F$  eigenvalue when it acts. The only way out is then that the ground states  $|a\rangle$  must split into two subsets, one with eigenvalue  $+$  and the other with eigenvalue  $-1$ .

It is quite easy to see how this insertion changes our expressions for the partition functions. Each fermionic oscillator now contributes a factor  $1 - q^r$  instead of  $1 + q^r$ . Hence

$$P_{\text{NS}}(q)_{\text{PBC}} = \text{Tr} (-1)^F q^{L_0 - \frac{1}{48}} = q^{-\frac{1}{48}} \prod_{r=\frac{1}{2}}^{\infty} (1 - q^r) . \quad (7.12)$$

In the Ramond sector we get

$$P_{\text{R}}(q)_{\text{PBC}} = 0 \quad (7.13)$$

since the trace over  $(-1)^F$  on the ground state manifestly vanishes. It is easy to see that the same is true for all excited states as well.

Now we can compute four kinds of fermionic partition functions, labelled by the periodicities along each of the two cycles of the torus. For  $N$  fermions the result is

AA	$\text{Tr}_{\text{NS}} q^{L_0 - \frac{c}{24}}$	$\left(\frac{\theta_3}{\eta}\right)^{N/2}$
AP	$\text{Tr}_{\text{NS}} (-1)^F q^{L_0 - \frac{c}{24}}$	$\left(\frac{\theta_4}{\eta}\right)^{N/2}$
PA	$\text{Tr}_{\text{R}} q^{L_0 - \frac{c}{24}}$	$\left(\frac{\theta_2}{\eta}\right)^{N/2}$
PP	$\text{Tr}_{\text{R}} (-1)^F q^{L_0 - \frac{c}{24}}$	$\left(\frac{\theta_1}{\eta}\right)^{N/2}$

Here the letters ‘‘AP’’ indicate anti-periodicity along the ‘‘space’’ direction and periodicity along the ‘‘time’’ direction on the torus, etc. It turns out that these four partition functions can be expressed in terms of standard mathematical functions, namely the Jacobi  $\theta$ -functions and the Dedekind  $\eta$ -function, which we have already seen. The  $\theta$ -functions are defined as follows

$$\theta \begin{bmatrix} a \\ b \end{bmatrix} (z|\tau) = \sum_n e^{i\pi[(n+a)^2\tau + 2(n+a)(z+b)]} \quad (7.14)$$

with the additional definitions

$$\theta_1 = \theta \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}; \quad \theta_2 = \theta \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}; \quad \theta_3 = \theta \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \quad \theta_4 = \theta \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} \quad (7.15)$$

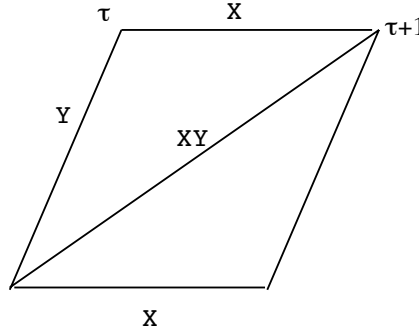
The last column above indicates the identification of each partition function with ratios of  $\theta$ - and  $\eta$ -functions. It is far from obvious that these ratios are equal to the formulas we derived above. But one may check by expanding in  $q^n$  that indeed equality holds for any finite number of levels. Of course the equality can be proved to all orders, but we will not do that here.

Note that the Jacobi  $\theta$ -functions have two arguments, but we are only using them at  $z = 0$  here. The function  $\theta_1(z|\tau)$  vanishes for  $z = 0$ , as does the partition function in the PP sector, but it can be made plausible that the identification given here is the correct one. To do this one may examine the dependence on the first argument, which can be done by considering tori with external legs. Then the contribution from the ‘‘PP-sector’’ does not vanish, and the identification with  $\theta_1$  makes sense.



## 7.5 Modular transformations

Now we discuss modular invariance. Clearly modular transformations change the fermion boundary conditions. For example, the transformation  $\tau \rightarrow -\frac{1}{\tau}$  interchanges the two cycles (“space” and “time”) on the torus, and hence it interchanges AP and PA. The transformation  $\tau \rightarrow \tau + 1$  maps XY to X(XY) as shown in the figure, where X and Y stand for A or P, and the multiplication rule is AA=P, AP=A and PP=P. In other words, it interchanges AA and AP. Since these two transformations generate the modular group, we generate all permutations of AA, AP and PA. On the other hand PP transforms into itself.



The explicit form of the modular transformations can be worked out by means of the Poisson re-summation formula, applied to the  $\theta$ -function form of the partition function. Note that this form is rather reminiscent of a summation over a lattice, namely the lattice of integers. Using this Poisson re-summation for this lattice it is rather easy to derive

$$\begin{aligned} \theta_1\left(-\frac{1}{\tau}\right) &= -i\sqrt{-i\tau}\theta_1(\tau); & \theta_2\left(-\frac{1}{\tau}\right) &= \sqrt{-i\tau}\theta_4(\tau); \\ \theta_3\left(-\frac{1}{\tau}\right) &= \sqrt{-i\tau}\theta_3(\tau); & \theta_4\left(-\frac{1}{\tau}\right) &= \sqrt{-i\tau}\theta_2(\tau) \end{aligned} \quad (7.16)$$

$$\begin{aligned} \theta_1(\tau + 1) &= e^{i\pi/4}\theta_1(\tau); & \theta_2(\tau + 1) &= e^{i\pi/4}\theta_2(\tau); \\ \theta_3(\tau + 1) &= \theta_4(\tau); & \theta_4(\tau + 1) &= \theta_3(\tau) \end{aligned} \quad (7.17)$$

We have already seen that

$$\eta\left(-\frac{1}{\tau}\right) = \sqrt{-i\tau}\eta(\tau); \quad \eta(\tau + 1) = e^{i\pi/12}\eta(\tau) \quad (7.18)$$

## 7.6 The modular invariant partition function

To construct something modular invariant we clearly need some sort of sum over the various boundary conditions. Let us consider the correct number of fields for the fermionic string, namely eight transverse bosons and fermions. The four partition functions are then of the form

$$P_i(\tau) = \frac{(\theta_i(\tau))^4}{\eta^{12}(\tau)} \quad (7.19)$$

Under modular transformations

$$\begin{aligned} P_1\left(-\frac{1}{\tau}\right) &= \tau^{-4}P_1(\tau) ; & P_2\left(-\frac{1}{\tau}\right) &= \tau^{-4}P_4(\tau) \\ P_4\left(-\frac{1}{\tau}\right) &= \tau^{-4}P_2(\tau) ; & P_3\left(-\frac{1}{\tau}\right) &= \tau^{-4}P_3(\tau) \end{aligned} \quad (7.20)$$

$$\begin{aligned} P_1(\tau + 1) &= P_1(\tau) ; & P_2(\tau + 1) &= P_2(\tau) \\ P_3(\tau + 1) &= -P_4(\tau) ; & P_4(\tau + 1) &= -P_3(\tau) \end{aligned} \quad (7.21)$$

The full partition function contains a factor  $[\text{Im}(\tau)]^{\frac{(2-D)}{2}}$  just as for the bosonic string. This cancels factors  $\tau^{-4}$  in the  $\tau \rightarrow -1/\tau$  transformation. Apart from that we see a sign in the transformation between  $P_3$  and  $P_4$ . This is cancelled by taking as the partition function

$$P_{\text{GSO}}(\tau) = \frac{1}{2}(P_3(\tau) - P_4(\tau) - P_2(\tau) + xP_1(\tau)) \quad (7.22)$$

Note that the coefficient of  $P_1$  is not determined by modular invariance, because it only transforms into itself. By considering modular invariance at two loops one does get a constraint on  $x$ , and it turns out that  $x = \pm 1$ .

The full partition function is then

$$P_{\text{II}}(\tau, \bar{\tau}) = [\text{Im}(\tau)]^{\frac{(2-D)}{2}} P_{\text{GSO}}(\tau) P_{\text{GSO}}(-\bar{\tau}) \quad (7.23)$$

Several points require some more discussion. In the foregoing paragraph ‘‘GSO’’ stands for Gliozzi, Scherk and Olive, who invented this construction in a somewhat different way. The subscript ‘‘II’’ stands for ‘‘type-II’’, the name given for historical reasons to this kind of string theory. Let us see what the linear combination does to the spectrum. In the Neveu-Schwarz sector the combination

$$\frac{1}{2}(P_3 - P_4) = \text{TrNS} \left[ \frac{1}{2}(1 - (-1)^F) \right] q^{L_0 - \frac{1}{2}} \quad (7.24)$$

is a partition function with a projection operator inserted. Indeed, the combination  $\frac{1}{2}(1 - (-1)^F)$  is a projection operator, that removes all states from the spectrum that are created by an *even* number of fermion operators (the ‘‘GSO-projection’’). This is good news, because the tachyon corresponds to the state  $|0\rangle$  with  $F = 0$ . Hence in this string theory the tachyon is gone! At the first excited level we find the states

$$b_{-1/2}^i |0\rangle \quad (7.25)$$

which stay in the spectrum (note that the normalization factor  $\frac{1}{2}$  ensures that there is precisely one such multiplet). As we have already seen, these states are massless. Together with the ones from the right-moving sector they produce a graviton, a dilaton and a  $B^{\mu\nu}$ .

In the Ramond sector a similar projection occurs. Here we have

$$\frac{1}{2}(P_2 \pm P_1) = \text{Tr}_R \left[ \frac{1}{2}(1 \pm (-1)^F) \right] q^{L_0} . \quad (7.26)$$

To interpret this we have to think about the action of  $(-1)^F$  on the ground state. We have seen that  $(-1)^F$  anti-commutes with all the  $d_0^i$  operators, and we have also seen that  $d_0^i$  can be represented – up to a factor  $\sqrt{2}$  – by the Dirac  $\gamma$  matrices. Then there is a natural candidate for  $(-1)^F$ , namely

$$\gamma^* = i^{N/2} \prod_i \gamma_i \tag{7.27}$$

(the factor  $i^{N/2}$  is inserted to make sure that  $\gamma^*$  is Hermitean). This matrix is the generalization to  $N$  Euclidean\* dimensions of the familiar operator  $\gamma^5$  in four dimension. In four dimensions the projection operators  $\frac{1}{2}(1 \pm \gamma^5)$  are also familiar. They project on fermions with chirality  $+$  or  $-$ . The former have their spin aligned with the direction of motion, and the latter have it anti-aligned. States with definite chirality must be massless, since otherwise we can flip the velocity relative to the spin by “overtaking” the particle.

In other dimensions the notion of spin changes, but chirality can still be defined, namely as the eigenvalue of  $\gamma^*$ . Hence the operators  $\frac{1}{2}(1 \pm \gamma^*)$  project on the two chiralities. In case of the 10-dimensional fermionic string the Ramond ground state has multiplicity  $2^{(2-D)/2} = 16$ . It consists out of 8 states with chirality  $+$  and eight with chirality  $-$ . We see thus that the GSO-projection selects half the states. We can choose the chirality by selecting the  $+$  sign or the  $-$  sign in front of  $P_1$ . Since we can do that for left- and right-moving modes separately, there are in fact four type-II strings, which we could denote as  $\text{II}_{++}$ ,  $\text{II}_{+-}$ ,  $\text{II}_{-+}$  and  $\text{II}_{--}$ . But by making a reflection in 10 dimensions we can flip all chiralities simultaneously, so that  $\text{II}_{++}$  is equivalent to  $\text{II}_{--}$  and  $\text{II}_{+-}$  to  $\text{II}_{-+}$ . Hence there are really only two string theories, called type-IIA (opposite chiralities) and type-IIB (the same chiralities).

Note that the Ramond partition function in  $P_{\text{GSO}}$  is multiplied with  $-1$ . This is good news, because we have seen that Ramond states are space-time fermions. A well-known feature of fermions is that they contribute to loop diagrams with a  $-1$  sign. This is closely related to the spin-statistics relation. Since the torus diagram is a one-loop diagram, bosons and fermions should contribute to it with opposite sign. Remarkably, this is automatically guaranteed by modular invariance.

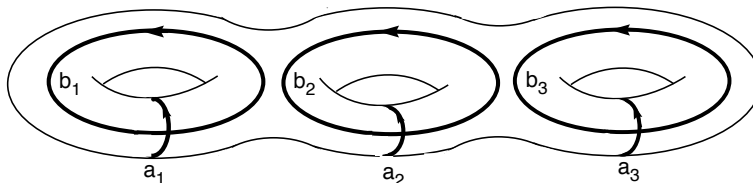
## 7.7 Multi-loop modular invariance

In principle one should also worry about higher loop modular invariance. It turns out that this gives very few constraints. From modular invariance for two-loop surfaces we do learn one interesting fact, the normalization of the “ $\theta_1$ ” contribution.

The essence of the argument goes as follows. Consider a two-loop diagram in the limit where it almost separates into two one-loop tori. In that limit the two-loop “partition function” factorizes into two one-loop partition functions, which are described by Jacobi  $\theta$ -functions. An  $n$ -loop surface has  $2n$  independent loops that cannot be contracted to a point. This is shown below for the three-torus.

---

\* Remember that we are only considering the transverse fermions here.



On each loop the fermions can have boundary condition  $+$  or  $-$ . This gives a total of  $2^{2n}$  boundary conditions on an  $n$ -loop surface: Four for the torus, as we have seen, 16 for the two-torus, etc. On the two-torus there are then 16 distinct terms in the partition function. In the limit two-torus  $\rightarrow$  torus  $\times$  torus those 16 partition functions approach the product  $P_i P_j$ . It turns out that modular transformations of the two-torus mix up the boundary conditions. The 10 combinations  $(i, j) = (2 \dots 4, 2 \dots 4)$  and  $(1, 1)$  are transformed into each other, and the six combinations  $(1, 2 \dots 4)$  and  $(2 \dots 4, 1)$  are also mixed with each other. Invariance under modular transformations at two loops thus requires the coefficient of  $P_1 P_1$ ,  $x^2$ , to be equal to the coefficient of  $P_3 P_3$ , namely 1 (there is no sign change in this transformation).

## 7.8 Superstrings

The complete massless spectrum consists out of states coming from four sectors: NS-NS, NS-R, R-NS and R-R. The NS-NS states produce the graviton, the dilaton and the  $B^{\mu\nu}$ , as mentioned above. The NS-R and R-NS produce, in light-cone coordinates, a chiral spinor times a vector. In the type-IIA string NS-R and R-NS sectors have states with opposite space-time chirality, whereas in the type-IIB theory they have the same chirality. A theory which is invariant under chirality-flip of all particles is called non-chiral. We see thus that type-IIA is non-chiral, whereas type-IIB is chiral. Note that the standard model is chiral: for example, left-handed fermions couple to the  $W$  boson whereas right-handed fermions do not.

Covariantly, the product “spinor times vector” produces a gravitino and an ordinary spinor with opposite chirality. In four dimensions a gravitino is a particle with spin  $3/2$ . In any dimension, the presence of a massless gravitino in the spectrum implies that the spectrum is supersymmetric.

## 7.9 Supersymmetry and Supergravity\*

We have already discussed supersymmetry in two dimensions. The basic idea is the same in any dimension. One introduces an infinitesimal transformation which takes bosons into fermions, and vice-versa. In general there can be several supersymmetries.

In gravitational theories the supersymmetry acts in particular on the graviton. Each supersymmetry produces one gravitino. In the present case we get in fact two gravitini: one from R-NS and one from NS-R. One says therefore that the type-IIA and -IIB theories have N=2 supersymmetry. Theories with gravity that are supersymmetric are in fact invariant under *local* supersymmetry. This means that they are invariant under infinitesimal transformations with a space-time dependent parameter  $\epsilon(x)$ . Such theories are called supergravity theories.

Given a number of supersymmetries one can work out the possible *super-multiplets*, combinations of bosons and fermions that are closed under the action of the supersymmetry. For example in two dimensions a boson and a fermion form a supermultiplet. As one may expect, the sizes of the multiplets increase with the number of supersymmetries. In four dimensions, the maximum number is N=4 if one wants to restrict to global supersymmetry (*i.e.* no graviton and gravitino), and N=8 if one does not want to have particles of spin larger than 2. It turns out to be impossible to make interacting theories with particles of spin larger than 2, so that the absolute maximum number of supersymmetries in four dimensions is 8. Similar maxima exist in other dimensions; for  $D = 5, 6$  the maximum is N=4; for  $D = 7, 8, 9$  and 10 the maximum is N=2, and for  $D = 11$  the maximum is 1. Beyond  $D = 11$  there are no supergravity theories at all!

## 7.10 Supersymmetry for the massive states

We see thus that the type-II theories produce as their massless spectrum N=2 supergravity theories. Indeed they produce precisely the only two known N=2 supergravity theories in ten dimensions. These theories were constructed independently of string theory. Supersymmetry implies that the number of physical bosonic states must be equal to the number of physical fermionic states. In light-cone gauge, the bosonic states from one side of the theory are a transverse vector  $b_{-1/2}^i |0\rangle$  with  $D - 2$  components, and the fermionic states are a chiral spinor  $|a\rangle$  with  $\frac{1}{2} \times 2^{(D-2)/2}$  components (there is a factor  $\frac{1}{2}$  because of the projection on one chirality). For  $D = 10$  these numbers are both equal to 8, for other dimensions they differ. String theory has in addition to these massless states infinitely many massive states, on which supersymmetry has to act as well. The action of supersymmetry has a spectacular effect on the partition functions. The  $\theta$ -functions satisfy the so-called “abstruse” identity of Jacobi:

$$(\theta_3)^4 - (\theta_4)^4 - (\theta_2)^4 = 0 \tag{7.28}$$

The word “abstruse” means “difficult to comprehend”, but to quote Gliozzi, Scherk and Olive: “This identity is not abstruse to us because it is a consequence of supersymmetry at the massive string levels”. Indeed, the identity is a consequence of the fact that at every level the number of fermions and bosons is equal, and that bosons and fermions contribute to the partition function with opposite sign!\*

---

\* Technical remark: Only physical boson and fermion degrees of freedom should be counted. These are also called the on-shell degrees of freedom. For example a massless vector boson has  $D-2$  physical degrees of freedom. In the light-cone gauge one only sees physical degrees of freedom, so that one automatically obtains the correct count.

## 7.11 Ramond-Ramond particles\*

The sector we have not considered yet is the RR-sector. The particles coming from this sector are products of fermions and fermions, and hence have “integer spin” (more precisely, they are in tensor representations of the Lorentz-group  $SO(D - 1, 1)$  rather than in spinor representations) and bosonic statistics. To work out their representations requires some knowledge about computing tensor products of two spinor representations in  $SO(N)$ . The answer is that in the type-IIA theory the RR-sector contains a vector boson and a rank-3 anti-symmetric tensor, while in the type-IIB theory one gets a scalar, an anti-symmetric rank-2 tensor and a self-dual anti-symmetric rank-4 tensor\*.

## 7.12 Other modular invariant string theories

In the foregoing discussion we have found a solution to the condition of modular invariance by imposing the condition on left- and right-movers separately. But this is not necessary. Instead we might consider

$$[\text{Im}(\tau)]^{\frac{(2-D)}{2}} \sum_{ij} a_{ij} P_i(\tau) P_j(-\bar{\tau}) \quad (7.29)$$

and require invariance under modular transformations. Assuming that the  $P_1$  contributions behave as before we find that there are two independent solutions, namely the type-II strings seen above and furthermore (with a  $P_1$ -term added)

$$P_0(\tau, \bar{\tau}) = \frac{1}{2} [\text{Im}(\tau)]^{\frac{(2-D)}{2}} [P_3(\tau)P_3(-\bar{\tau}) + P_4(\tau)P_4(-\bar{\tau}) + P_2(\tau)P_2(-\bar{\tau}) \pm P_1(\tau)P_1(-\bar{\tau})] \quad (7.30)$$

This is called the type-0 string theory. It also comes in two kinds, namely type-0A and type-0B, which have opposite signs in the last term. These theories have a tachyon, because the NS ground states is not projected out, and they have no space-time fermions since there are no NS-R and R-NS sectors. They are not supersymmetric. In fact it can be shown that supersymmetry forbids a tachyon. The converse is not true: there do exist non-supersymmetric strings without tachyons.

## 7.13 Internal fermions

When we discussed fermions on the world-sheet we started with fermions  $\psi^A$  that have an “internal” index A rather than a space-time index  $\mu$ . If one builds theories with such fermions, the issue of modular invariance also arises.

We are now back at the bosonic string. We have seen that if we have  $2N$  internal fermions, the resulting string theory lives in  $26 - N$  dimensions. The partition function

---

\* perhaps surprisingly, rank-4 tensors in 10 dimension can have a chirality + or -, just like fermions, and one finds a rank-4 tensor of chirality + or - (depending in whether one considers the  $\text{II}_{++}$  or the  $\text{II}_{--}$  version of the IIB string). Such tensors are called self-dual or anti-self-dual.

can be built out of functions

$$F_i(\tau) = \frac{\theta_i(\tau)}{\eta(\tau)}, \quad (7.31)$$

and similarly for the right-moving modes. A function  $F_i$  represents a *pair* of fermions. A partition function of a fermionic string contains some combination of  $N$  such functions times a product of  $24 - N$  functions  $\eta^{-1}$  originating from the transverse bosons  $X^i$ . Hence the full partition function is a linear combination of terms of the form

$$\frac{1}{\eta(\tau)^{24-N}} \prod_{m=1}^N F_{i_m}(\tau) = \frac{1}{\eta(\tau)^{24}} \prod_{m=1}^N \theta_{i_m}(\tau) \quad (7.32)$$

The  $\eta$ -function to the power 24 is completely modular invariant, and hence the  $\theta$ -functions must be modular invariant as well (note that we may ignore the factors  $\sqrt{\tau}$ , since they will cancel no matter what we do). For simplicity we will try to make the other factors cancel separately for left- and right-movers. From the transformation of  $\theta_2$  under  $\tau \rightarrow \tau + 1$  we see then that  $N$  must be a multiple of 8: otherwise there can be no  $\theta_2$  term, and then the others must vanish as well. If  $N = 8$  there is indeed a simple solution, which we write in terms of the ratio's  $F$  as

$$F_8 = \frac{1}{2} [(F_3)^8 + (F_4)^8 + (F_2)^8 \pm (F_1)^8] \quad (7.33)$$

If  $N = 16$  there are two solutions, namely  $(F_8)^2$  and  $F_{16}$ , the latter defined in the obvious way.

The free fermion Lagrangian for  $2N$  fermions is symmetric under  $SO(2N)$  rotations

$$\psi_A = \sum_B O_A^B \psi_B \quad (7.34)$$

where  $O$  is an orthogonal matrix. (Note that since  $\psi_A$  is real  $SU(2N)$  is not a symmetry). This symmetry can be seen in the spectrum provided that the boundary conditions respect it. In other words, only if all  $2N$  fermions always have the same boundary conditions. If we consider for example the  $F_{16}$  theory we expect the states in the spectrum to form  $SO(32)$  multiplets. Indeed,  $b_{1/2}^A$  transforms as an  $SO(32)$  vector. For the  $(F_8)^2$  theory we only expect an  $SO(16) \times SO(16)$  symmetry.

Consider now a theory in 18 dimensions plus 16 internal fermions. The only way to construct a partition function that is separately modular invariant on the right and the left is

$$(\text{Im } \tau)^{-8} \eta(\tau)^{-16} \eta(-\bar{\tau})^{-16} F_8(\tau) F_8(-\bar{\tau}) \quad (7.35)$$

Let us compute the spectrum of this theory, starting with the left-moving NS sector. Since the relative sign between  $F_3$  and  $F_4$  is positive, we keep all the states created by an even number of  $b$ 's. Hence the lowest states are the tachyon and the states  $b_{-1/2}^A b_{-1/2}^B |0\rangle$ . The latter are massless and form an anti-symmetric tensor of the symmetry group,  $SO(16)$ , with dimension 120. In the Ramond sector we get as the lowest state a chiral  $SO(16)$

spinor. This has  $\frac{1}{2}2^8 = 128$  components. Furthermore there is a space-time vector created by  $\alpha_{-1}^\mu$ . When we combine left and right sector we get the following massless states: 248+248 vector bosons,  $248 \times 248$  scalars, and of course the graviton, the dilaton and a  $B_{\mu\nu}$ .

This is probably reminiscent of something we have seen before (see appendix C): these are precisely the states that one obtains if one compactifies the bosonic string to 18 dimensions on the  $(E_8)_L \times (E_8)_R$  Lorentzian ESDL. It is now not a very bold conjecture that perhaps these theories are the same. This is indeed true, and we can get additional evidence for this conjecture by showing that the state multiplicities are the same at any level, as a consequence of the following identity

$$\frac{1}{2} \left( \left( \frac{\theta_3}{\eta} \right)^8 + \left( \frac{\theta_4}{\eta} \right)^8 + \left( \frac{\theta_2}{\eta} \right)^8 \right) = \frac{1}{\eta^8} \sum_{\vec{v} \in E_8} e^{\pi i \tau \vec{v}^2} \quad (7.36)$$

This identity is easy to prove if we write  $\theta_i$  as a sum, as in 7.14). This equivalence is an example of bosonization. On the left hand side we have a theory constructed out of 16 free fermions, on the right hand side a theory constructed out of 8 free bosons. Indeed, in general a pair of fermions with appropriate boundary conditions is equivalent to a free boson on an appropriate torus.

Similarly the two fermionic theories in  $26 - 16 = 10$  dimensions can be identified with the  $E_8 \times E_8$  and  $D_{16}$  Euclidean ESDL's in 16 dimensions. If we combine left and right-movers we can in fact built three theories:  $(E_8 \times E_8)_L \times (E_8 \times E_8)_R$ ,  $(D_{16})_L \times (D_{16})_R$  and also the ‘‘hybrid’’ combination  $(E_8 \times E_8)_L \times (D_{16})_R$ . The partition functions are of course

$$(\text{Im } \tau)^{-4} \eta(\tau)^{-8} \eta(-\bar{\tau})^{-8} F_{16}(\tau) F_{16}(-\bar{\tau}) \quad (7.37)$$

for the  $(D_{16})_L \times (D_{16})_R$  theory, and obvious variations on this theme for the other two.

It should be emphasized that there are far more bosonic string compactifications; we have only considered here the ones that have separate modular invariance in the left and right.



## 8 Heterotic strings

The foregoing discussion leads us naturally to the idea of combining the right sector of the type-II string with the left sector of one of the 10-dimensional compactified bosonic strings. Since both are separately modular invariant in the left and right sectors, cutting each theory in half and gluing the parts back together should give a modular invariant theory. This thought is indeed correct, and the result was so important that a new name was given to it: the *heterotic string*.

Since there are two distinct left-moving partition functions in the 10-dimensional bosonic string, we can build two heterotic strings. The partition functions are given by

$$\frac{1}{(\operatorname{Im} \tau)^4} P_{\text{GSO}}(-\bar{\tau}) \eta(\tau)^{-8} F_8(\tau) F_8(\tau) \quad (8.1)$$

and

$$\frac{1}{(\operatorname{Im} \tau)^4} P_{\text{GSO}}(-\bar{\tau}) \eta(\tau)^{-8} F_{16}(\tau) \quad (8.2)$$

These two are called respectively the  $E_8 \times E_8$  heterotic string and the  $SO(32)$  heterotic string. In principle we also have the option of choosing the sign in front of  $P_1$  in  $P_{\text{GSO}}$ . But this only flips the chirality of all the fermions, and has no observable effect, because there is no other chirality to compare with.

### 8.1 The massless spectrum

Note first of all that the spectrum does not contain a tachyon. The left-moving bosonic sector contains one, but the right-moving superstring does not. Hence there is no physical tachyon state, because the condition  $M_L = M_R$  cannot be satisfied.

Analyzing the spectrum of the heterotic strings is easy. The right sector gives us a vector representation and a spinor representation of the transverse  $SO(8)$  group in 10 dimensions. The left sector gives as a vector of  $SO(8)$  plus a number of scalars from the  $(F_8)^2$  or  $F_{16}$  part of the partition function. These scalars come from the NS parts in both cases, and from the R part in the case of  $F_8$ . The total number is

$$\begin{aligned} (F_8)^2 : \quad & 2 \times \left( \frac{1}{2} (16 \times 15)_{\text{NS}} + \frac{1}{2} (2^8)_{\text{R}} \right) = 496 \\ (D_{16}) : \quad & \frac{1}{2} (32 \times 31)_{\text{NS}} = 496 \end{aligned} \quad (8.3)$$

Note that in the first case the subscript ‘‘R’’ refers only to the first (or second) group of 16 fermions. The other group of 16 is always in the NS ground state. The total ground state mass contribution for the 8 bosons, 16 NS fermions and 16 R fermions is thus  $(8 \times (-\frac{1}{24}) + 16 \times (-\frac{1}{48}) + 16 \times (+\frac{1}{24})) = 0$ . For the second case the analogous computation yields  $(8 \times (-\frac{1}{24}) + 32 \times (+\frac{1}{24})) = 1$ , and hence the Ramond ground state is massive.

The complete massless spectrum is thus ( $V$  =vector and  $S$  =spinor)

$$(V + S) \times (V + 496) \quad (8.4)$$

which yields the usual graviton,  $B_{\mu\nu}$  and dilaton (from  $V \times V$ ), 496 massless vector bosons, a gravitino and an ordinary spinor with opposite chirality (from  $S \times V$ ) and 496 additional spinors.

## 8.2 Supersymmetry\*

The presence of the gravitino suggests that there must be supersymmetry. Indeed, this theory has  $N = 1$  supersymmetry (only  $N = 1$  because in this case there is only one gravitino). As for type-II it is in fact a local supersymmetry, *i.e.* supergravity. Any symmetry must act on all fields in a theory. Since supersymmetry switches bosons and fermions, it cannot act trivially. Hence it must act non-trivially on the 496 vector bosons. Obviously, they are transformed into the 496 fermions and vice-versa. In any supersymmetric theory vector bosons are accompanied by fermions in the same representation (*i.e.* the adjoint representation). These fermions are called *gaugino's* (in the supersymmetric extensions of the standard model one speaks of photino's, gluino's, Wino's and Zino's; experiments are looking for such particles, but so far they haven't been found).

## 8.3 Gauge symmetry\*

Just as the presence of a “vector-spinor” or gravitino implies supergravity, in string theory the presence of a vector boson always implies the existence of a local internal symmetry or, in other words, a *gauge symmetry*. Gauge interactions can be seen as a generalization of electrodynamics (which is a  $U(1)$  gauge theory) to non-abelian symmetries (*i.e.* symmetry algebras whose generators do not commute with each other). In this context, the non-abelian symmetries always form a (semi-)simple Lie-algebra (the most familiar example is  $SU(2)$ ). The number of gauge bosons is always equal to the number of generators of the Lie-algebra. One can write down a gauge-invariant action for the vector bosons, and one can couple them to other particles. For example, the coupling to fermions  $\psi^I$  is as

$$- \sum_{I,J} \bar{\psi}^I \gamma^\mu (\partial_\mu - ig A_\mu^a T_{IJ}^a) \psi^J \quad (8.5)$$

The sum is over all components of the fermion field. The matrices  $T^a$  form a *representation* of the Lie-algebra, which means that they satisfy the commutation relations. The number of components of  $\psi^I$  is called the *dimension* of the representation.

In the present case, the gauge symmetry includes the global  $SO(2N)$  symmetry that rotates the world-sheet fermions into each other. In the case of the  $SO(32)$  heterotic string this is indeed precisely the full gauge symmetry. However, in the case of the  $E_8 \times E_8$  heterotic string one would expect only an  $SO(16) \times SO(16)$  symmetry on the basis of this argument. Here something special happens. From the Ramond sector we get (two times) 128 extra massless vector bosons. As a result, the gauge symmetry gets extended as well. The extension is described by a Lie algebra called  $E_8$ , which has 248 generators and contains  $SO(16)$  as a sub-algebra. The gaugino's are in the same representation as the gauge bosons, the adjoint representation.

The spinor  $S$  of  $SO(8)$  is chiral because of the GSO-projection in the right sector. It follows that the gaugino's are chiral as well (as is in fact the gravitino).

## 8.4 Compactification

When it was discovered in 1984, the heterotic string was almost immediately seen as an excellent theory for particle physics. The main reasons are that it has a sufficiently large gauge symmetry to contain the standard model, and that it has massless chiral fermions. For some people even the presence of supersymmetry in the spectrum is an advantage, although it has not (yet) been seen in nature. The reason is that already at that time supersymmetry was considered as an attractive mechanism to solve the so called *hierarchy problem*. Simply stated, this is the question how to explain the large discrepancy between the weak interaction scale ( $\approx 100$  GeV) and the natural scale of quantum gravity, the Planck scale ( $\approx 10^{19}$  GeV) (in fact, supersymmetry does not explain that discrepancy, but explains why it is stable against quantum fluctuations). Obviously, since supersymmetry is not observed it can at best be an approximate or “broken” symmetry, but from the Planck scale point of view it is to first approximation exact, and the breaking is only a higher order correction, which is usually ignored in string phenomenology.

The most striking disagreement of the heterotic string description with our world is the number of space-time dimensions. But we have already seen that the number can be reduced to 4 by compactifying 6 dimensions. Just as the bosonic string, the heterotic string can be compactified on a torus, but it turns out that this is not a good idea. The reason is that in torus compactifications all higher-dimensional fields are decomposed into lower-dimensional ones according to simple group-theoretical rules:

$$SO(D-1, 1) \rightarrow SO(D-N-1, 1) \times SO(N) \quad (8.6)$$

(in our case  $D = 10$  and  $N = 6$ ) In such a decomposition, a vector  $V$  decomposes as  $(V, 1) + (1, V)$ . In matrix notation this is simply a matter of decomposing the matrix into blocks

$$\begin{pmatrix} SO(3, 1) \text{ matrix} & 0 \\ 0 & SO(6) \text{ matrix} \end{pmatrix} \quad (8.7)$$

An observer in four dimensions only sees the first four components, and views the internal components as a multiplicity. Hence he/she sees a vector and  $N$  scalars.

Now we have to determine how this works for spinors. To simplify the notation we do this in Euclidean space, for a decomposition

$$SO(D) \rightarrow SO(D-N) \times SO(N) . \quad (8.8)$$

The only difference with the Minkowski case are a few irrelevant factors of  $i$  in some expressions. We assume that both  $D$  and  $N$  are even. In that case there are two kinds of fundamental spinor representations. They are distinguished by the eigenvalue of the matrix  $\gamma^*$ , the product of all Dirac  $\gamma$  matrices (usually called  $\gamma_5$  in four dimensions).

Infinitesimal rotations of these spinors are generated by the matrices

$$\Sigma^{\mu\nu} = -\frac{1}{4}i [\gamma^\mu, \gamma^\nu] , \quad (8.9)$$

where  $\gamma^\mu$  are the Dirac matrices. In  $N$  even dimensions, these are  $2^{N/2} \times 2^{N/2}$  matrices. Since we have three different spaces, we have to introduce three different set of Dirac matrices:

$$\begin{aligned} SO(D) : & \Gamma^M \\ SO(D-N) : & \gamma^\mu \\ SO(N) : & \gamma^I \end{aligned} \quad (8.10)$$

All these matrices can be written down explicitly in many different ways, but it can be shown that all choices are equivalent.

Since the dimensions of these spinor spaces are  $2^{D/2}$ ,  $2^{(D-N)/2}$  and  $2^{N/2}$  respectively, it seems natural to try and map the spinor index of  $\Gamma^M$ ,  $A$ , to a product of those of  $\gamma^\mu$  and  $\gamma^I$ ,  $\alpha$  and  $a$  respectively. So we map the labels  $A$  to a pair of labels  $\alpha a$ . We introduce a *Kronecker product* of two matrices  $A_{\alpha\beta}$  and  $B_{ab}$  as follows

$$C_{AB} \equiv C_{\alpha a, \beta b} \equiv A_{\alpha\beta} B_{ab}$$

and we denote  $C$  as follows

$$C = A \otimes B$$

Some useful property of such a product are

$$\begin{aligned} (A \otimes B)(C \otimes D) &= AC \otimes BD \\ \{(A \otimes B), (C \otimes D)\} &= \frac{1}{2} \{A, C\} \otimes \{B, D\} + \frac{1}{2} [A, C] \otimes [B, D] \\ [(A \otimes B), (C \otimes D)] &= \frac{1}{2} [A, C] \otimes \{B, D\} + \frac{1}{2} \{A, C\} \otimes [B, D] \end{aligned}$$

Suppose we have a valid set of Dirac matrices in  $SO(D-N)$  and  $SO(N)$ . Then it is easy to check that the following is a valid set in  $SO(D)$ :

$$\Gamma^\mu = \gamma^\mu \otimes \mathbf{1} \quad (8.11)$$

$$\Gamma^I = \gamma^* \otimes \gamma^I \quad (8.12)$$

where we have split the index  $M$  into  $(\mu, I)$ . It is easy to see that this  $\Gamma^M$  indeed satisfies the Clifford algebra relation  $\{\Gamma^M, \gamma^N\} = 2\delta^{MN}$ , so it is indeed a valid choice.

Now we compute  $\Gamma^*$  as the product of all Dirac matrices, and we get

$$\Gamma^* = (-1)^{N/2} \gamma_{D-N}^* \gamma_N^* \quad (8.13)$$

The two chirality eigenspaces of  $\Gamma^*$  are projected out by the operators

$$P_D^\pm = \frac{1}{2}(\mathbf{1} \pm \Gamma^*) \quad (8.14)$$

Writing this in terms of the subgroup we find

$$P_D^+ = P_{D-N}^+ \otimes P_N^+ + P_{D-N}^- \otimes P_N^- \quad \text{for } N = 0 \pmod{4} \quad (8.15)$$

and

$$P_D^+ = P_{D-N}^+ \otimes P_N^- + P_{D-N}^- \otimes P_N^+ \quad \text{for } N = 2 \pmod{4} \quad (8.16)$$

Denoting the spinor representations as  $S$  (chirality  $+$ ) and  $C$  (chirality  $-$ ) we may write this in the case of interest as

$$S_{10} \rightarrow (S_4, C_6) + (C_4, S_6) \quad (8.17)$$

An observer in four dimensions sees a 10-dimensional chiral spinor as 4 chiral spinors  $S_4$  and 4 anti-chiral ones  $C_4$ . The 4 components are the dimensions of  $S_6$  and  $C_6$ ; an observer in the four uncompactified dimensions cannot rotate the internal components of the spinor, but observes the number of components.

However, we are only interested in massless spinors, so we have to work out the mass that is observed. This is determined by solving the Dirac equation. Its internal space eigenvalues are observed as a mass (of course we start with a massless spinor in 10 dimensions). A four-dimensional massless fermion is a ten-dimensional massless spinor that does not get a mass contribution from the internal dimensions:

$$[(\gamma^\mu \otimes \mathbf{1})\partial_\mu + (\gamma^* \otimes \gamma^I)\partial_I] \psi = 0 \quad (8.18)$$

If the second term vanishes, this is observed as zero mass in four dimensions.\* The Dirac equation in internal space is just  $\gamma^I \partial_I \psi = 0$ , and it has  $2^{N/2}$  solutions (of both chiralities):  $\psi_\alpha^a = \text{constant} \times \delta_{a\alpha}$ ,  $a = 1 \dots 2^{N/2}$ , where  $\alpha$  is a spinor index and  $a$  labels the different solutions. These are all the independent constant vectors in spinor space. The chiral and anti-chiral spinors are in the same representation of any gauge group, and hence we end up with a non-chiral theory: we cannot get the standard model.

The number of solutions in internal space gives the number of observed massless particles in four dimensions. These particles are distinguished from each other by having a different wave function in the compactified dimensions.

The foregoing problem is fortunately special to torus compactifications. If one compactifies on other manifolds, a chiral spinor decomposes in a certain number  $N_L$  of chiral spinors and another number  $N_R$  of anti-chiral ones.

On other spaces one has to generalize (8.18) to take into account curvature. This implies that the ordinary derivatives  $\partial^I$  are replaced by covariant derivatives which contain the spin connection of the manifold (this was trivial for the torus, since a torus is flat). An

---

\* Note that in general it is a bit misleading to speak of eigenvalues, because the Dirac matrix acts off-diagonally; however, if the eigenvalue is zero, this does not matter.

additional complication is that sometimes one allows some of the fields in the space-time theory to have classical values (an example of that was considered in the discussion of torus compactification where we allowed a non-vanishing – but constant – background field  $B_{IJ}$ ). Often one chooses a non-vanishing classical gauge field, and in that case the covariant derivative must also contain a gauge potential term. All of these issues are beyond the scope of these lectures. However, it is clear that in general a different, more complicated equation must be solved in the internal dimensions. In general  $N_L \neq N_R$  so that a chiral theory may be obtained after compactification.

## 8.5 Calabi-Yau compactification\*

One cannot simply compactify space-time on any manifold. String theory imposes constraints on the space-times it can propagate in. These constraints follow from conformal invariance. A very popular class of solutions is a class of six-dimensional manifolds called *Calabi-Yau* spaces. They satisfy the additional requirement that in compactification one supersymmetry is preserved. Note that the heterotic string in 10 dimensions has  $N = 1$  supersymmetry: there is one gravitino. The rules for decomposing a gravitino in torus compactifications are the same as for “ordinary” fermions, as discussed above. Hence in four dimensions the torus-compactified heterotic string has four chiral and four anti-chiral gravitino’s. In four dimensions the anti-chiral ones are just the anti-particles of the chiral ones, so effectively there are four gravitino’s and one calls this  $N = 4$  supersymmetry. It turns out that chiral theories in four dimensions are possible only for  $N = 0$  and  $N = 1$  supersymmetry.

The numbers  $N_L$  and  $N_R$  in Calabi-Yau compactification can be computed. They are topological, which means that they do not change under continuous deformations of the manifold. The difference is equal to half the Euler number of the manifold.

## 8.6 The standard model in string theory\*

We must first review a few basic properties of the standard model. It contains particles of spin 1/2, the quarks and leptons, particles of spin 1, the carriers of the electromagnetic, strong and weak forces, and most likely a particle of spin 0, the Higgs boson. The spin 1 bosons belong to a gauge group  $SU(3) \times SU(2) \times U(1)$ , which below about 100 GeV is broken down to  $SU(3) \times U(1)$ , QCD and electrodynamics. If we ignore the weak interactions the standard model is non-chiral: QED and QCD couple in the same way to both chiralities. However, the weak interactions do distinguish chirality, and couple only to the left-handed and not to the right-handed electron.

A remarkable feature of the standard model is that the particles are grouped in families: it is natural to associate  $(e, \nu_e, u, d)$ ,  $(\mu, \nu_\mu, c, s)$  and  $(\tau, \nu_\tau, t, b)$  with each other. Although this grouping is rather arbitrary (it is determined only by the mass hierarchy, which is not understood, and is furthermore disturbed by mixing), it is certainly true that the basic  $SU(3) \times SU(2) \times U(1)$  quantum number repeat three times.

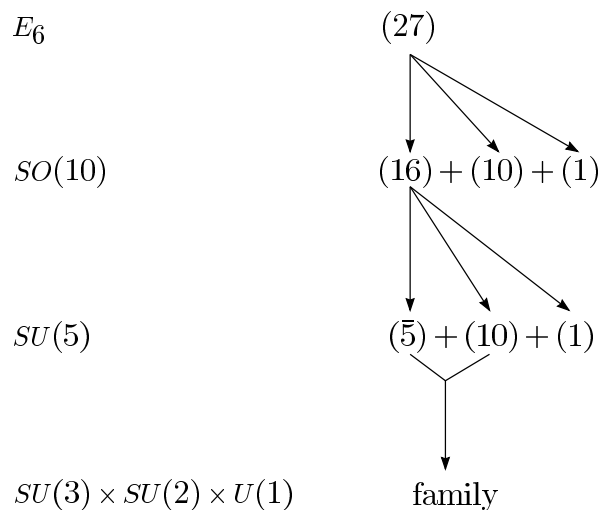
When writing one family of the standard model, it is sometimes useful to replace

certain particles by their anti-particles. In general, anti-particles have opposite charge and complex conjugate non-abelian group representations. Furthermore in four dimensions (and in all  $4m$  dimensions) particles and anti-particles have opposite chirality (in  $4m + 2$  dimensions they have the same chirality). For massless particles chirality is equivalent to helicity: a particle with chirality  $+$  is left-handed. Hence, for example, the anti-particle of  $e_R^-$  is  $e_L^+$ . This can be used to make all particles in a standard model family left-handed. If one does that, the  $SU(3) \times SU(2) \times U(1)$  representations are (we are ignoring right-handed neutrinos here)

Representation	Family 1	Family 2	Family 3
$(3, 2, \frac{1}{6})$	$u, d$	$c, s$	$t, b$
$(\bar{3}, 1, -\frac{2}{3})$	$\bar{u}$	$\bar{c}$	$\bar{t}$
$(\bar{3}, 1, \frac{1}{3})$	$\bar{d}$	$\bar{s}$	$\bar{b}$
$(1, 2, -\frac{1}{2})$	$e^-, \nu_e$	$\mu^-, \nu_\mu$	$\tau^-, \nu_\tau$
$(1, 2, 1)$	$e^+$	$\mu^+$	$\tau^+$

Here the “bar” denotes the conjugate of a representation, and representations are denoted as (dimension).

The standard model gauge group can be embedded into larger groups. The most familiar embedding chain is



The idea is that just as the low energy gauge group  $SU(3) \times U(1)$  is extended at 100 GeV to the standard model gauge group  $SU(3) \times SU(2) \times U(1)$ , perhaps at still higher energies the standard model gauge group is enlarged further. This can even be done in such a way that there is just one gauge group and one representation per family. From this point of view  $SO(10)$  looks perhaps most attractive, but the next step,  $E_6$ , has also been studied in the literature.

The point is now that there is a class of compactifications of the  $E_8 \times E_8$  heterotic string in which the gauge group is broken to  $E_6 \times E_8$ . Furthermore in these compactifications

the massless fermions are automatically in the representation (27). One can have  $N_L$  left-handed fermions and  $N_R$  right-handed ones, with  $N_L \neq N_R$ . If we represent the right-handed ones by anti-particles, we may also write that as  $N_L(27)_L + N_R(\overline{27})_L$ , *i.e.* with only left-handed particles.

In switching from particles to anti-particles one has to be careful with the notion of chirality of a theory. The rule is simple: if after flipping all chiralities one obtains a different theory, the theory is chiral. However, if we can write one theory in the form of another one by trading some particles for anti-particles, then those two theories are not different. For example

$$\begin{array}{l|l|l|l}
 (27)_L & \text{chirality flip} & (27)_R \Rightarrow (\overline{27})_L & \text{chiral} \\
 (27)_L + (\overline{27})_L & \text{chirality flip} & (27)_R + (\overline{27})_R \Rightarrow (\overline{27})_L + (27)_L & \text{non-chiral} \\
 (10)_L & \text{chirality flip} & (10)_R \Rightarrow (\overline{10})_L = (10)_L & \text{non-chiral}
 \end{array}$$

Here the double arrow indicates a particle - anti-particle interchange. In the last line the representation (10) of  $SO(10)$  appears. This is the vector representation, and it is real:  $(10) = (\overline{10})$ . Hence a four-dimensional fermion in such a representation can never be chiral.

Chiral particles cannot have a mass. The reason is that in a mass term of the form  $m\bar{\psi}\psi$  the chiral projection operator is flipped:

$$\bar{\psi}\frac{1}{2}(1 + \gamma_5)\psi \equiv \psi^\dagger\gamma_4\frac{1}{2}(1 + \gamma_5)\psi = \psi^\dagger\frac{1}{2}(1 - \gamma_5)\gamma_4\psi = \left(\frac{1}{2}(1 - \gamma_5)\psi\right)^\dagger\gamma_4\psi \quad (8.19)$$

Hence if the second  $\psi$  has chirality +, the first must have chirality -. This implies that one cannot write down a mass term for  $(27)_L$  unless there is also a  $(27)_R$ . On the other hand, if both  $(27)_L$  and  $(27)_R$  are present a mass term is allowed.

An often used principle in quantum field theory is “anything that is allowed is obligatory”. In other words, if there is no reason why a certain parameter should be zero, then it is reasonable to assume that it is not zero. On the basis of this principle one usually assumes that if for example  $N_L > N_R$ , then  $N_R$   $(27)_R$ 's combine with  $N_L$   $(27)_L$ 's to acquire a large mass (in this context large means “too large to be observed”), so that there are  $N_L - N_R$  unpaired  $(27)$ 's left over. They cannot get a mass, and hence they must appear as particles in the low-energy spectrum. By the same logic, when  $E_6$  break to  $SO(10)$  the leftover  $(10) + (1)$  are assumed to acquire a mass of order the symmetry breaking scale (which must be much larger than 100 GeV). The leftover (1) in the breaking to  $SU(5)$  can also become massive, so that finally only the standard model particles remain massless.

If we accept all this, then we must be looking for a Calabi-Yau manifold with  $|N_L - N_R| = 3$ . One of the best studied examples has  $N_L = 1$  and  $N_R = 101$ , but thousands of Calabi-Yau manifolds have been found, including some yielding three families. In addition there are many variations on the basic idea. These allow breaking  $E_6$  to something closer to the standard model gauge group, while at the same time lowering the number of



families. Although there are many examples that get rather close to the standard model, there is none that really “fits like a glove”. On the other hand, nothing like a systematic search has been possible so far due to the overwhelming number of possibilities.

Furthermore it is difficult to do detailed calculations. For example, one may wish to compute the quark and lepton masses. They get their masses from a three-point coupling to the Higgs scalar,  $\lambda_x \bar{\psi}_x \phi \psi_x$ , where “ $x$ ” denote a particle, and  $\lambda_x$  is a coupling constant. This is called a Yukawa coupling. The mass of a particle is equal to the vacuum expectation value of the Higgs – a universal parameter with the dimension of mass – times  $\lambda_x$ . The values of  $\lambda_x$  can be computed in string theory, but the problem is that the massless spectrum typically contains a few hundred scalars, so that it is not clear which one is going to play the rôle of the Higgs scalar  $\phi$ . In addition the Higgs potential and the whole symmetry breaking mechanism is poorly understood within string theory. One has tried to remedy this by adding some field-theoretic assumptions, but pretty soon any predictive power is completely lost. It seems that in making contact with reality one still has a long way to go.

However, it is encouraging that the gross features of the standard model are reproduced, and that they come out essentially automatically from a theory that in principle could only hope to be a theory of gravity. These features are

- Chiral fermions
- A sufficiently large gauge group
- The correct family structure
- Replication of families

## 8.7 Other ten-dimensional Heterotic strings\*

It turns out that one can find other solutions to the conditions of modular invariance, by dropping the requirement of separate modular invariance in the left- and right-moving sectors. An example is

$$\frac{1}{4} \frac{[\eta(\tau)\eta(-\bar{\tau})]^8}{(\text{Im } \tau)^4} \left\{ \begin{aligned} & (F_3^4 - F_4^4)_R [(F_3^8 + F_4^8)^2 + (F_2^8 - F_1^8)^2]_L \\ & + (F_3^4 + F_4^4)_R [(F_3^8 - F_4^8)(F_1^8 + F_2^8) + (F_1^8 + F_2^8)(F_3^8 - F_4^8)]_L \\ & + (F_2^4 + F_1^4)_R [(F_3^8 - F_4^8)^2 + (F_2^8 + F_1^8)^2]_L \\ & + (F_2^4 - F_1^4)_R [(F_3^8 + F_4^8)(F_1^8 - F_2^8) + (F_1^8 - F_2^8)(F_3^8 + F_4^8)]_L \end{aligned} \right\} \quad (8.20)$$

Here the right-moving factors (with subscript R) have an implicit argument  $\bar{\tau}$  whereas the left-moving ones (L) have argument  $\tau$ . In the left-moving sector the fermions are grouped into two groups of 16, and *e.g.* in  $(F_3^8 - F_4^8)^2$  the first factor refers to the first group of 16 and the second factor to the second.

From the partition function one reads off that the gauge group is  $SO(16) \times SO(16)$ , and is not extended, and that there are no gravitino's, and hence no supersymmetry. Furthermore there is no tachyon. This shows that it is not necessary to have supersymmetry to get rid of the tachyon.

There are six other non-supersymmetric heterotic string theories in ten dimensions, but the other six all have a tachyon.

## 8.8 Four-dimensional strings\*

The main idea behind four-dimensional strings is to drop the requirement that all right-moving fermions carry a space-time index  $\mu$ . For conformal invariance we still need to have a total of 10 fermions, but we can divide them up into four space-time and six internal fermions. Of course we do the same with the 10 right-moving bosons. In the left sector we have 26 bosons, which were previously divided into 10 space-time and 16 internal. Instead we choose now 4 space-time and 22 internal.

It should be clear from the foregoing discussion that a group of  $2N$  internal fermions gives rise to an  $SO(2N)$  symmetry if and only if all  $2N$  fermions always have the same boundary conditions. The same is true for fermions with a space-time index. The reason the theories constructed so far are ten-dimensional strings is the fact that in light-cone gauge there is an  $SO(8)$  symmetry: all partition functions depend on  $(F_i(-\bar{\tau}))^4$ . If instead we allow combinations like  $F_{i_1}F_{i_2}F_{i_3}F_{i_4}$  with different labels  $i_m$ , then clearly the  $SO(8)$  symmetry is broken, in general to  $SO(2)^4$  (one  $SO(2)$  for each pair of fermions). If we associate the first factor  $F_{i_1}$  with the light-cone components of the fermions  $\psi^\mu$ , then the resulting string theory is four-dimensional.

The difficulty in constructing such a theory is to construct a modular invariant partition function out of the internal degrees of freedom. Concretely, suppose one the right the 10-dimensional space-time index is split into  $(\mu, M)$ ,  $M = 1 \dots, 6$  and on the left the 26-dimensional space-time index is split into  $(\mu, I)$ ,  $I = 1 \dots, 22$ , then the internal degrees of freedom are  $\psi^M$ ,  $X^M$  and  $X^I$ . Now the internal degrees of freedom are partly bosons and partly fermions, and it is difficult to build modular invariant partition functions in such a hybrid system. The solution is *bosonization*: in 2 dimension one (compactified) boson can be replaced by a pair of fermions or vice-versa. Hence we can either replace  $\psi^M$  by 3 bosons and construct a modular invariant partition function out of free bosons (this leads to a description in terms of Lorentzian even self-dual lattices) or replace  $X^M$  and  $X^I$  by 12 and 44 fermions. Then one proceeds as above, by trying to construct modular invariant combinations of fermion partition functions. A technical complication in both cases is how to preserve the two-dimensional supersymmetry that relates  $\psi^M$  and  $X^M$ .

We will not give further details here. Both methods work, and yield huge numbers of solutions. So many solutions that a complete listing is impossible. In any case one can get chiral and non-chiral theories, a large number of possible gauge groups, theories with and without tachyons and with any allowed number of supersymmetries.

It may seem that there are two different ways to get to four dimensions: compactify a ten-dimensional string, or construct a string theory directly in 4 dimensions. It turns out

that these two methods are closely related. In many examples a given four-dimensional string theory yields exactly the same spectrum as a definite Calabi-Yau compactification. It is believed that then not only the spectrum but also all interactions are the same: the two theories are identical.

There are other methods for constructing strings in four dimensions, but they are all (believed to be) connected to each other. The full set of four-dimensional theories is a complicated parameter space that is still to a large extent unexplored.

## 9 Dualities

Also in ten dimensions there are connections among seemingly different string theories. The theories involved in these “dualities” are the ten-dimensional supersymmetric string theories: they include the type-IIA, type-IIB, heterotic  $SO(32)$  and heterotic  $E_8 \times E_8$  theories introduced in the previous chapter. Before discussing dualities, we have to introduce one more family member, namely

### 9.1 Open strings

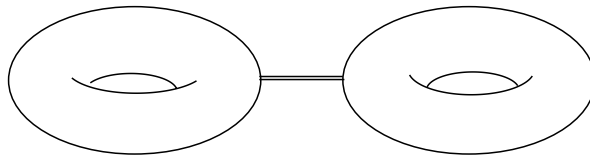
All theories seen so far are oriented closed strings. There exists also a theory of unoriented open and closed strings in ten dimensions that satisfies all known consistency conditions. Just as one of the heterotic strings, this theory has  $N = 1$  supersymmetry and an  $SO(32)$  gauge group. It is called the type-I string. The reason why  $SO(32)$  appears is rather different than in the case of heterotic strings. In that case the reason was modular invariance.

### 9.2 Finiteness

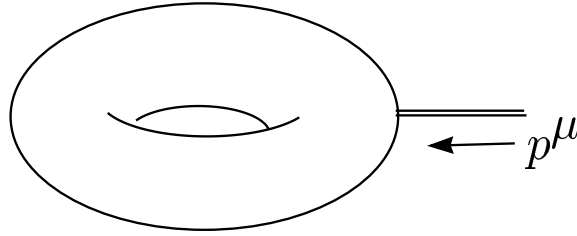
As we have seen, modular invariance for closed strings allowed us to restrict the one-loop integration to one “modular domain”, depicted in section (5.9). Furthermore we have seen that this domain excludes the dangerous  $\text{Im } \tau = 0$  limit where the field theory divergences are. Then, if there is no tachyon, the one loop integral is finite. Indeed, the two supersymmetric heterotic strings have finite one-loop integral, and so does the non-supersymmetric  $SO(16) \times SO(16)$  string.

Furthermore the one-loop integral (without external legs) is not just finite, but zero for superstrings, because the partition function (the integrand) vanishes due to the “abstruse identity”. This is important for higher loop divergences.

Consider for example the two-loop diagram, the double torus. The modular integration for the double torus contains a limit where the double torus looks like this



One may show that in the limit where the size of the tube connecting the tori goes to zero and the length of that tube goes to infinity, only one particle contributes: the dilaton. In that limit the amplitude factorizes into two so-called *tadpole* diagrams.

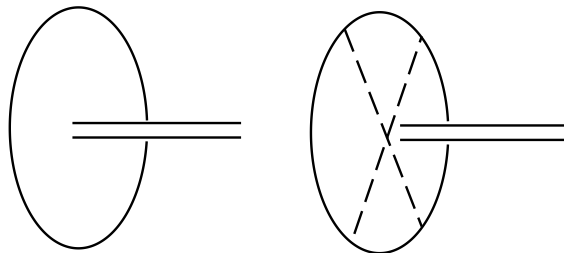


A tadpole is a diagram with a single external line. Four-momentum conservation requires that the four-momentum into that line vanishes:  $p^\mu = 0$ . Then  $p^2 = -M^2 = 0$ , so this must be a massless particle. Lorentz-invariance requires that it must be a scalar. Then it can only be the dilaton.

It can be shown that if the one-loop diagram without external lines give a contribution  $\Lambda$ , then the dilaton tadpole is proportional to  $\Lambda$ . The the limit of the double torus diagram, treated as a field theory Feynman diagram with one propagator ending on both sides on a one-point vertex, yields something like  $\Lambda[1/p^2]\Lambda$ . However, momentum conservation tells us that  $p^2 = 0$ . This implies that the double torus contribution becomes infinite, unless  $\Lambda = 0$ . In that case this simple argument is inconclusive.

Although the discussion here was not very precise, the argument is indeed correct. The  $SO(16) \times SO(16)$  heterotic string has a finite, non-zero  $\Lambda$ . Therefore it is finite at one loop, but divergent at two loops. The supersymmetric heterotic strings have  $\Lambda = 0$  at one loop. It turns out that the full two-loop integral again yields zero, which does not follow from the foregoing discussion, but is consistent with it. In fact the corresponding integral vanishes at any loop order. Diagrams with external lines do not vanish, but give a finite answer at any loop order (or at least that is what is generally believed, but only partly proved).

In the case of open and unoriented strings the same arguments apply to the closed sub-sector, *i.e.* the torus and higher order generalizations, but the open and non-oriented diagrams are a separate story. At the lowest non-trivial order, Euler number 1, there two surfaces, characterized by  $b = 1$  and another by  $c = 1$  (see eq. (6.12)). To both surfaces one can attach an external dilaton, so that one obtains two dilaton tadpole diagrams (disk and crosscap)



The point is now that the contributions of these two dilaton tadpoles must cancel. If not, then they lead to a divergence at Euler number zero due to a dilaton propagator in an

intermediate line. At Euler number zero the relevant diagrams are annulus, Klein bottle and Möbius strip. It turns out that the integrations corresponding to these diagrams have limits that look like a dilaton propagation between two disks (annulus), between two crosscaps (Klein bottle) and between a disk and a crosscap (Möbius strip). Summing these diagrams we get, schematically

$$D\left[\frac{1}{p^2}\right]D + C\left[\frac{1}{p^2}\right]C + C\left[\frac{1}{p^2}\right]D + D\left[\frac{1}{p^2}\right]C = (D + C)\left[\frac{1}{p^2}\right](D + C) . \quad (9.1)$$

The dangerous  $1/p^2$  divergence cancels if  $D + C = 0$ .

Direct computation shows that in suitable units  $C = -32$ . It turns out that each boundary can appear with a certain multiplicity factor  $N$ , the Chan-Paton multiplicity that was introduced before. Then, in the same units, the disk contribution  $D$  is equal to  $N$ , and the tadpoles cancel if  $N = 32$ . In that case one finds precisely the gauge bosons of  $SO(32)$  in the spectrum.

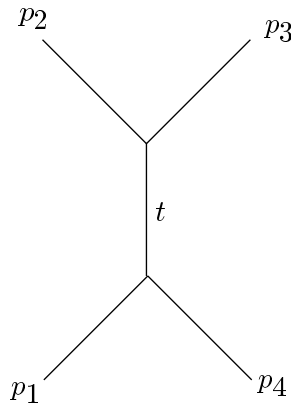
It should be emphasized that this argument is not as nice as the one used for heterotic strings. There finiteness was a consequence of a sound principle, modular invariance, whereas here it has to be imposed “by hand”.

We will not discuss the  $SO(32)$  open string theory in more detail here, except to remark that the massless spectrum is identical to that of the heterotic  $SO(32)$  string. In particular both have  $N = 1$  supersymmetry. However, the massive spectra are different. For example, the heterotic string has massive states in the  $SO(32)$  spinor representation, whereas the open string theory only has symmetric and anti-symmetric tensors in its spectrum.

### 9.3 Dual models

The term “duality” is very frequently used in string theory to describe a variety of phenomena. The first use of the term goes back to 1968, and is related to the historical origin of string theory.

This notion of duality occurs in four point functions at tree level. For example, consider two particles that scatter by the exchange of some other particle, as in the following diagram:



Two-to-two particle scattering is kinematically described in terms of the Mandelstam variables  $s$ ,  $t$  and  $u$ :

$$s = -(p_1 + p_2)^2, \quad t = -(p_2 + p_3)^2, \quad u = -(p_1 + p_3)^2, \quad (9.2)$$

where all momenta are assumed to be incoming (in other words, the final state particles have outgoing momenta  $-p_3$  and  $-p_4$ ). If the exchanged particle has spin 0, this particular diagram has an amplitude of the form

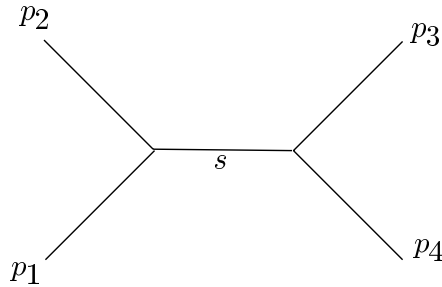
$$A = -g^2 \frac{1}{t - M^2} \quad (9.3)$$

where  $g$  is a coupling constant and  $M$  the mass of the exchanged particle.

String theory was born by considering the scattering of hadrons by exchanging mesons. If one considers the exchange of several mesons of different spins  $J$ , the amplitude generalizes to

$$A_t(s, t) = - \sum_J g_J^2 \frac{s^J}{t - M_J^2} \quad (9.4)$$

One may also consider annihilation diagrams of the form



This yields an amplitude of the form

$$A_s(s, t) = - \sum_J g_J^2 \frac{t^J}{s - M_J^2} \quad (9.5)$$

It was observed empirically that in hadronic scattering  $A_s(s, t) \approx A_t(s, t)$ . This led Veneziano in to a formula that explicitly had this “ $s$ - $t$  duality” built in:

$$A(s, t) = \frac{\Gamma(-\alpha(s))\Gamma(-\alpha(t))}{\Gamma(-\alpha(s) - \alpha(t))} \quad (9.6)$$

with

$$\alpha(s) = \alpha(0) + \alpha' s \quad (9.7)$$

The  $\Gamma$  function satisfies

$$\Gamma(u + 1) = u\Gamma(u), \quad \Gamma(1) = 1 \quad (9.8)$$

so that  $\Gamma(u) = (u + 1)!$  for positive integer arguments. For negative arguments there are singularities at  $u = 0, -1, -2, \dots$ , and near  $u = -n$  the behavior is

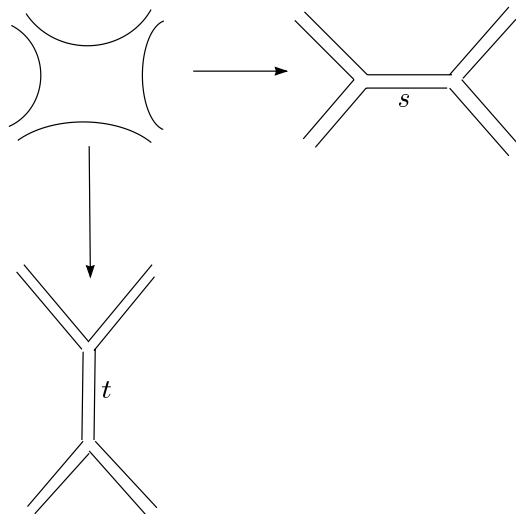
$$\Gamma(u) \approx \frac{(-1)^n}{n!(u + n)} \quad (9.9)$$

We see then that the Veneziano amplitude has poles if  $\alpha(s) = n$  or  $\alpha(t) = n$ , or

$$s = \frac{1}{\alpha'}(n - \alpha(0)) , \quad t = \frac{1}{\alpha'}(n - \alpha(0)) \quad (9.10)$$

If one interprets these poles as particles exchanged in the  $t$ -channel ( $t = M^2$ ) or the  $s$ -channel ( $s = M^2$ ), then we recognize here precisely the open string spectrum (3.63) (apart from multiplicities and spin-dependence, which with a more careful analysis also come out correctly). The parameter  $\alpha(0)$  is called the *intercept* and plays the role of the vacuum energy subtraction  $a$ . The parameter  $\alpha'$  is called the *Regge slope* (this explains finally why the fundamental length scale of string theory was called  $\alpha'$ , and in particular why there is a prime).

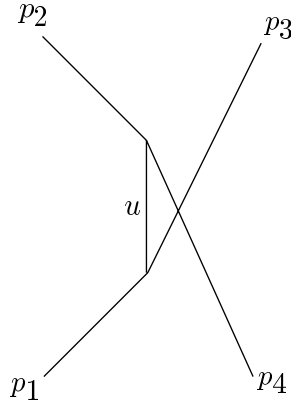
The foregoing remarks suggest that the Veneziano amplitude has a string theory interpretation, and indeed it does. Pictorially the observation in the previous paragraph is



Although it is beyond the scope of these lectures, the four-point string diagram is not hard to compute, and yields indeed the Veneziano amplitude. In special limits (namely near the poles in  $s$  or near the poles in  $t$ ) this single amplitude behaves at the same time as sum of an infinite number of  $f$ -channel exchanges, or as the sum of an infinite number of  $s$ -channel annihilations. It is historically rather strange that first the amplitude was written down, and only afterwards its origin was understood, but that is indeed what happened.



Note that by deforming the open string tree diagram one can get the  $s$  and  $t$  diagrams but not the  $u$  diagram



since we cannot interchange external lines. However, if we work with closed instead of open strings we can interchange all lines, since they are simple four tubes attached to a sphere. The tubes can be moved around each other on the sphere. Consequently the four-point amplitude for closed strings has a symmetry in  $s$ ,  $t$  and  $u$ , and has the form

$$A(s, t, u) = \frac{\Gamma(-\frac{1}{4}\alpha(s))\Gamma(-\frac{1}{4}\alpha(t))\Gamma(-\frac{1}{4}\alpha(u))}{\Gamma(-\frac{1}{4}\alpha(s) - \frac{1}{4}\alpha(t) - \frac{1}{4}\alpha(u))} \quad (9.11)$$

The factors  $\frac{1}{4}$  reflect the factor of 4 difference in spacing between the open and closed string spectra.

## 9.4 Historical remarks

After the discovery of the Veneziano amplitude and the realization that it had a string interpretation there was a short but intense period of trying to develop it as a strong interaction model. These models were called “dual models”, because of the duality property described above. These models ran into serious difficulties when the tachyon was found and the critical dimension turned out to be 26.

Around 1972 QCD was formulated as a theory of the strong interactions, and it quickly gained in popularity over dual models. Meanwhile fermionic strings were constructed, and in the early 70’s, completely independently of string theory, supersymmetry and supergravity were discovered. In 1976 Gliozzi, Scherk and Olive formulated their projection on the fermionic string spectrum, which connected these developments for the first time: superstrings were born, and the supergravity theories appeared as their low-energy limits.

An important shift of perspective took place when it was realized that the infinite tower of string states contained a massless particle of spin 2 (a symmetric tensor in 10 or 26 dimensions). It was in particular Scherk who pushed the idea the string theory should be considered as a theory of gravity rather than the strong interactions. But after 1975 the interest in string theory diminished rapidly, and all attention turned to the development of the standard model, which was going through an exciting period.

In 1984 the “second string revolution” began, after an almost ten-year period of near inactivity. The Heterotic strings were constructed, and the  $E_8 \times E_8$  theory looked especially promising for phenomenology. In the years that followed there was a huge outburst of activity. Calabi-Yau compactifications were studied, and many other compactifications and four-dimensional string constructions were investigated.

After another period of relatively little activity (although much more than between 1975-1984) in 1994 the “third string revolution” occurred. This was sparked by the discovery of a set of string dualities, although a rather different kind of dualities than those discussed above.

## 9.5 T-duality\*

In the section on bosonic string torus compactification we saw that the compactified string has a remarkable duality under  $R \rightarrow \alpha'/R$ . Similar dualities can be studied for the compactified type-II and heterotic strings. If we compactify on a circle to 9 dimensions, it turns out that the two compactified type-II strings are mapped into each other, and that the two heterotic strings are also mapped to each other. We will not study this in detail, but only give a rough idea how this works.

The two type-II strings only differ in the relative chirality of the left and right GSO partition functions. However, chirality is lost if one make a torus compactification, as we have seen in the previous chapter. Therefore it is not a big surprise that the difference between IIA and IIB is lost upon compactification.

For the Heterotic compactification we can be more explicit. A general compactification of one dimension can be described by a  $\Gamma_{1,1}$  Lorentzian even self-dual lattice. If we combine this with the 16-dimensional lattice in the left sector (using the bosonic description of that sector) we get a Lorentzian even self-dual lattice  $\Gamma_{17,1}$ . A general theorem on such lattices says that they are all related by Lorentz transformations. It can be shown that in string theory these Lorentz-transformations are parametrized by the moduli of the compactification torus, which are massless scalar fields. Giving these scalars a vacuum expectation value we move from one  $\Gamma_{17,1}$  lattice to another, and in particular we can move from  $E_8 \times E_8 \times \Gamma_{1,1}$  to  $D_{16} \times \Gamma_{1,1}$ . This shows that these two theories are “on the same moduli space”, and hence they are clearly related. That the relation is actually a T-duality requires a bit more discussion.

## 9.6 Perturbative and non-perturbative physics\*

String theory defines a perturbation series. Any amplitude comes out as an infinite sum of the form  $\sum_n a_n g^n$ , where  $g$  is the string coupling, related to the exponential of the vacuum expectation value of the dilaton.

It is not obvious that this is all there is. Suppose the exact answer is of the form

$$F(g) = \sum_n a_n g^n + A e^{-\frac{1}{g^2}} \tag{9.12}$$

The second term has the property that all its derivatives at  $g = 0$  vanish. Hence its Taylor expansion around  $g = 0$  is trivial. One can never see such a term in perturbation theory.

Therefore, no matter how far we expand, we will never be able to compute the coefficient  $A$  from a loop diagram. Effects that behave like this are called non-perturbative effects. It is known that such contributions occur, for example, in QCD.

In QCD the appearance of such effects is related to the existence of classical solutions to the equations of motion (without sources). The first terms in (9.12) are coming from an expansion of the action around the trivial classical solution (with all fields equal to zero), whereas the second term comes from an expansion around the non-trivial solution (there are additional perturbative corrections to the last term). In string theory it is *a priori* much more difficult to find such classical solutions. This is because in string theory we *only* have a perturbative expansion.

To understand this note that in QCD the Feynman diagram are derived from an action, and that action can also be used to derive classical equations of motion that can be solve. In string theory we only have the “Feynman diagrams”, but no action to derive them from (we would need an action in target space, not on the world-sheet).

One way to get around this is to derive from string theory a low-energy effective action. This is an action for the massless fields only, such that the interactions derived from it are the same as those derived directly from string theory. This action includes of course the Einstein action for gravity, as well as the usual actions of the gauge bosons. Then one can solve the equations of motion of this low-energy effective action, and postulate that it is the first order approximation to a genuine string theory classical solution.

Classical solutions to field theory equations of motion can be classified according to their extension in space. A solution can be point-like, meaning that its energy density is concentrated around a point in space; analogously it can be string-like or membrane-like. In general we call a solution a *p-brane* if it has  $p$  spatial dimensions. Hence a particle-like solution is a 0-brane, a string-like solution a 1-brane, etc.

## 9.7 D-branes\*

There exists an exception to the statement that we cannot describe classical solutions exactly in string theory. These are the so-called D-branes (where “D” stands for Dirichlet).

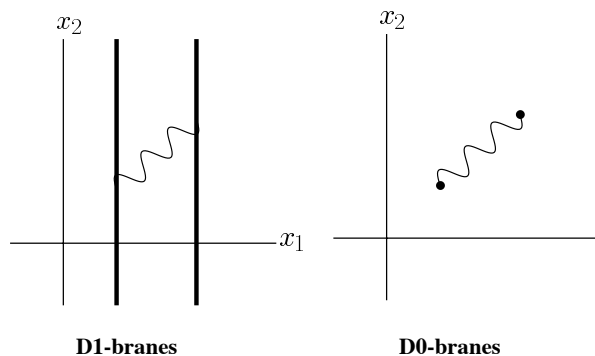
As we have seen, open strings can have two kinds of boundary condition. A Neumann boundary condition has the property that  $\partial_\sigma X$  vanishes at the boundary. This amounts to conservation of momentum at the boundary. For a Dirichlet boundary condition  $\partial_\tau X = 0$  at the boundary. This means that  $X(0)$  and  $X(\pi)$  are fixed as a function of time.

One may consider situations where – for the bosonic string –  $D$  space-like coordinates have Dirichlet boundary conditions, and the remaining  $26 - D$  have Neumann boundary conditions. Then we have

$$X^i(0) = v^i, \quad X^i(\pi) = w^i, \quad i = 1, \dots, D \quad (9.13)$$

where  $v^i$  and  $w^i$  are some fixed vectors. The remaining space-like directions are not fixed. Below we illustrate this for a total of 3 dimensions instead of 26. We plot the two

space-like dimensions  $x^1$  and  $x^2$



In the first picture we have Dirichlet boundary conditions in the  $x^1$  direction. Hence the open string has its endpoints fixed to two lines at fixed  $x^1$ . The endpoints are allowed to move in the  $x^2$  direction. In the second picture we have Dirichlet boundary conditions in both directions. Then the endpoints are completely fixed to points in space. Finally we may take Neumann conditions in all directions. Then the endpoints are free to move in all space directions. This is what we normally mean by “open strings”.

In the first two cases translation invariance is broken in one or two directions. This can be understood in terms of objects that are present in space. A particle fixed at some position obviously breaks translation invariance in all directions. A wall breaks translation invariance in the direction orthogonal to it. Therefore we interpret the first picture in terms of two lines, or 1-branes, whose presence breaks translation invariance, and the second picture is interpreted in terms of two points or 0-branes. The space-time momentum of modes on the string is not conserved because it can be transferred to these objects. Because their presence is due to Dirichlet boundary conditions these objects are called D-branes or  $D_p$ -branes: they are  $p$ -branes defined by the property that open strings can end on them.

This only makes sense if we view the D-branes as dynamical objects, that can themselves move around and vibrate. This point of view is indeed correct. It turns out that these D-branes are the exact string analog of certain approximate field theory solutions. There may be additional solutions, but those are harder to describe.

Note that the D-branes are objects that are not obtained in string perturbation theory. There are precise rules for which kind of D-branes can occur in which string theories. There are no D-branes in heterotic strings, but they do exist in type-II strings. This may look surprising at first, since type-II strings were constructed as closed string theories. However, that was a perturbative description. It turns out that these theories nevertheless contain non-perturbative states that are described by open strings.

The rules are that type-IIA theories can have  $D_{\text{even}}$ -branes whereas type-IIB theories can have  $D_{\text{odd}}$ -branes. Of special interest are the  $D_0$ -branes in type-IIA and the  $D_9$  branes in type-IIB. Let us start with the latter. A  $D_9$  brane fills all of space, and the open strings ending on it have endpoints that can move freely through space. Such strings are traditional open strings with only Neumann conditions, and no Dirichlet conditions.

The type-IIB string either has no  $D_9$  branes at all, or it must have exactly 32. In the latter case one is forced to allow also different world-sheet orientations, and one is led to the type-I string. One can assign a label to each of the 32 branes. This corresponds to the Chan-Paton labels. From this point of view the type-I string lives in a space built out of 32  $D_9$  branes stacked on top of each other.

## 9.8 S-duality\*

The  $D_0$ -branes of type-IIA correspond to extra, non-perturbative particle-like states in the spectrum. These states turn out to carry a charge with respect to the massless vector boson  $A_\mu$  in the perturbative spectrum. They are the only states that are charged, because it is easy to show that all perturbative states are neutral.

One may also compute the masses of the  $D_0$  particles, and then one finds that they are  $\propto \frac{|q|}{g}$ , where  $q$  is the charge and  $g$  the string coupling constant.

Witten made the observation that these particles could be interpreted as Kaluza-Klein modes of a circle compactification of 11-dimensional supergravity.

The latter theory has two bosonic fields: a graviton, and a rank-3 anti-symmetric tensor. Upon compactification (as we have seen in section (3.6)) the graviton yields a 10-dimensional graviton, a photon and a scalar. The rank-3 tensor yields a rank-3 tensor and a rank-2 anti-symmetric tensor. These are precisely the bosonic fields of the ten-dimensional type-II string at the massless level.

But Kaluza-Klein compactification yields more than that. If we expand any field in modes, we get also massive excitations. These excitations turn out to carry charge with respect to the photon, and turn out to match precisely the  $D_0$ -particle spectrum. Furthermore the masses of these excitations are inversely proportional to the radius of the circle: the smaller the circle, the more energy it costs to excite a mode in the compact direction.

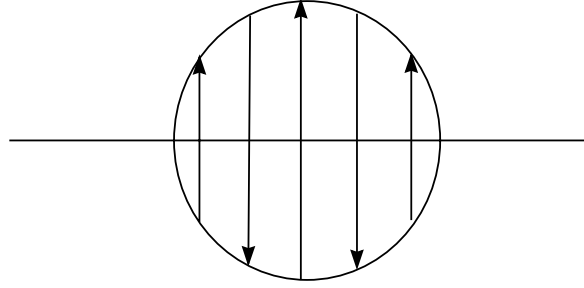
This suggests to identify the radius of the circle with the strength of the type-IIA coupling constant. At small coupling constant the masses of the Kaluza-Klein modes or  $D_0$ -particles become large and can be neglected; this correspond to the  $D = 11$  theory compactified on a very small circle. In the limit of large coupling the  $D_0$  particles become massless, and the radius goes to infinity. In this limit we approach 11 flat dimensions.

The conclusion – which is supported by other arguments – is that the strong coupling limit of the type-IIA theory is an 11-dimensional theory. This is definitely not a string theory. It is also definitely not 11-dimensional supergravity: that theory has infinities in its perturbation theory, which should be absent if it is the strong coupling limit of a finite string theory. The speculation is now that it is an unknown theory, called “M-theory” whose low-energy limit is  $D = 11$  supergravity.

One can also consider the strong coupling limit of all other string theories and one finds

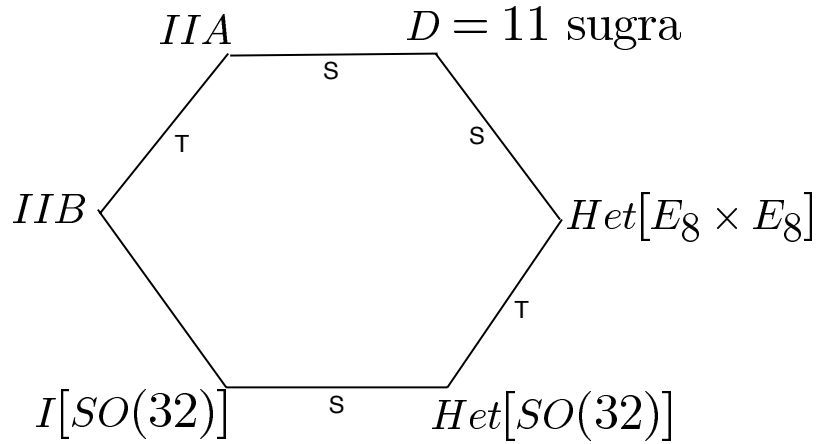
$$\begin{aligned}
 \text{Heterotic } SO(32) &\leftrightarrow \text{Type-I } SO(32) \\
 \text{Type-IIB} &\leftrightarrow \text{Type-IIB} \\
 \text{Heterotic } E_8 \times E_8 &\leftrightarrow \text{11-dimensional supergravity on } S_1/Z_2
 \end{aligned}
 \tag{9.14}$$

The first duality requires the existence of  $SO(32)$  spinors in the spectrum of the type-I string. They can only because of non-perturbative effects. The second case is, for obvious reasons, called self-duality. The strong coupling limit of the  $E_8 \times E_8$  heterotic string is  $D = 11$  supergravity compactified on a circle with mirror image points identified to each other, as shown here:



The resulting space is a line segment, so that the 11-dimensional space becomes two planes separated by the length of the segment. The two  $E_8$  factors can be associated with the two planes.

There is a lot more to say about dualities in 10 and 11 dimensions, and even more in less than 10 dimensions. But we will stop here and conclude with the duality picture that summarizes the present understanding.



## A Functional methods

In classical field theory we deal with dynamical systems with an infinite number of “coordinates”  $q_i$ . For example, suppose we want to describe the classical motion of gauge fields  $A_\mu(\vec{x}, t)$  in electrodynamics. Then the coordinates that enter into the familiar Lagrangian or Hamiltonian description are  $A_\mu(\vec{x})$ .

An additional complication is that these coordinates are continuously infinite: the label ‘ $i$ ’ of  $q_i$  gets replaced by the set  $(\mu, \vec{x})$ . This implies that summations become integrations, and differentiations are replaced by *functional derivatives*.

Let us first review the kind of formulas we want to generalize to classical field theory. A system with a finite number of degrees of freedom is usually described by a Lagrangian  $L(q_i, \dot{q}_i)$ . To every path with initial time  $t_0$  and final time  $t_1$  one associates an action

$$S = \int_{t_0}^{t_1} dt L(q(t), \dot{q}(t)) \quad (\text{A.1})$$

Hamilton’s principle states that the classical trajectory followed by the system is given by an extremum of the action:  $\delta S[q_i(t)] = 0$ . From this one derives the equations of motion, the Euler-Lagrange equations,

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \left( \frac{\partial L}{\partial q_i} \right) = 0 . \quad (\text{A.2})$$

To quantize such a theory one defines canonical momenta

$$p_i = \frac{\partial L}{\partial \dot{q}_i} , \quad (\text{A.3})$$

and a Hamiltonian

$$H = \sum_i p_i \dot{q}_i - L . \quad (\text{A.4})$$

For two classical quantities  $A$  and  $B$  one may define a Poisson bracket

$$\{A, B\}_{\text{PB}} = \sum_\ell \left( \frac{\partial A}{\partial q_\ell} \frac{\partial B}{\partial p_\ell} - \frac{\partial A}{\partial p_\ell} \frac{\partial B}{\partial q_\ell} \right) \quad (\text{A.5})$$

To go to the quantum theory one replace  $A$  and  $B$  by Hilbert space operators, and furthermore one replaces the Poisson brackets  $\{A, B\}_{\text{PB}}$  by the commutator  $\frac{1}{i\hbar}[A, B]$ . For  $A = q_i$  and  $B = p_j$  this procedure leads to the well-known Heisenberg relation

$$[q_i, p_j] = i\hbar \delta_{ij} . \quad (\text{A.6})$$

A classical field theory is given by a set of “coordinates”  $\phi^a(\vec{x})$ , where  $\vec{x}$  is a space-time point and  $a$  some additional discrete label, for example a space-time vector index or an index of some internal symmetry of the theory. The dynamics is governed by a Lagrangian

$$L(\phi^a(\vec{x}, t), \dot{\phi}^a(\vec{x}, t)) \quad (\text{A.7})$$

from which one can compute the action as  $S = \int dt L$ . To write down the Euler-Lagrange equations one needs a functional derivative. A *functional* is an object that assigns a number to a function. To only example we will ever encounter is an integral. A functional derivative can be viewed as the derivative of the functional with respect to the function.

For example, consider a function  $\phi(\vec{x})$  and the functional

$$F[\phi] = \int d^3x f(\phi(\vec{x})) \quad (\text{A.8})$$

This should be thought of as a continuous version of

$$F(q_i) = \sum_i f(q_i) \quad (\text{A.9})$$

To differentiate this, all we need is the rule

$$\frac{\partial q_i}{\partial q_j} = \delta_{ij} \quad (\text{A.10})$$

In the present case this generalizes to

$$\frac{\delta\phi(\vec{x})}{\delta\phi(\vec{y})} = \delta(\vec{x} - \vec{y}) , \quad (\text{A.11})$$

where we use a  $\delta$  instead of  $\partial$  to distinguish a functional derivative from an ordinary one. Apart from the notation, the rules for functional differentiation are quite similar to those of ordinary differentiation. For example, one may use the chain rule, and consequently the functional derivative of  $F$  is

$$\begin{aligned} \frac{\delta F[\phi]}{\delta\phi(\vec{y})} &= \int d^3x f'(\phi(\vec{x})) \frac{\delta\phi(\vec{x})}{\delta\phi(\vec{y})} \\ &= \int d^3x f'(\phi(\vec{x})) \delta(\vec{x} - \vec{y}) \\ &= f'(\phi(\vec{y})) . \end{aligned} \quad (\text{A.12})$$

The Euler-Lagrange equations are in this case

$$\frac{d}{dt} \left( \frac{\delta L}{\delta\dot{\phi}^a(\vec{y})} \right) - \left( \frac{\delta L}{\delta\phi^a(\vec{y})} \right) = 0 . \quad (\text{A.13})$$

This is not the most convenient form, since in Lorentz covariant field theories time derivatives appear in combination with space derivatives, as  $\partial_\mu$ . The symmetry between space and time is restored by defining a Lagrangian density, which is simply the integrand of the Lagrangian

$$L = \int d^3x \mathcal{L} \quad (\text{A.14})$$



In terms of  $\mathcal{L}$  the Euler Lagrange equations are

$$\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0, \quad (\text{A.15})$$

These are ordinary derivatives, because we see  $\mathcal{L}$  as a function (not a functional) of  $\phi$  and  $\partial_\mu \phi$ . This version of the Euler-Lagrangian can either be derived from (A.13), or directly from the variation of  $S = \int d^4x \mathcal{L}$ .

Canonical momenta are defined as

$$\pi^a(\vec{x}) = \frac{\delta L}{\delta \dot{\phi}^a(\vec{x})} \quad (\text{A.16})$$

A bit more care is needed when there is a distinction between upper and lower indices. If  $g_{ab}$  is a symmetric matrix and the action depends on combinations  $\phi^a g_{ab} \phi^b$  it is customary to define  $\phi_a = \sum_b g_{ab} \phi^b$ . Indices can be raised by using  $g^{ab}$ , which is the inverse of  $g_{ab}$ :  $\sum_b g_{ab} g^{bc} = \delta_a^c$ . Suppose we have a functional  $F = \int d^3x V_c \phi^c(\vec{x})$ , with  $V_c$  a constant vector. Then

$$\frac{\delta F}{\delta \phi^c(\vec{x})} = V_c, \quad (\text{A.17})$$

using

$$\frac{\delta \phi^a(\vec{x})}{\delta \phi^b(\vec{y})} = \delta_b^a \delta(\vec{x} - \vec{y}) \quad (\text{A.18})$$

One sees this more explicitly by leaving all indices upper, and writing the matrix  $g_{ab}$  explicitly. Hence the derivative with respect to a quantity with an upper (lower) index yields a quantity with a lower (upper) index.

This applies also to canonical momenta and Poisson brackets. The canonical momentum of  $\phi^a$  is then (if there is a distinction between upper and lower indices)

$$\pi_a = \frac{\delta L}{\delta \dot{\phi}^a(\vec{x})} \quad (\text{A.19})$$

The Hamiltonian is

$$H = \int d^3x \sum_a \pi_a(\vec{x}) \dot{\phi}^a(\vec{x}) - L \quad (\text{A.20})$$

and canonical quantization yields

$$[\pi_a(\vec{x}), \phi^b(\vec{y})] = i\hbar \delta_a^b \delta(\vec{x} - \vec{y}) \quad (\text{A.21})$$

Often one converts the first index to an upper index, and then the result is of course

$$[\pi^a(\vec{x}), \phi^b(\vec{y})] = i\hbar g^{ab} \delta(\vec{x} - \vec{y}) \quad (\text{A.22})$$

## A.1 The Euclidean action

The Euclidean action is obtained by a continuation from Minkowski space, by making time imaginary. This could mean either  $t = it_E$  or  $t = -it_E$ . The choice between these two options is determined by the requirement that the analytic continuation should make the path integral a damped exponential. Consider the weight factor in the path integral,  $\exp(iS)$ . In this expression  $S = \int dt L$ , and we consider a simple Lagrangian of the form  $\frac{1}{2}\dot{q}^2 - V(q)$ , so that the Hamiltonian is  $H = \frac{1}{2}p^2 + V(q)$ . Once we know the correct continuation for this action, we know it for all action.

If the potential  $V$  is bounded from below, as it should in a physical system, then clearly the correct definition for the potential term alone is  $t = -it_E$ . If the Minkowski time integral is from 0 to  $T$  we get now

$$iS = i \int_0^T dt \left( \frac{1}{2} \left( \frac{dq}{dt} \right)^2 - V(q) \right) = \int_0^{T_E} dt_E \left( -\frac{1}{2} \left( \frac{dq}{dt_E} \right)^2 - V(q) \right) \quad (\text{A.23})$$

with  $T = -iT_E$ . We see that both the kinetic and the potential term give a negative exponential. It is conventional to choose the Euclidean action in such a way that the path-integral is dominated by its minima (and not its maxima). Hence one defines

$$\exp(iS(t = -it_E)) \equiv \exp(-S_E) , \quad (\text{A.24})$$

with  $S_E$  a positive definite expression. In our case

$$S_E = \int_0^{T_E} dt_E \left( \frac{1}{2} \left( \frac{dq}{dt_E} \right)^2 + V(q) \right) \quad (\text{A.25})$$

Note that this definition contains an extra minus sign compared to the naively expected relation, and that in particular the  $V(q)$  contribution appears with opposite sign in  $S$  and  $S_E$ :

$$S_E = - \int_0^{T_E} dt_E L_M(t = -it_E) , \quad (\text{A.26})$$

where  $L_M$  is the Lagrangian in Minkowski space.

## B Path integrals

The derivation of (6.22) goes as follows (this discussion is based on appendix A of [2]). We consider a quantum mechanical system with one coordinate  $q$  and a momentum  $p$ . The corresponding quantum operators  $\hat{q}$  and  $\hat{p}$  satisfy  $[\hat{q}, \hat{p}] = i$ . Furthermore we define eigenstates  $|q\rangle$  (with  $\hat{q}|q\rangle = q|q\rangle$ ) and  $|p\rangle$  (with  $\hat{p}|p\rangle = p|p\rangle$ ), and we have

$$\begin{aligned} \langle q|p\rangle &= e^{ipq} \\ \langle q|q'\rangle &= \delta(q - q') \\ \langle p|p'\rangle &= 2\pi\delta(p - p') \end{aligned} \quad (\text{B.1})$$

The integration measures for position and momentum integral are respectively

$$\int dq |q\rangle \langle q| = 1; \quad \int \frac{dp}{2\pi} |p\rangle \langle p| = 1. \quad (\text{B.2})$$

We divide the time-interval  $T$  into  $N$  equal steps  $\epsilon$ ,  $T = N\epsilon$ , and define  $t_m = m\epsilon$ . At  $t = 0$  a complete set of states is  $|q\rangle$ . We define a set of states  $|q, t\rangle$  as

$$|q, t\rangle = e^{iHt} |q\rangle \quad (\text{B.3})$$

This introduces a complete set of states at every time  $t_m$ . In general we have the completeness relation

$$\langle A|B\rangle = \int dq \langle A|q, t\rangle \langle q, t|B\rangle. \quad (\text{B.4})$$

Applying this to each intermediate time  $t_m$  yields for a transition amplitude

$$\langle q_f, T|q_i, 0\rangle = \int dq_1 \dots dq_{N-1} \prod_{m=0}^{N-1} \langle q_{m+1}, t_{m+1}|q_m, t_m\rangle \quad (\text{B.5})$$

Now we express all  $|q, t\rangle$  in terms of  $|q\rangle \equiv |q, 0\rangle$  using (B.3):

$$\begin{aligned} \langle q_{m+1}, t_{m+1}|q_m, t_m\rangle &= \langle q_{m+1}| e^{-iHt_{m+1}} e^{iHt_m} |q_m\rangle \\ &= \langle q_{m+1}| e^{-iH\epsilon} |q_m\rangle \end{aligned} \quad (\text{B.6})$$

Now we insert a complete set of momentum states

$$\langle q_{m+1}, t_{m+1}|q_m, t_m\rangle = \int \frac{dp_m}{2\pi} \langle q_{m+1}|p_m\rangle \langle p_m| e^{-iH\epsilon} |q_m\rangle \quad (\text{B.7})$$

The classical Hamiltonian is a function of  $p$  and  $q$ ,  $H(p, q)$ . The quantum Hamiltonian is ambiguous because of the ordering of  $\hat{p}$  and  $\hat{q}$ . We will define it by requiring that all  $\hat{p}$  are always to the left of  $\hat{q}$ . Then

$$\langle p_m| H(\hat{p}, \hat{q}) |q_m\rangle = H(p_m, q_m) \langle q|p\rangle \quad (\text{B.8})$$

Even if  $\hat{p}$  and  $\hat{q}$  are ordered correctly in  $H$ , the exponential  $\exp(-\epsilon H)$  still contains improperly ordered terms. However, they are always of order  $\epsilon^2$  and will be dropped. Then we get

$$\begin{aligned} \langle q_{m+1}, t_{m+1}|q_m, t_m\rangle &= \int \frac{dp_m}{2\pi} e^{-iH(p_m, q_m)\epsilon} \langle q_{m+1}|p_m\rangle \langle p_m|q_m\rangle \\ &= \int \frac{dp_m}{2\pi} e^{-i[H(p_m, q_m)\epsilon - p_m(q_{m+1} - q_m)]}, \end{aligned} \quad (\text{B.9})$$

where in the last step (B.1) was used.

We will only be interested in Hamiltonians of the form  $H = \frac{1}{2}ap^2 + V(q)$ , where  $a$  is some constant. In that case the  $p_m$  integrations in (B.9) can be done explicitly. At this point it is better to go to Euclidean space, by defining  $\epsilon = -i\delta$ . The sign choice is dictated by the requirement that we get a damped Gaussian integral. The relevant integral is

$$\begin{aligned} \int \frac{dp}{2\pi} \exp(-\frac{1}{2}a\delta p^2 + ipd) &= \int \frac{dp}{2\pi} \exp(-\frac{1}{2}[a\delta(p - id/a\delta)^2 + d^2/a\delta]) \\ &= \frac{1}{\sqrt{2\pi a\delta}} \exp(-\frac{1}{2}d^2/a\delta), \end{aligned} \quad (\text{B.10})$$

with  $d = q_{m+1} - q_m$ . Now we get

$$\begin{aligned} \langle q_f, T|q_i, 0 \rangle &= \int \prod_{i=1}^{N-1} dq_i \left[ \frac{1}{\sqrt{2\pi a\delta}} \right]^N \exp\left(-\sum_{m=0}^{N-1} \left[ \frac{1}{2}(q_{m+1} - q_m)^2/a\delta - V(q_m)\delta \right]\right) \\ &= \int \prod_{i=1}^{N-1} dq_i \left[ \frac{1}{\sqrt{2\pi a\delta}} \right]^N \exp\left(-\sum_{m=0}^{N-1} \delta \left[ \frac{1}{2} \left( \frac{q_{m+1} - q_m}{\delta} \right)^2/a - V(q_m) \right]\right) \end{aligned} \quad (\text{B.11})$$

The limit  $\delta \rightarrow 0$  defines the path-integral as (writing  $T = -i\beta$ , so that  $\beta = N\delta$ )

$$\begin{aligned} \langle q_f, \beta|q_i, 0 \rangle &= \int \mathcal{D}q \exp \int_0^\beta dt \left[ -\frac{1}{2a} \dot{q}^2 - V(q) \right] \\ &= \int \mathcal{D}q \exp(-S_E(\dot{q}, q)) \end{aligned} \quad (\text{B.12})$$

Here we have absorbed the prefactor into the definition of  $\mathcal{D}q$ . Note that the normalization is anyhow fixed by the limit  $\beta \rightarrow 0$ , which should yield the normalization of the state. The result can also be written as

$$\langle q_f, 0|e^{-H\beta}|q_i, 0 \rangle = \int \mathcal{D}q \exp(-S_E) \quad (\text{B.13})$$

The path integral is over all paths  $q(t)$  with  $q(0) = q_i$  and  $q(\beta) = q_f$ . Now we take  $q_i = q_f \equiv q$  and integrate over all  $q_i$ . Then we get

$$\text{Tr } e^{-\beta H} \equiv \int dq \langle q, 0|e^{-H\beta}|q, 0 \rangle = \int_{PBC} \mathcal{D}q \exp(-S_E) \quad (\text{B.14})$$

The extra integral has extended the path-integral with an integral over the initial position. We are now integrating over all paths satisfying  $q(0) = q(\beta) = q$ , with an integration over  $q$ . Hence this is a path integral over all paths satisfying a periodic boundary condition.

## C Self-dual lattices

Here we summarize a few results on even self-dual lattices.

Lorentzian even self-dual lattices  $\Gamma_{p,q}$  are defined by the requirement that all vectors have even length, and the  $\Gamma_{p,q}$  is self-dual. Both length and duality are defined with an inner product of the form

$$(\vec{v}, \vec{w})^2 = \sum_{i=1}^p v_i^2 - \sum_{j=1}^q w_j^2, \quad (\text{C.1})$$

where  $(\vec{v}, \vec{w})$  is a vector on the lattice.

Lorentzian even self-dual lattices (ESDL's) are classified modulo Lorentz transformations. Up to such transformations there is just one lattice, *i.e.* all ESDL's are related to each other by Lorentz transformations. Physically this information is useless. Particle spectra depend on  $\vec{v}^2$  and  $\vec{w}^2$  separately, not just on their difference. Only rotations in the first  $p$  or the last  $q$  components make no difference. In other words, the condition for modular invariance is Lorentz invariant, but the spectrum is not.

Since all Lorentzian ESDL's are Lorentz rotations of each other, the parameters of the Lorentz transformation are the moduli of the compactification. The full parameter space ("moduli-space") is in fact the *coset space*

$$\frac{SO(p, q)}{SO(p) \times SO(q)}, \quad (\text{C.2})$$

This means that we consider the Lorentz-transformations in  $SO(p, q)$ , but we identify those Lorentz-transformations that are related to each other by rotations in the first  $p$  or the last  $q$  components. The precise moduli space is still a bit smaller: because of T-duality and its generalizations on higher-dimensional tori some parts of the moduli space are identified with each other.

The foregoing holds if  $p \neq 0, q \neq 0$ . If  $q = 0$ , ESDL's exist only if  $p$  is a multiple of 8. The first case is  $p = 8$ , and it turns out that the lattice is then unique and given by

$$\begin{aligned} &(0, \dots, 0, \pm 1, 0, \dots, 0, \pm 1, 0, \dots, 0) \\ &\left(\frac{1}{2}, \dots, \frac{1}{2}\right) \end{aligned} \quad (\text{C.3})$$

This is an over-complete basis, but in any case the idea is that one builds a lattice by adding all these vectors in any integer linear combination. The vectors shown here have length 2, and in total there are 240 of them: 112 of the first type, and 128 of the second type (note that by adding the first set one can get all even sign changes of the second vector). This lattice is called  $E_8$  (it is in fact the root lattice of the Lie-algebra  $E_8$ ).

By taking this lattice for left- and right-movers we can build a lorentzian ESDL  $\Gamma_{8,8} = (E_8)_L \times (E_8)_R$ . The vectors of this lattice are of the form  $(\vec{v}_L, \vec{v}_R)$  with  $\vec{v}_L$  and  $\vec{v}_R$  vectors of the above form. On this lattice we may compactify the bosonic string to 16 dimensions. In the resulting spectrum the states of the form  $\bar{\alpha}_{-1}^i |v_L, 0\rangle$  and  $\alpha_{-1}^I \bar{\alpha}_{-1}^i |0, 0\rangle$  (with  $v_L^2 = 2$  and  $I = 1, \dots, 8$ ) are massless vector bosons. There are  $240 + 8 = 248$  such bosons from the left-moving and the same number from the right-moving modes. It turns out that massless vector bosons in string theory are always gauge bosons. In this case they are gauge bosons with a Lie-algebra  $(E_8)_L \times (E_8)_R$ . Here  $E_8$  is a Lie-algebra specified, just

as  $SU(2)$ , by a set of commutation relations. We will not explain its construction here. It has of course precisely 248 generators.

If  $p = 16$  there are two solutions:  $(E_8)^2$  and a lattice one might call  $E_{16}$ . It is defined as  $E_8$  above, but with 16-component vectors. Because of this the second vector,  $(\frac{1}{2}, \dots, \frac{1}{2})$  now has length 4 and would not produce a massless gauge boson when we compactify on  $\Gamma_{16,16} = (E_{16})_L \times (E_{16})_R$ . Indeed, now there are 496 massless gauge bosons from the left and the same number from the right. They all originate from the first type of lattice vector in (C.3), plus the 16 oscillators  $\alpha_{-1}^I$ . The number of length-2 vectors is  $\frac{1}{2} \times (16 \times 15) \times 4 = 480$ . The gauge group is in fact  $SO(32)_L \times SO(32)_R$ . The name  $E_{16}$  is not standard. Usually one calls the lattice without the half-integer vectors  $D_{16}$ : it is the root lattice of  $SO(32)$ . The half-integer spin vector is an  $SO(32)$  spinor weight. So the correct description of the “ $E_{16}$ ” lattice would be the “ $D_{16}$  root lattice with a spinor conjugacy class added”.

For profound reasons  $496 = 2 \times 248$ , so that both  $p = 16$  lattices give the same number of gauge bosons, but different gauge groups.

If  $p = 24$  there are 24 ESDL’s, and for  $p = 32$  the number explodes to an unknown number larger than  $10^7$  (and undoubtedly much larger)!

## D Lie algebra dictionary

Here we collect some formulas and conventions for Lie-algebras. This is not a review of group theory, but rather an “encyclopedia dictionary” of some relevant facts with few explanations.

### D.1 The algebra

We will mainly use compact groups (see below). Their Lie algebras can be characterized by a set of  $\dim(A)$  Hermitean generators  $T^a$ ,  $a = 1, \dots, \dim(A)$ , where  $A$  stands for “adjoint”. Provided a suitable basis choice is made, the generators satisfy the following algebra

$$[T^a, T^b] = i f^{abc} T_c, \tag{D.1}$$

with structure constants  $f_{abc}$  that are real and completely anti-symmetric.

### D.2 Exponentiation

Locally, near the identity, the corresponding Lie group can be obtained by exponentiation

$$g(\alpha^a) = e^{i\alpha^a T^a}. \tag{D.2}$$

The global properties of the group, involving element not “close” to 1, are not fully described by the Lie-algebra alone, but will not be discussed here. The space formed by all the group elements is called the *group manifold*.

### D.3 Real forms

A Lie-algebra is a vector space of dimension  $\dim(A)$  with an additional operation, the commutator. An arbitrary element of the vector space has the form  $\sum_a \alpha_a T^a$ . In applications to physics  $\alpha_a$  is either a real or a complex number. If the coefficients  $\alpha^a$  are all real and the generators Hermitian, the group manifold is a compact space. For a given compact group there is a unique complex Lie-algebra, which is obtained simply by allowing all coefficients  $\alpha_a$  to be complex. Within the complex algebra there are several real sub-algebras, called *real forms*. The generators of such a sub-algebra can be chosen so that (D.1) is satisfied with all structure constants real, but with generators that are not necessarily Hermitian. One can always obtain the real forms from the compact real form (which has hermitian generators) by choosing a basis so that the generators split into two sets,  $\mathcal{H}$  and  $\mathcal{K}$ , so that  $[\mathcal{H}, \mathcal{H}] \in \mathcal{H}$  and  $[\mathcal{K}, \mathcal{K}] \in \mathcal{H}$ . Then one may consistently replace all generators  $K \in \mathcal{K}$  by  $iK$  without affecting the reality of the coefficients  $f^{abc}$ . The most common case in physics are the real forms  $SO(n, m)$  of the compact real form  $SO(n + m)$ . Most of the following results hold for the compact real form of the algebra, unless an explicit statement about non-compact forms is made.

### D.4 The classical Lie groups

The group  $SU(N)$  is the group of unitary  $N \times N$  matrices with determinant 1;  $SO(N)$  is the group of real orthogonal matrices with determinant 1, and  $Sp(2r)$  the group of real  $2r \times 2r$  matrices  $S$  that satisfies  $S^T M S = M$ , where  $M$  is a matrix which is block-diagonal in terms of  $2 \times 2$  blocks of the form

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{D.3}$$

Mathematicians (and some physicist) write  $Sp(r)$  instead of  $Sp(2r)$ .

### D.5 Simple Lie-algebras

Lie-algebras are in general a “product” (or, more accurately, a direct sum) of a semi-simple Lie-algebra and some  $U(1)$ ’s. The latter require no further discussion. Semi-simple algebras are a product of various simple ones; the simple Lie-algebras have been classified completely, see below.

### D.6 The Cartan sub-algebra

This is the maximal set of commuting generators of the simple algebra. All such sets can be shown to be equivalent. The dimension of this space is called the *rank* (denoted  $r$ ) of the algebra.

## D.7 Roots

If we denote the Cartan sub-algebra generators as  $H_i, i = 1, \dots, r$ , then the remaining generators can be chosen so that

$$[H_i, E_{\vec{\alpha}}] = \alpha^i E_{\vec{\alpha}} . \quad (\text{D.4})$$

The eigenvalues with respect to the Cartan sub-algebra are vectors in a space of dimension  $r$ . We label the remaining generators by their eigenvalues  $\vec{\alpha}$ . These eigenvalues are called the *roots* of the algebra.

## D.8 Positive roots

A positive root is a root whose first component  $\alpha_1$  is positive in some fixed basis. This basis must be chosen so that, for all roots,  $\alpha_1 \neq 0$ .

## D.9 Simple roots

Simple roots are positive roots that cannot be written as positive linear combinations of other positive roots. There are precisely  $r$  of them. They form a basis of the vector space of all the roots. The set of simple roots of a given algebra is unique up to  $O(r)$  rotations. In particular it does not depend on the choice of the Cartan sub-algebra or the basis choice in “root space”. This set is thus completely specified by their relative lengths and mutual inner products. The inner product used here, denoted  $\vec{\alpha} \cdot \vec{\beta}$ , is the straightforward Euclidean one.

## D.10 The Cartan matrix

The Cartan matrix is defined as

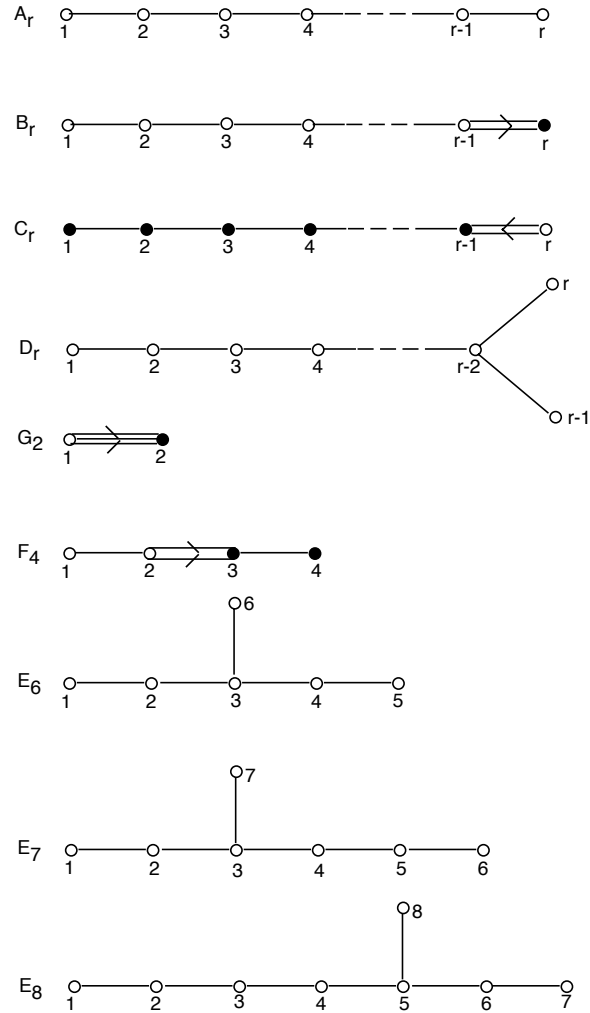
$$A^{ij} = 2 \frac{\vec{\alpha}_i \cdot \vec{\alpha}_j}{\vec{\alpha}_j \cdot \vec{\alpha}_j} , \quad (\text{D.5})$$

where  $\vec{\alpha}_i$  is a simple root. This matrix is unique for a given algebra, up to permutations of the simple roots. One of the non-trivial results of Cartan’s classification of the simple Lie algebras is that all elements of  $A$  are integers. The diagonal elements are all equal to 2 by construction; the off-diagonal ones are equal to 0,  $-1$ ,  $-2$  or  $-3$ .

## D.11 Dynkin diagrams

are a graphical representation of the Cartan matrix. Each root is represented by a dot. The dots are connected by  $n$  lines, where  $n$  is the maximum of  $|A_{ij}|$  and  $|A_{ji}|$ . If  $|A_{ij}| > |A_{ji}|$  an arrow from root  $i$  to root  $j$  is added to the line. The simple algebras are divided into 7 classes, labelled A–G, with Dynkin diagrams as shown below.





## D.12 Long and short roots

If a line from  $i$  to  $j$  has an arrow,  $A_{ij} \neq A_{ji}$  and hence the lengths of roots  $i$  and  $j$  are not the same. An arrow points always from a root to another root with smaller length. Lines without arrows connect roots of equal length. There is at most one line with an arrow per diagram, and therefore there are at most two different lengths. This is not only true for the simple roots, but for all roots. One frequently used convention is to give all the long roots length-squared equal to two. Then the short roots have length 1 if they are connected to the long ones by a double line, and length-squared  $\frac{2}{3}$  if they are connected by a triple line. Often the short roots are labelled by closed dots, and the long ones by open dots, although this is strictly speaking superfluous. If all roots have the same length the algebra is called *simply laced*. This is true for types A, D and E.

### D.13 Realizations

The compact Lie-algebras corresponding to types  $A - D$  are realized by the algebras  $SU(n)$ ,  $SO(n)$  and  $Sp(n)$ . The correspondence is as follows

$$\begin{aligned} A_r &: SU(r+1) \\ B_r &: SO(2r+1) \\ C_r &: Sp(2r) \\ D_r &: SO(2r) \end{aligned} \tag{D.6}$$

There is no such simple characterization for the algebras of types  $E, F$  and  $G$ , the exceptional algebras.

### D.14 Representations

A set of unitary  $N \times N$  matrices satisfying the algebra (D.1) is said to form a (unitary matrix) representation of dimension  $N$ .

### D.15 Equivalence

If a set of hermitean generators  $T^a$  satisfy the algebra, then so do  $\tilde{T}^a = U^\dagger T^a U$ , if  $U$  is unitary. Then  $T^a$  and  $\tilde{T}^a$  are called equivalent.

### D.16 Real, complex and pseudo-real representations

The *complex conjugate representation* is the set of generators  $-(T^a)^*$ , which obviously satisfy the algebra if  $T^a$  does. A representation is real if a basis exists so that for all  $a$   $-(T^a)^* = T^a$  (in other words, if a  $\tilde{T}^a = U^\dagger T^a U$  exists so that all generators are purely imaginary). An example of a real representation is the adjoint representation. A representation is pseudo-real if it is not real, but only real up to equivalence, *i.e.*  $-(T^a)^* = C^\dagger T^a C$  for some unitary matrix  $C$ . Otherwise a representation is called complex.

A frequently occurring example of a pseudo-real representation is the two-dimensional one of  $SU(2)$ . The generators are the Pauli matrices, and only  $\sigma_2$  is purely imaginary. However, if one conjugates with  $U = i\sigma_2$  the other two matrices change sign, so that indeed  $-\sigma_i^* = U^\dagger \sigma_i U$ .

### D.17 Irreducible representations

If a non-trivial subspace of the vector space on which a representation acts is mapped onto itself (an “invariant subspace”) the representation is called reducible. Then all  $T$ ’s can be simultaneously block-diagonalized, and each block is by itself a representation. If there are no invariant subspaces the representation is called irreducible.

## D.18 Weights

In any representation the matrices representing the Cartan sub-algebra generators  $H_i$  can be diagonalized simultaneously. The space on which the representation acts decomposes in this way into eigenspaces with a set of eigenvalues  $\vec{\lambda}$ , *i.e.*  $H_i v_{\vec{\lambda}} = \lambda_i v_{\vec{\lambda}}$ . The  $\vec{\lambda}$ 's, which are vectors in the vector space spanned by the roots, are called weights. The vector space is usually called weight space.

## D.19 Weight space versus representation space

We are now working in two quite different vector spaces: the  $r$ -dimensional weight space, and the  $N$  dimensional space on which the representation matrices act. The former is a real space, the latter in general a complex space. Often the vectors in the latter space are referred to as “states”, a terminology borrowed from quantum mechanics. Although this may be somewhat misleading in applications to classical physics, it has the advantage of avoiding confusion between the two spaces.

## D.20 Weight multiplicities

In the basis in which all Cartan sub-algebra generators are simultaneously diagonal each state in a representation are characterized by some weight vector  $\lambda$ . However, this does not characterize states completely, since several states can have the same weight. The number of states in a representation  $R$  that have weight  $\lambda$  is called the multiplicity of  $\lambda$  in  $R$ .

## D.21 Co-roots

Co-roots  $\hat{\alpha}$  are defined as  $\hat{\alpha} = \frac{2\vec{\alpha}}{\alpha \cdot \alpha}$ .

## D.22 Dynkin labels

For any vector  $\lambda$  in weight space we can define Dynkin labels  $l_i$  as  $l_i = \lambda \cdot \hat{\alpha}_i = 2 \frac{\lambda \cdot \alpha_i}{\alpha_i \cdot \alpha_i}$ . Since the simple (co)roots form a complete basis, these Dynkin labels are nothing but the components of a weight written with respect to a different basis. The advantage of this basis is that it can be shown that for any unitary representation of the algebra the Dynkin labels are integers.

## D.23 Highest weights

Every irreducible representation of a simple Lie-algebra has a unique weight  $\lambda$  so that on the corresponding weight vector  $E_{\alpha} v_{\lambda} = 0$  for all positive roots  $\alpha$ . Then  $\lambda$  is called the highest weight of the representation. Its Dynkin labels are non-negative integers. Furthermore for every set of non-negative Dynkin labels there is precisely one irreducible representation whose highest weight has these Dynkin labels.

## D.24 The irreducible representations of a simple Lie-algebra

can thus be enumerated by writing a non-negative integer next to each node of the Dynkin diagram. The states in a representation can all be constructed by acting with the generators  $E_\alpha$  on the highest weight state. This state always has multiplicity 1.

## D.25 Special representations

- *Fundamental representations*

The representations with Dynkin labels  $(0, 0, \dots, 0, 1, 0, \dots, 0)$  are called the fundamental representations.

- *The adjoint representation*

The adjoint representation is the set of generators  $(T^a)_{bc} = -if^{abc}$ ; it has dimension  $\dim(A)$ .

- *Vector representations*

The  $N \times N$  matrices that were used above to define the classical Lie groups form the vector representation of those groups; the expansion of these matrices around the identity yields the vector representation of the corresponding Lie-algebra.

- *Fundamental spinor representations*

They are defined only for  $SO(N)$ . If  $N$  is odd, they have Dynkin label  $(0, 0, \dots, 0, 1)$ . If  $N$  is even there are two fundamental spinor representations with Dynkin labels  $(0, 0, 0, \dots, 1, 0)$  and  $(0, 0, 0, \dots, 0, 1)$ .

## D.26 Tensor products

If  $V_{i_1}$  transforms according to some representation  $R_1$  and  $W_{i_2}$  according to some representation  $R_2$ , then obviously the set of products  $V_{i_1}W_{i_2}$  forms a representation as well. This is called the tensor product representation  $R_1 \times R_2$ ; it has dimension  $\dim R_1 \dim R_2$ . This representation is usually not irreducible. It can thus be decomposed into irreducible representations:

$$R_1 \times R_2 = \sum_j N_{12j} R_j, \quad (\text{D.7})$$

where  $N_{12j}$  is the number of times  $R_j$  appears in the tensor product.

## References

- [1] M. Green, J. Schwarz and E. Witten, *Superstring Theory*, parts 1 and 2, Cambridge Univ. Press, (1986).
- [2] J. Polchinski *String Theory*, parts 1 and 2, Cambridge Univ. Press, (1998).

- [3] D. Lüst and S. Theisen, *Lectures on String Theory*. Springer (1989).
- [4] B. Zwiebach, *A first course in String Theory*. Cambridge University Press, Cambridge. 2005.
- [5] W. Lerche, A. Schellekens and N. Warner, *Lattices and Strings*, Physics Reports 177 (1989) 1.
- [6] B. de Wit and J. Smith, *Field Theory and Particle Physics*,(North Holland).