Introduction to String Theory

Angel M. Uranga

Contents

Part I

Introductory Overview

Chapter 1

Motivation

1.1 Standard Model and beyond

1.1.1 Our Model of Elementary Particles and Interactions

Our description of particles and interactions treats strong-electroweak interactions and gravitational interactions in a very different way.

• Electromagnetic, weak and strong interactions are described by a **quantum gauge field theory**. Interactions are mediated by gauge vector bosons, associated with the gauge group

$$SU(3)_c \times SU(2)_W \times U(1)_Y \tag{1.1}$$

While matter is described by left-handed Weyl fermions in the following representation of the gauge group

$$\begin{array}{ll} 3\,[\,(3,2)_{1/6} \,+\, (\overline{3},1)_{1/3} \,+\, (\overline{3},1)_{-2/3} + & Q_L\,,\,U\,,\,D \\ \quad +\,(1,2)_{-1/2} \,+\, (1,1)_1\,] + 3(1,1)_0 & E\,,\,L\,,\,\nu_R \end{array} \tag{1.2}$$

where the subscript denotes $U(1)_Y$ charge (hypercharge), and where we have also included right-handed neutrinos (although they have not been observed experimentally).

An important property of these fermions is their chirality (this is at the heart of parity violation in the Standard Model). There are no left-handed Weyl fermions with conjugate quantum numbers (if there would be, we could rewrite the pair as a left-handed and a right-handed Weyl fermion, both with equal quantum numbers; this is called a vector-like pair, and does not violate parity, it is non-chiral).

Our description considers all these objects to be pointlike. This assumption works as far as the model has been tested experimentally, i.e. up to energies about 1 TeV.

In order to break the electroweak symmetry $SU(2)_W \times U(1)_Y$ down to the $U(1)$ of electromagnetism, the model contains a Higgs sector, given by a complex scalar ϕ with quantum numbers

$$(2,1)_{-1/2} \tag{1.3}$$

The theory contains a scale M_W, which is the scale of pontaneous breaking of the symmetry [1]. It is fixed by the vacuum expectation value $<\phi>$ acquired by the scalar, as determined by a potential of the form

$$V(\phi) \,=\, -m^2\,\phi^*\phi \,+\, \lambda\,(\phi^*\phi)^2 \tag{1.4}$$

The electroweak scale is then

$$M_W \simeq <\phi> \simeq \frac{m}{\sqrt{\lambda}} \simeq 10^2\,\mathrm{GeV} \tag{1.5}$$

Chirality of the fermions forbid writing a Dirac mass term for them. The only way for them to get a mass is via coupling to the Higgs multiplet via Yukawa couplings schematically of the form

$$Q_L\,U\,\phi \quad ; \quad Q_L\,D\,\phi^* \quad ; \quad L\,E\,\phi \tag{1.6}$$

so the scale of fermion masses is linked to the scale of electroweak symmetry breaking.

This theory is well defined at the quantum mechanical level, it is unitary, renormalizable (leaving the issue of 'triviality' of the Higgs sector aside), etc...

• On the other hand, the gravitational interactions are described by the classical theory of general relativity. Interactions are encoded in the

[1] To be fair, there is also a further scale in the model, the QCD scale around 1 GeV, which is understood in terms of dimensional transmutation, i.e. it is the energy at which the $SU(3)$ coupling constant becomes strong.

spacetime metric $G_{\mu\nu}$ via the principle of diffeomorphism (or coordinate reparametrization) invariance of the physics. This leads to an action of the form

$$S_{grav} \,=\, M_P^2 \int_{X_4} R\sqrt{-G}\, d^4x \tag{1.7}$$

with a typical scale of

$$M_P \simeq 10^{19}\,\mathrm{GeV} \tag{1.8}$$

Four-dimensional Einstein theory has been tested experimentally to be good description of the gravitational interactions down to length scales of about 10^{-7} m.

Since the interaction contains an explicit dimensionful coupling, it is difficult to make sense of the theory at the quantum level. They theory is non-renormalizable, it presents loss of unitarity at loop levels, it cannot be quantized in the usual fashion, it is not well defined in the ultraviolet.

The modern viewpoint is that Einstein theory should be regarded as an effective field theory, which is a good approximation at energies below M_P (or some other cutoff scale at which four-dimensional classical Einstein theory ceases to be valid). There should exist an underlying, quantum mechanically well-defined, theory which exists for all ranges of energy, and reduces to classical Einstein at low energies, below the cutoff scale. Such a theory would be called an ultraviolet completion of Einstein theory (which by itself is ill-defined in the ultraviolet).

1.1.2 Theoretical questions raised by this description

There are many such questions, and have led to a great creative effort by the high energy physics (and general relativity) communities. To be fair, most of them have not been successfully answered, so the quest for solutions goes on. These are some of these questions

• The description is completely schizophrenic! We would like to make gravitational interactions consistent at the quantum mechanical level. Can this really be done? and how?

• Are all interactions described together in a unified setup? Or do they remain as intrinsecally different, up to arbitrary energies? Is there a microscopic quantum theory that underlies the gravitational and the Standard Model gauge interactions? Is there a more modest description which at least

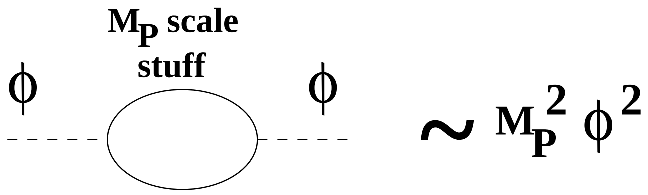


Figure 1.1: Quantum corrections to the Higgs mass due to Planck scale stuff.

unifies the gauge interactions of the Standard Model (leaving for the moment gravity aside)?

• Why are there two different scales, M_W and M_P? Why are there so widely separated? Are they related in any way, and if so, which?

• Why M_W, which is fixed by the mass of the Higgs scalar, is not modified by quantum loops of stuff related to physics at the scale M_P? Power counting would suggest that the natural value of these corrections is of order M_P^2, which would then push the electroweak scale up to the Planck scale.

• Are there other scales between M_W and M_P? or is there just a big desert in energies in between? (there are some suggestions of intermediate masses, for instance from the see-saw mechanism for neutrino masses, which points to new physics at an energy scale of 10^{12} GeV).

• Why the gauge sector is precisely as it is? Why three gauge factors, why these fermion representations, why three families? How are these features determined from an underlying microscopic theory that includes gravity?

• Are global symmetries of the Standard Model exact symmetries of the underlying theory? Or just accidental symmetries? Is baryon number really conserved? Why is the proton stable, and if not what new physics mediates its decay?

• Why are there four dimensions? Is it true that there are just four dimensions? Does this follow from any consistency condition of the theory supposedly underlying gauge and gravitational interactions?

• ..., ..., ... ?

1.1.3 Some proposals for physics beyond the Standard Model

These and other similar questions lie at the origin of many of the ideas of physics beyond the Standard Model. Let us review some of them (keeping in mind that they do not exclude each other, and mixed scenarios are often the most attractive). For a review along similar lines, see e.g. [1].

Grand Unification Theories (GUTs)

See for instance [2, 3].

In this setup the Standard Model gauge group is a low-energy remnant of a larger gauge group. This group G_{GUT} is usually taken to be simple (contains only one factor) like $SU(5)$, $SO(10)$, or E_6, and so unifies all low-energy gauge interactions into a unique kind. The GUT group is broken spontanously by a Higgs mechanism (different form that of the Standard Model, of course) at a large scale M_{GUT}, of about 10^{16}-10^{17} GeV.

This idea leads to a partial explanation of the fermion family gauge quantum numbers, since the different fermions are also unified into a smaller number of representations of G_{GUT}. For $SU(5)$ a Standard Model family fits into a representation $10+\overline{5}$; for $SO(10)$ it fit within an irreducible representation, the 16.

A disadvantage is that the breaking of G_{GUT} down to the Standard Model group requires a complicated scalar Higgs sector. In minimal $SU(5)$ theories, the GUT-Higgs belongs to a 24-dimensional representation; $SO(10)$ is even more involved.

Additional interesting features of these theories are

• Extra gauge interactions in G_{GUT} mediate processes of proton decay (violate baryon number), which are suppressed by inverse powers of M_{GUT}. The rough proton lifetime in these models is around 10^{32} years, which is close to the experimental lower bounds. In fact, some models like minimal $SU(5)$ are already experimentally ruled out because they predict a too fast proton decay.

• If we assume no new physics between M_W and M_{GUT} (desert hypothesis), the Standard Model gauge couplings run with scale towards a unified value at a scale around 10^{16} GeV. This may suggest that the different low-energy interactions are unified at high energies.

Besides these nice features, it is fair to say that grand unified theories do not address the fundamental problem of gravity at the quantum level, or the relation between gravity and the other interactions.

Supersymmetry (susy)

See graduate course by A. Casas, also review like e.g. [4]

Supersymmetry is a global symmetry that relates bosonic and fermionic degrees of freedom in a theory. Infinitesimal supersymmetry transformations are associated so (super)generators (also called supercharges), which are operators whose algebra is defined in terms of anticommutation (rather than commutation) relations (these are the so-called superalgebras, and generate supergroups). The minimal supersymmetry in four dimensions (so-called $D = 4$ $N = 1$ supersymmetry is generated by a set of such fermionic operators Q_α, which transform as a left-handed Weyl spinor under the 4d Lorentz group. The supersymmetry algebra is

$$\{Q_\alpha, Q_\beta\} = (\sigma^\mu)_{\alpha\beta} P_\mu \tag{1.9}$$

where $\sigma^\mu = (\mathbf{1}_2, \sigma^i)$ are Pauli matrices, and P_μ is the four-momentum operator.

A simple realization of supersymmetry transformations is: consider a four-dimensional Weyl fermion ψ^α and a complex scalar ϕ, and realize Q_α acting as

$$\begin{aligned} Q_\alpha \phi &= \psi_\alpha \\ Q_\beta \psi_\alpha &= i(\sigma^\mu)_{\alpha\beta} \partial_\mu \phi \end{aligned} \tag{1.10}$$

The algebra closes on these fields, so the (super)representation (also called supermultiplet) contains a 4d Weyl fermion and a complex scalar. Such multiplet is known as the chiral multiplet. Another popular multiplet of $N = 1$ susy) is the vector multiplet, which contains a four-dimensional massless vector boson and a 4d Weyl fermion (the latter is often re-written as a 4d Majorana fermion).

There exist superalgebras generated by more supercharges, they are called extended supersymmetries. The N-extended supersymmetry is generated by supercharges Q_α^a with $a = 1, \ldots, N$. Any supersymmetry with $N > 1$ is inconsistent with chiral fermions (any multiplet contains fermions with both chiralities, i.e. is vector-like), so such theories have limited phenomenological applications and we will skip them here.

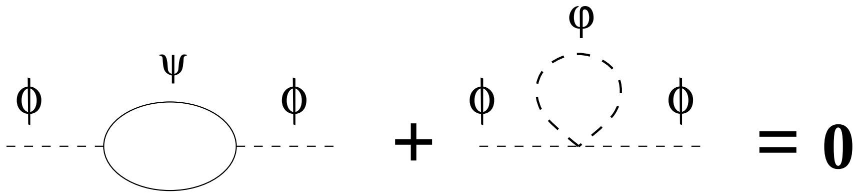


Figure 1.2: Fermionic and bosonic loop corrections to the higgs mass cancel in a supersymmetric theory.

The reason why susy may be of phenomenological interest is that it relates scalars (like the Higgs) with chiral fermions, and the symmetry requires them to have equal mass. The mass of a chiral fermion is forced to be zero by chirality, so the mass of a scalar like the Higgs is protected against getting large $O(M_P)$ corrections, so supersymmetry stabilizes M_W against M_P.

Diagrammatically, any corrections to the Higgs mass due to fermions in the theory are cancelled against corrections to the Higgs mass due to their boson superpartners. There is a non-renormalization theorem of certain couplings in the lagrangian (like scalar masses) which guarantees this to any order in perturbation theory.

SUSY commutes with gauge symmetries. So in trying to build a supersymmetric version of the standard model the simplest possibility is to add superpartners to all observed particles: fermion superpartners (gauginos) for gauge bosons to promote them to vector multiplets; boson superpartners (squarks and sleptons) for the quark and leptons, to promote them to chiral multiplets; and fermion superpartner (higgssino) for the scalar Higgs (for technical reasons, like anomaly cancellation, a second Higgs chiral multiplet must be included). Interactions are dictated by gauge symmetry and supersymmetry. Such model is known as the minimal supersymmetric standard model (MSSM).

However, superpartners have not been observed in Nature, so it is clear that they are not mass-degenerate with usual matter. Supersymmetry is not an exact symmetry of Nature and must be broken. The most successful way to do so, without spoiling the absence of quadratic corrections to the Higgs mass is explicit breaking. That is, to introduce explicitly non-supersymmetric terms of a certain kind (so-called soft terms) in the MSSM lagrangian. These terms render superpartners more massive than standard model fields. Cancellation of loop contributions to the Higgs mass is not exact, but is not quadratically dependent on M_P, only logarithmically. In order to retain 10^2 GeV as a natural scale, superpartner mass scale (supersymmetry breaking scale in the MSSM) should be around 1 TeV or so.

The MSSM is a theoretically well motivated proposal for physics beyond the Standard Model, it is concrete enough and experimentally accessible. It addresses the question of the relation between M_W and M_P. On the other hand, it leaves many others of our questions unanswered.

Supergravity (sugra)

See for instance [5].

It is natural to consider theories where supersymmetry is realized as a local gauge symmetry. Given the susy algebra (1.10), this means that the four-momentum operator P_μ, which generates global translations, is also promoted to a gauge generator. Local translations are equivalent to coordinate reparametrization (or diffeomorphism) invariance

$$x^\mu \to x^\mu + \xi(x) \tag{1.11}$$

so the resulting theories are generalizations of general relativity, and hence contain gravity. They are called supergravities.

A very important 4d $N = 1$ supermultiplet is the gravity multiplet, which contains a spin-2 graviton $G_{\mu\nu}$ and its spin-3/2 superpartner (gravitino) ψ_α^μ (also called Rarita-Schwinger field) . Other multiplets are like in global susy, the chiral and vector multiplets. The sugra lagrangian is basically obtained from the global susy one by adding the Einstein term for the graviton, a kinetic term for the gravitino, and coupling the graviton to the susy theory stress-enery tensor,and coupling the gravitino to the susy theory supercurrent (current associated to the supersymmetry).

In applications to phenomenology, a nice feature of supergravity is that spontaneous breaking of local supersymmetry becomes, in the limit of energies much below M_P, explicit breaking of global supersymmetry by soft terms. A popular scenario is to construct models with a MSSM sector (visible sector), a second sector (hidden sector) decoupled from the MSSM (except by gravitational interactions) and which breaks local supersymmetry at a scale of $M_{hidden} = 10^{12}$ GeV. Transmission of supersymmetry breaking to the visible sector is manifest at a lower scale M_{hidden}/M_P of around 1 TeV, i.e. the right superpartner mass scale.

Supergravity is a nice and inspiring idea, which attempts to incorporate gravity. However, it does not make gravity consistent at the quantum level, supergravity is neither finite nor renormalizable, so it does not provide an ultraviolet completion of Einstein theory.

Extra dimensions

There are many scenarios which propose that spacetime has more than four dimensions, the addibional ones being unobservable because they are compact and of very small size. We briefly mention two ideas, which differ by whether the usual Standard Model matter is able to propagate in the new dimensions or not. Again, mixed scenarios are often very popular and interesting.

• **Kaluza-Klein idea**

Kaluza-Klein theories propose the appearance of four-dimensional gauge bosons as components of the metric tensor in a higher-dimensional spacetime. The prototypical example is provided by considering a 5d spacetime with topology $M_4 \times S^1$ and endowed with a 5d metric G_{MN}, $M, N = 1, \ldots, 5$. From the viewpoint of the low-energy four-dimensional theory (at energies much lower than the compactification scale $M_c = 1/R$, with R the circle radius) he 5d metric decomposes as

$$\begin{array}{llll} G_{MN} \to & G_{\mu\nu} \quad \mu,\nu = 0,\ldots,3 & G_{\mu\nu} & \text{4d graviton} \\ & G_{\mu 4} & A_\mu & \text{4d gauge boson} \\ & G_{44} & \phi & \text{4d scalar (modulus)} \end{array} \qquad (1.12)$$

We obtain a 4d metric tensor, a 4d massless vector boson and a 4d massless scalar. Moreover, diffeomorphism invariance in the fifth dimension implies gauge invariance of the interactions of the 4d vector boson (so it is a $U(1)$ gauge boson).

The idea generalizes to d extra dimensions. Take (4+d)-dimensional spacetime of the form $M_4 \times X_d$. The metric in $(4 + d)$ dimensions gives rise to a 4d metric and to gauge bosons associated to a gauge group which is the isometry group of X_d. Specifically, let k_a^M be a set of Killing vectors in X_d; the 4d gauge bosons are obtained as $A_\mu^a = G_{\mu N} k_a^N$.

The Kaluza-Klein idea is beautiful, but it is difficult to use for phenomenology. It is not easy to construct manifolds with isometry group that of the Standard Model. Moreover, a generic difficulty first pointed out by

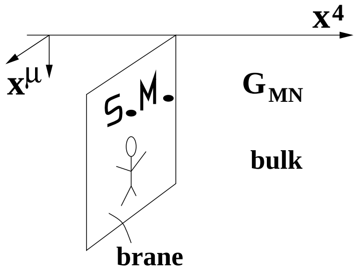


Figure 1.3: Schematic picture of the brane-world idea.

Witten (see [16]) is how to obtain chiral 4d fermions in this setup. For this to be possible one needs to include elementary gauge fields already in the higher-dimensional theory, so much of the beauty of the idea is lost.

On top of that, although the idea involves gravity, it still suffers from quantum inconsistencies, so it does not provide an ultraviolet completion of Einstein theory, consistent at the quantum level.

• **Brane-world idea**

This is a recent proposal (see e.g. [106]), building on the idea of extra dimensions, but with an interesting new ingredient. It is based on the observation that it is conceivable that extra dimensions exist, but that the Standard Model fields do not propagate on them, and that only gravity does. In modern jargon, the Standard Model is said to live on a 'brane' (generalization of a membrane embedded in a higher dimensional spacetime), while gravity propagates in the 'bulk' of spacetime.

In such a scenario, Standard Model physics is four-dimensional up to energies around the TeV, even if the extra dimensions have sizes larger than $(\text{TeV})^{-1}$. The best experiments able to probe the extra dimensions are measurements of deviations from four-dimensional Newton's law in Cavendish experiments, to put a bound at the length scale at which gravity starts being five- or higher-dimensional. The present bound implies that extra dimensions should be smaller than 0.1 mm. This energy scale is surprisingly small, still we do not detect these extra dimensions.

This scenario allows for an alternative interpretation of the four-dimensional Planck scale. Starting with a fundamental Planck scale M_d in the $(4+d)$ dimensional theory, the 4d Planck scale is

$$M_P^2 = (M_d)^{d+2} V_{X_d} \tag{1.13}$$

where V_{X_d} is the volume of the internal manifold. The scenario allows for a low value of the fundamental $(4+d)$ Planck scale, keeping a large 4d M_P by taking a large volume compactification. In usual Kaluza-Klein, such large volumes would imply light Kaluza-Klein excitation of Standard Model fields, in conflict with experiment. In the brane-world scenario, such fields do not propagate in the bulk so they do not have Kaluza-Klein replicas. In certain models, it is possible to set $M_{4+d} \simeq$ TeV, obtaining $M_P \simeq 10^{19}$ GeV as a derived quantity, due to a choice of large volume for the internal manifold. Is is therefore a possible alternative explanation for the hierarchy between M_W and M_P.

Again, it is fair to emphasize that this setup does not provide a ultraviolet completion of Einstain gravity, gravity is treated classically. Moreover, it is not clear to start with that a quantum field theory on a slice of full spacetime can be consistently defined at the quantum level.

1.1.4 String theory as a theory beyond the Standard Model

String theory is also a proposal for physics beyond the Standard Model. It differs from the above in that it addresses precisely the toughest of all issues: it provides a quantum mechanically well-defined theory underlying gauge and gravitational interactions. Hence it provides an ultraviolet completion of Einstein theory, which is finite order by order in perturbation theory. Einstein theory is recovered as a low-energy effective theory for energies below a typical scale, the string scale M_s. That is the beautiful feature of string theory.

Moreover, string theory incorporates gauge interactions, and is able to lead to four-dimensional theories with chiral fermions. In addition, string theory incorporates many of the ingredients of the previous proposals beyond the standard model, now embedded in a consistent and well-defined framework, and leading to physical theories very similar to the Standard Model at energies below a typical scale of the theory (the string scale M_s).

Finally, string theory contains physical phenomena which are new and quite different from expectations from other proposals beyond the standard model. As a theory of quantum gravity, it has the potential to give us some insight into questions like the nature of spacetime, the black hole information paradox. As a theory underlying gauge interactions, it has the potential to explain what is the origin of the number of families in theories like the Standard Model, how do chiral fermions arise, etc...

String theory is an extremely rich structure, from the mathematical, theoretical and phenomenological viewpoints. It is certainly worth being studied in a graduate course in high energy physics!

Chapter 2

Overview of string theory in perturbation theory

To be honest, we still do not have a complete description of string theory at the non-perturbative level (this will become clear in coming lectures). Still, the perturbative picture is very complete, and is the best starting point to study the theory.

2.1 Basic ideas

2.1.1 What are strings?

String theory proposes that elementary particles are *not* pointlike, but rather they are small 1-dimensional extended objects (strings), of typical size $L_s = 1/M_s$. They can be open or closed strings, as shown in figure 2.1. At energies well below the string scale M_s, there is not enough resolution to see the spatial extension of the objects, so they look like point particles, and usual point particle physics should be recovered as an effective description.

Experimentally, our description of elementary particles as pointlike works nicely up to energies or order 1 TeV, so $M_s >$ TeV. In many string models, however, the string scale turns out to be related to the 4d Planck scale, so we have $M_s \simeq 10^{18}$ GeV. This corresponds to string of typical size of 10^{-33} cm, really tiny.

Strings can vibrate. Different oscillation modes of a unique kind of underlying object, the string, are observed as different particles, with differ-

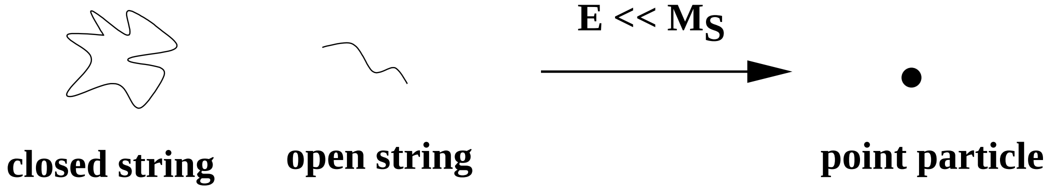


Figure 2.1: According to string theory, elementary particles are 1-dimensional extended objects (strings).

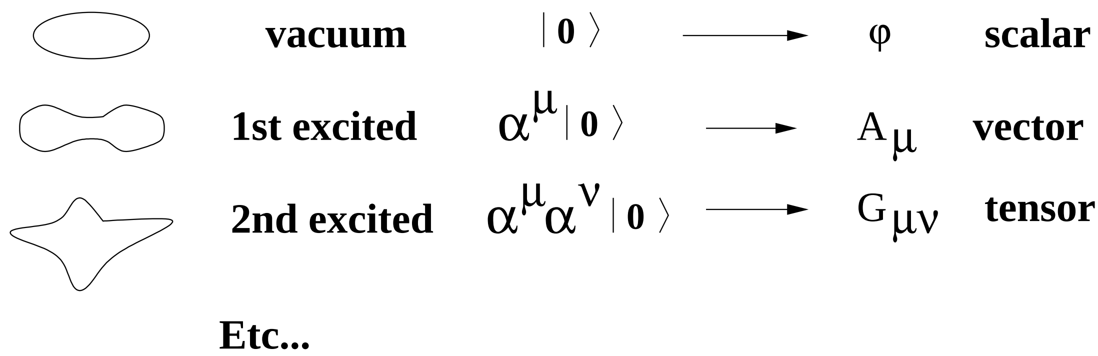


Figure 2.2: Different oscillation modes of unique type of string correspond to different kinds of particles, with e.g. different Lorentz quantum numbers.

ent Lorentz (and gauge and global) symmetry quantum numbers. This is schematically shown in figure 2.2 for closed string states.

The mass of the corresponding particle increases with the number of oscillator modes that we are exciting. So the vibration modes of the string give rise to an infinite tower of particles, with masses increasing in steps of order M_s. Since M_s is so large, only the particles with masses of order zero (to leading order) can correspond to the observed ones.

Upon explicit computation of this spectrum of particles, the massless sector always contains a 2-index symmetric tensor $G_{\mu\nu}$. Later on we will see that this field behaves as a graviton, so string theories automatically contain gravity. But before we can explain interactions in string theory we need some further ingredients.

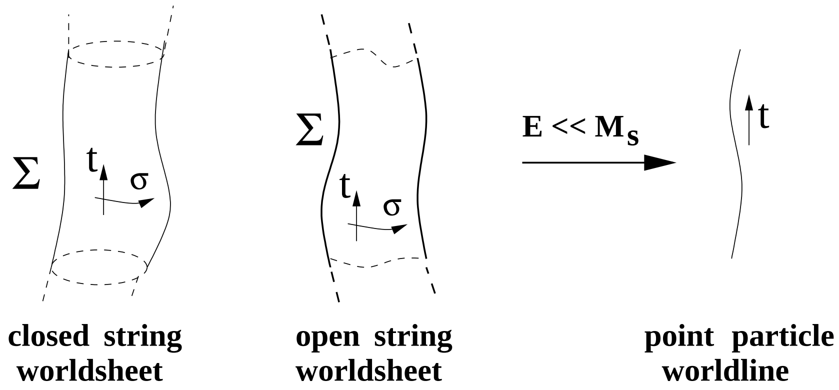


Figure 2.3: Worldsheets for closed and open strings. They reduce to worldlines in the point particle (low energies) limit.

2.1.2 The worldsheet

As a string evolves in time, it sweeps out a two-dimensional surface in spacetime Σ, known as the worldsheet, and which is the analog of the worldline of a point particle in spacetime. Closed string correspond to worldsheets with no boundary, while open string sweep out worldsheets with boundaries. Any point in the worldsheet is labeled by two coordinates, t the 'time' coordinate just as for the point particle worldline, and σ, which parametrizes the extended spatial dimension of the string at fixed t.

A classical string configuration in d-dimensional Minkowski space M_d is given by a set of functions $X^\mu(\sigma, t)$ with $\mu = 0, \ldots, d-1$, which specify the coordinates in M_d of the point corresponding to the string worldsheet point (σ, t).

This can be expressed by saying that the functions $X^\mu(\sigma, t)$ provide a map from a two-dimensional surface (the abstract worldsheet), parametrized by (σ, t) to a d-dimensional space M_d (spacetime, also known as target space of the embedding functions).

$$\begin{array}{rccc} X^\mu : & \Sigma & \to & M_d \\ & (\sigma, t) & \to & X^\mu(\sigma, t) \end{array} \tag{2.1}$$

This is pictorially shown in figure 2.4.

A natural definition for the classical action for a string configuration is given by the total area spanned by the worldsheet (in analogy with the

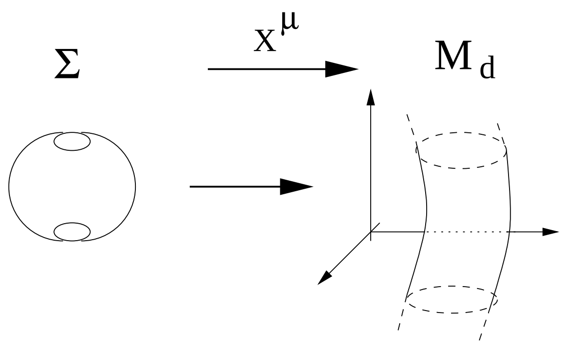


Figure 2.4: The functions $X^\mu(\sigma, t)$ define a map, an embedding, of a 2-dimensional surface into the target space M_d.

worldline interval length as action for a point particle).

$$S_{\rm NG} \ = \ -T \int_\Sigma dA \tag{2.2}$$

where T is the string tension, related to M_s by $T = M_s^2$. One also often introduces the quantity α', with dimensions of length squared, defined by $T = M_s^2 = \frac{1}{2\pi\alpha'}$.

In terms of the embedding functions $X^\mu(\sigma, t)$, the action (2.2) can be written as

$$S_{\rm NG} \ = \ -T \int_\Sigma (\, \partial_\tau X^\mu \, \partial_\tau X_\mu \ - \ \partial_\sigma X^\mu \, \partial_\sigma X_\mu \,)^{1/2} \, d\sigma \, dt \tag{2.3}$$

This is the so-called Nambu-Goto action. It is difficult to quantize, so quantization is simpler if carried out starting with a different, but classically equivalent action, known as the Polyakov action

$$S_{\rm Polyakov} \ = \ -T/2 \int_\Sigma \sqrt{-g} \, g^{\alpha\beta}(\sigma, t) \, \partial_\alpha X^\mu \, \partial_\beta X^\nu \eta_{\mu\nu} \, d\sigma \, dt \tag{2.4}$$

where we have introduced an additional function $g(\sigma, t)$. It does not have interpretation as an embedding. The most geometrical interpretation it receives is that it is a metric in the abstract worldsheet Σ. At this point it is useful to imagine the worldsheet as an abstract two-dimensional world which is embedded in physical spacetime M_d via the functions X^μ. But which to some extent makes sense by itself.

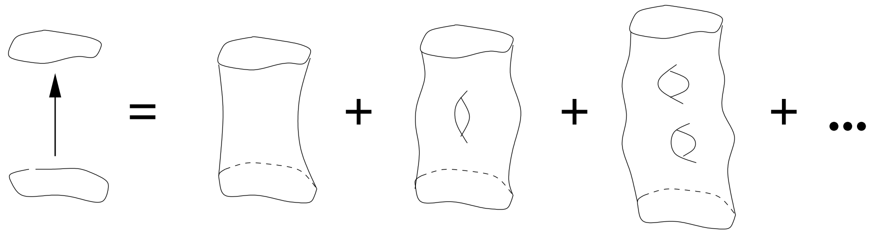


Figure 2.5: The genus expansion for closed string theories .

The important fact we would like to emphasize is that this looks like the action for a two-dimensional field theory coupled to two-dimensional gravity. Many of the wonderful properties of string theory arise from subtle relation between the 'physics' of this two-dimensional world and the physics of spacetime.

The two-dimensional field theory has a lot of gauge and global symmetries, which will be studied later on. For the moment let us simply say that after fixing the gauge the 2d action becomes

$$S_P[X(\sigma,t)] \;=\; -T/2\int_\Sigma \partial_\alpha X^i\,\partial_\alpha X^i, \qquad i=2,\ldots,d-1 \qquad (2.5)$$

It is just a two-dimensional quantum field theory of $d-2$ free scalar fields. This is easy to quantize, and gives just a bunch of decoupled harmonic oscillators, which are the string oscillation modes mentioned before. It is important to notice that the fact that the worldsheet theory is a free theory does *not* imply that there are no interactions between strings in spacetime. There are interactions, as we discuss in the following.

Before concluding, let us emphasize a crucial property of the worldsheet field theory, its conformal invariance. This property is at the heart of the finiteness of string theory, as we discuss below.

2.1.3 String interactions

A nice discussion is in section 3.1. of [55]

The quantum amplitudes between string configurations are obtained by performing a path integral, namely summing over all possible worldsheets which interpolate between the configurations, see figures 2.5, 2.6.

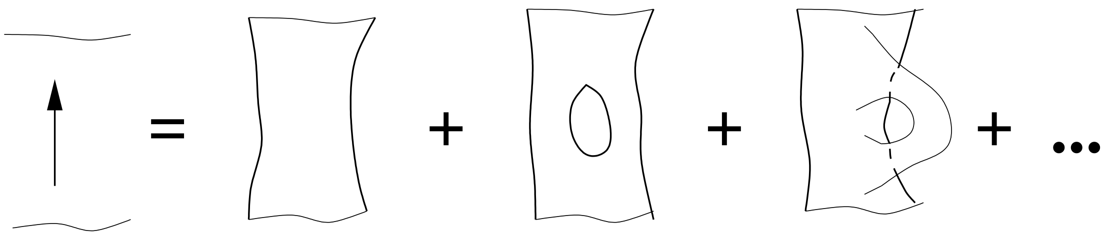


Figure 2.6: The genus expansion for theories with open strings. Notice that one must include handles and boundaries .

The sum organizes into a sum over worldsheet topologies, with increasing number of handles and of boundaries (for theories with open strings) This is the so-called genus expansion (the genus of a closed Riemann surface is the number of handles. In general it is more useful to classify 2d surfaces (possibly with boundaries) by their Euler number, defined by $\xi = 2-2g-n_b$, with g and n_b the numbers of handles and boundaries, respectively).

Formally, the amplitude is given by

$$\langle b|\text{evolution}|a\rangle \;=\; \sum_{\text{worldsheets}} \int [\mathcal{D}X]\, e^{-S_P[X]}\, \mathcal{O}_a[X]\, \mathcal{O}_b[X] \qquad (2.6)$$

where $\mathcal{O}_i[X]$ are the so-called vertex operators, which put in the information about the incoming and outgoing state. They are very important in tring theory and conformal field theory but we will not discuss them much in these lectures.

Notice that the quantity (2.6) is basically a quantum correlation function between two operators in the 2d field theory. However, notice the striking fact that (2.6) is in fact a sum of such correlators for 2d field theories living in 2d spaces with different topologies. Certainly it is a strange prescription, a strange quantity, in the language of 2d field theory. However, it is the prescription that arises naturally from the spacetime point of view.

The basic string interaction processes and their strengths are shown in figure 2.7. It is important to notice that these vertices are delocalized in a spacetime region of typical size L_s. At low energies $E \ll M_s$ they reduce to usual point particle interaction vertices.

There is also one vertex, shown in figure 2.8. It couples two open strings with one closed string. It is important to notice that the process that turns

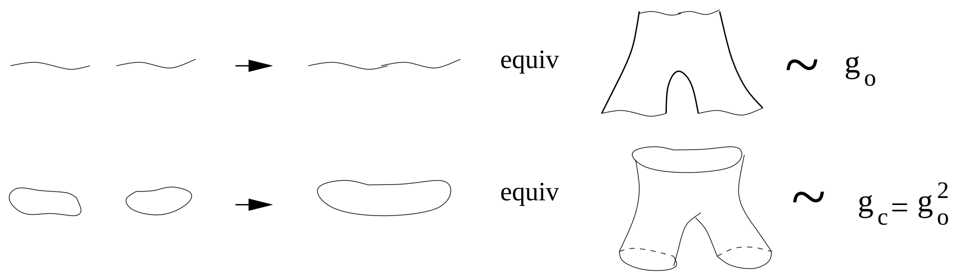


Figure 2.7: Basic interaction vertices in string theory.

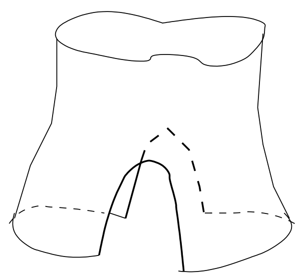


Figure 2.8: String vertex coupling open strings to closed strings. It implies that theories with open strings necessarily contain closed strings.

the closed strings into a closed one corresponds locally on the worldsheet exactly to joining two open string endpoints (twice). This coupling cannot be forbidden in a theory of interacting open strings (since this process also mediates the coupling of three open strings), so it implies that any theory of interacting open strings necessarily contains closed strings. (The reverse statement is not valid, it is possible to have interacting theories of closed strings without open strings).

A fundamental property of string theory is that the amplitudes of the theory are finite order by order in perturbation theory. This, along with other nice properties of string interactions (like unitarity, etc) implies that string theory provides a theory which is consistent at the quantum level, it is well defined in the ultraviolet. There are several ways to understand why string theory if free from the ultraviolet divergences of quantum field theory:

a) In quantum field theory, ultraviolet divergences occur when two interaction vertices coincide at the same point in spacetime. In string theory,

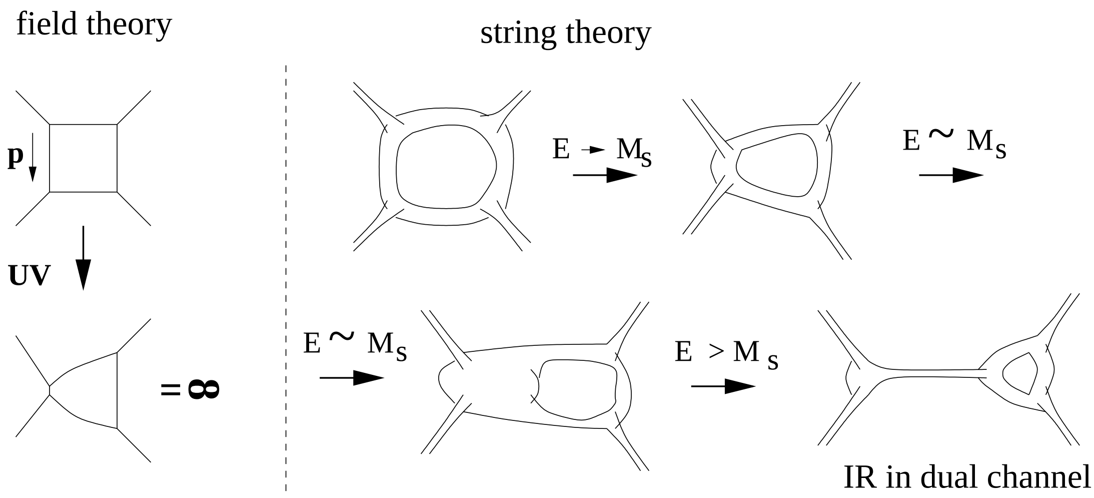


Figure 2.9: Different ultraviolet behaviours in quantum field theory and in string theory. When high energy modes exchanged in the loop reach energies of order M_s, long strings start being exchanged and dominate the amplitude. So at those energies the behaviour differs from the quantum field theory divergence, which is effectively cut-off by M_s. The ultra-high energy regime corresponds to exchange of very long strings, which can be interpreted as the infrared regime of a 'dual channel diagram'. .

vertices are delocalized in a region of size L_s, so L_s acts as a cutoff for the would-be divergences.

b) As is pictorially shown in figure 2.9, going to very high energies in some loop, the ultraviolet behaviour starts differing from the quantum field theory behaviour as soon as energies of order M_s are reached. This is so because longer and longer string states start being exchanged, and this leads to a limit which corresponds not to a ultraviolet divergence, but to an infrared limit in a dual channel.

c) More formally, using conformal invariance on the worldsheet, any limit in which a string diagram contains coincident or very close interaction vertices can be mapped to a diagram with well-separated vertices and an infinitely long dual channel. This is a formalization of the above pictorial argument.

Using the above rules for amplitudes, it is possible to compute interactions between the massless oscillation modes of string theory. These interactions turn out to be invariant under gauge and diffeomorphism transformations for spacetime fields. This means that the massless 2-index tensor $G_{\mu\nu}$ contains only two physical polarization states, and that it indeed interacts as a graviton. Also, massless vector bosons A_μ have only two physical polarizations, and interact exactly as gauge bosons. We will not discuss these issues in the present lectures, but a good description can be found in [9] or [55].

Hence, string theory provides a unified description of gauge and gravitational interactions, which is consistent at the quantum level. It provides a unified ultraviolet completion for these theories. This is why we love string theory!

2.1.4 Critical dimension

Conformal invariance in the 2d worldsheet theory is a crucial property for the consistency of the theory. However, this symmetry of the classical 2d field theory on the worldsheet may in principle not be preserved in the 2d quantum field theory, it may suffer what is called an anomaly (a classical symmetry which is not preserved at the quantum level), see discussion in chapter 3 of [55].

As is usual in quantum field theories with potential anomalies, the anomaly disappear for very specific choices of the field content of the theory. In the case of the conformal anomaly of the 2d worldsheet field theory, the field content is given by d bosonic fields, the fields $X^\mu(\sigma, t)$. In order to cancel the conformal anomaly, it is possible to show that the number of fields in the 2d theory must be 26 bosonic fields, so this is the number of X^μ fields that we need to consider to have a consistent string theory.

Notice that this is very striking, because the number of fields X^μ is precisely the number of spacetime dimensions where the string propagates. The self-consistency of the theory forces us to admit that the spacetime for this string theory has 26 dimensions. This is the first situation where we see that properties of spacetime are constrained from properties of the worldsheet theory. In a sense, in string perturbation theory the worldsheet theory is more fundamental than physical spacetime, the latter being a derived concept.

Finally, let us point out that there exist other string theories where the worldsheet theory contains other fields which are not just bosons (superstring theories, to be studied later on). In those theories the anomaly is different and the number of spacetime dimensions is fixed to be 10.

2.1.5 Overview of closed bosonic string theory

In this section we review the low-lying states of the bosonic string theory introduced above (defined by 26 bosonic degrees of freedom in the worldsheet, with Polyakov action), and their interactions.

The lightest states in the theory are

- the string goundstate, which is a spacetime scalar field $T(X)$, with tachyonic mass $\alpha' M^2 = -2$. This tachyon indicates that bosonic string theory is unstable, it is sitting at the top of some potential. The theory will tend to generate a vacuum expectation value for this tachyon field and roll down the slope of the potential. It is not know whether there is a minimum for this potential or not; if there is, it is not know what kind of theory corresponds to the configuration at the potential minimum. The theories we will center on in later lectures, superstrings, do not have such tachyonic fields, so they are under better control.

- a two-index tensor field, which can be decomposed in its symmetric (traceless) part, its antisymmetric part, and its trace. All these fields are massless, and correspond to a 26d graviton $G_{MN}(X)$, a 26d 2-form $B_{MN}(X)$ and a 26d massless scalar $\phi(X)$, known as the dilaton. These fields are also present in other string theories.

Forgetting the tachyon for the moment, it is possible to compute scattering amplitudes. It is possible to define a spacetime action for these fields, whose tree-level amplitudes reproduce the string theory amplitudes in the low energy limit $E \ll M_s$, usually denoted point particle limit or $\alpha' \to 0$. This action should therefore be regarded as an effective action for the dynamics of the theory at energies below M_s. Clearly, the theory has a cutoff M_s where the effective theory ceases to be a good approximation. At that scale, full-fledged string theory takes over and softens the UV behaviour of the effective field theory.

The spacetime effective theory for the string massless modes is

$$S_{\text{eff.}} = \frac{1}{2k_0^2} \int d^{26}X\, (-G)^{1/2}\, e^{-2\phi}\, \{\, R - \frac{1}{12}\, H_{MNP}\, H^{MNP} \, + \, 4\partial_M\phi\partial^M\phi\, \} \, + \mathcal{O}(\alpha') \tag{2.7}$$

where $M, N, P = 0, \ldots, 25$, and $H_{MNP} = \partial_{[M} B_{NP]}$. Notice that very remarkably this effective action is invariant under coordinate transformations in 26d, and under the gauge invariance (with 1-form gauge parameter $\Lambda_M(X)$)

$$B_{MN}(X) \to B_{MN}(X) + \partial_{[M}\Lambda_{N]}(X) \tag{2.8}$$

(which in the language of differential forms reads $B \to B + d\Lambda$).

Notice that the coupling constant of the theory k_0 can be changed if the scalar field ϕ acquires a vacuum expectation value ϕ_0. Hence, the spacetime string coupling strength (the g_c in our interaction vertices) is not an arbitrary external parameter, but it is a vacuum expection value for a dynamical spacetime field of the theory. In many other situations, string models contain this kind of 'parameters' which are actually not external parameters, but vevs for dynamical fields of the theory. This is the familiar statement that string theory does not contain external dimensionless parameters.

These fields, like the dilaton and others, are known as moduli, and typically have no potential in their effective action (so they can take any vev, in principle). This also leads to phenomenological problems, because we do not observe such kind of massless scalars in the real world, whereas they are ubiquitous in string theory.

The above action is said to be written in the string frame (which means that the field variables we are using are those naturally associated with the vertex operators one constructs from the 2d conformal field theory viewpoint. From the specetime viewpoint, it is most convenient to redefine the fields as

$$\tilde{G} = e^{\phi_0 - \phi} \quad ; \quad \tilde{\phi} = \phi - \phi_0 \tag{2.9}$$

to obtain the action

$$S_{\text{eff.}} = \frac{1}{2k^2} \int d^{26}X \, (-\tilde{G})^{1/2} \{ \tilde{R} + \frac{1}{12} e^{-\tilde{\phi}/12} \, H_{MNP} \, H^{MNP} - \frac{1}{6} \partial_M \tilde{\phi} \partial^M \tilde{\phi} \} + \mathcal{O}(\alpha') \tag{2.10}$$

with indices raised by $\tilde{G}$. This action is said to be written in the Einstein frame, because it contains the gravity action in the canoncial Einstein form. Notice that the change between frames is just a relabeling of fields, not a coordinate change or anything like that.

So we have obtained an effective action which reduces basically to Einstein gravity (plus some additional fields). The 26d Planck mass is given by $M_{26d}^{24} = M_s^{24}/g_c^2$. This effective theory is not renormalizable, and is valid only up to energies M_s, which is the physical cutoff of the effective theory; there is however an underlying theory which is well defined at the quantum level, valid at all energies (UV finite) and which reduces to the effective theory below M_s. String theory has succeeded in providing a consistent UV completion of Einstein theory.

It is also important to point out that this version of quantum gravity is also consistent with gauge invariance, for instance with the gauge invariance of the 2-form fields. Other string theories (with open strings, or some superstrings) also contain vector gauge bosons, with effective action given by Yang-Mills. So the theory contains gauge and gravitational interactions consistently at the quantum level.

Let us conclude by mentioning some of the not-so-nice features of the theory at hand.

- First, it lives in 26 dimensions. We will solve this issue in subsequent lectures by the process known as compactifications

- The theory does not contain fermions. This will be solved by introducing a more interesting kind of string theory (by modifying the worldsheet field content), the superstrings. These theories still live in 10 dimensions so they need to be compactified as well

- The theory does not contatin non-abelian vector gauge bosons. Such gauge bosons are however present in some superstring theories (heterotic and type I, and in type II theories in the presence of topological defects).

- Other questions which remain unsolved (like supersymmetry and supersymmetry breaking, or the moduli and vacuum degeneracy problems) will also appear along the way.

One issue that can be addressed at this point is to obtain four-dimensional physics (at low energies) from a theory originally with more dimensions. The standard technique to do so is known as compactification, and can be applied not only to reduce the closed bosonic string theory to four-dimensions, but also to other more interesting string theories. For this reason, it is interesting to study compactification right now. However, before that, we need to take a small detour and learn how to formulate string theory in spacetimes more complicated than Minkowski space.

2.1.6 String theory in curved spaces

See for instance sect. 3.7 in [55].

We have obtained an effective action for the low-lying modes of string theory. In principle, configurations of these fields which satisfy the corresponding (classical) equations of motion should correspond to classical backgrounds where strings can propagate.

However, the worldsheet description we provided is only valid when the background is trivial (26d Minkowski space). It is a natural question to ask how the worldsheet theory is modified so that it describes propagation of a string in a spacetime with non-trivial metric $G_{MN}(X)$, and non-trivial background for the two-index antisymmetric tensor field $B_{MN}(X)$ and the dilaton ϕ.

The effect of the metric is relatively simple: The string action is still the worldsheet area, now computed using the new metric in spacetime. Using the Polyakov version of the worldsheet action, eq (9.1) generalizes to

$$S_P^G\,[X(\sigma,t),g(\sigma,t)] \;=\; \frac{1}{4\pi\alpha'}\int_\Sigma d\sigma\,dt\,\sqrt{-g}G_{MN}[X(\sigma,t]]\,g^{\alpha,\beta}\,\partial_\alpha X^M(\sigma,t)\,\partial_\beta X^N(\sigma,t) \quad (2.11)$$

Where $G_{MN}(X)$ is a function(al) of $X(\sigma,t)$. This action is also known as non-linear sigma model, for historical reasons not to be discussed here.

One may wonder about the double role played that the spacetime graviton in string theory. On one hand, we have claimed that the graviton arises as one of the states in the string spectrum in flat space. On the other, a background metric, made out of gravitons, appears explicitly in the worldsheet action of a string propagating in curved space. (This issue is related to the discussion on how to split a field configuration as a background plus a fluctuation around it.)

This dicotomy can be understood in detail for metrics which are small perturbations of flat space metric

$$G_{MN} = \eta_{MN} + \delta G_{MN} \quad (2.12)$$

Replacing this into the worldsheet action (2.11), we obtain an expansion around the flat space action. In a path integral, expanding the exponential as well one gets that amplitudes in curved space can be regarded as amplitudes in flat space with corrections due to graviton insertions

$$\begin{aligned}\int [\mathcal{D}X]\, e^{-S_P^G} \;=\;& \textstyle\int [\mathcal{D}X]\, e^{-S_P^\eta} + \int [\mathcal{D}X]\, e^{-S_P^\eta}\,\mathcal{O}_\mathcal{G}[\mathcal{X}] \,+ \\ & \textstyle\int [\mathcal{D}X]\, e^{-S_P^\eta}\,\mathcal{O}_\mathcal{G}[\mathcal{X}]\,\mathcal{O}_\mathcal{G}[\mathcal{X}] + \ldots\end{aligned} \quad (2.13)$$

where $\mathcal{O}_\mathcal{G}[\mathcal{X}]$ is the vertex operator for the graviton, as a state in the string spectrum. Recalling that a path integral with a vertex operator insertion corresponds to addint an external leg, the situation is pictorially shown in figure 2.10

Even for metrics which cannot be regarded as deformations of flat space (for instance, if the corresponding manifolds are topologically different from flat space), then (2.11) is the natural prescription.

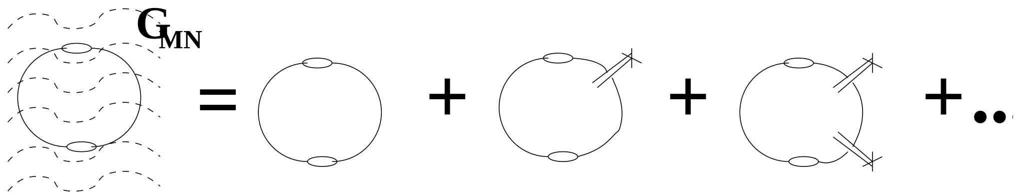


Figure 2.10: Amplitudes in curved space can be regarded as a resummation of amplitudes of amplitudes in flat space, with increasing number of graviton insertions. Hence the curved background can be regarded as built out of gravitons, in quite an explicit way.

Since there are also other massless fields in the spectrum of the string, it is natural to couple them to the worldsheet, so as to obtain a worldsheet action for strings propagating on non-trivial backgrounds. The resulting action is

$$S_P^G\,[X(\sigma,t),g(\sigma,t)] = \; \tfrac{1}{4\pi\alpha'}\int_\Sigma d\sigma\,dt\,\sqrt{-g}\,[\,G_{MN}[X(\sigma,t]\,g^{\alpha\beta}\,\partial_\alpha X^M(\sigma,t)\,\partial_\beta X^N(\sigma,t)\;+ \\ +\,B_{MN}[X(\sigma,t]]\,\epsilon^{\alpha,\beta}\,\partial_\alpha X^M(\sigma,t)\,\partial_\beta X^N(\sigma,t)\;+\;\alpha' R[g]\phi\,] \quad (2.14)$$

It satisfies the criterion that for backgrounds near the trivial one it expands as resummation over insertions of the corresponding vertex operators. Moreover, the different terms have a nice interpretation also in the form (2.14).

• We have already explained that the piece depending on G_{MN} is simply the area of the worldsheet as measured with the curved spacetime metric. That is, the natural generalization of the Nambu-Goto idea.

• The term that depends on B_{MN} is exactly the result of interpreting the two-index tensor as a 2-form $B_2 = B_{MN}dX^M \wedge dX^N$ in spacetime, and integrating it over the 2-dimensional surface given by the world-sheet. In the language of differential forms

$$S_B \;=\; \frac{1}{4\pi\alpha'}\int_\Sigma B_2 \qquad (2.15)$$

Notice that the term is purely topological in spacetime, it does not depend on the spacetime metric.

The physical interpretation of this term is that strings are charged objects with respect to B_2, when the latter is regarded as a gauge potential (recall the gauge invariance $B_2 \to B_2 + d\Lambda$). It is the analog of the minimal coupling of a point particle to a vector gauge potential A_1, given by integrating A_1 over the particle worldline).

• The term that depends on ϕ is very special. In principle it corresponds to an Einstein term for the 2-dimensional worldsheet metric $g_{\alpha\beta}(\sigma, t)$. However, 2d gravity is very special, is almost topological. This means that in 2 dimensions, the integral of the curvature scalar over a surface is, by Gauss theorem, just a number, determined by the topology of the surface. This number is simply the Euler number of the surface, given by

$$\xi = 2 - 2g - n_b \tag{2.16}$$

where g is the number of handles and n_b is the number of boundaries.

Insertion of this term in an amplitude corresponds exactly to weighting it by a factor $e^{-\phi\xi}$. It is possible to check that the power of $e^{-2\phi}$ appearing in the amplitude for a given diagram (worldsheet topology) is exactly the same power as for the closed string coupling g_c (in theories with open strings, the same is true for powers of $e^{-\phi}$ and of the open string coupling g_o (recall $g_c = g_o^2$). This is an alternative way of rediscovering that the vev for the dilaton plays the role of the string coupling constant.

Again we see that string theory does not contain external adimensional parameters. All parameters are in fact vevs for dynamical fields.

It is important to realize that in the presence of non-trivial backgrounds the worldsheet action, regarded as a 2d field theory, is no longer a free field theory. From this viewpoint, it is natural to study it in perturbation theory around the free theory. The expansion parameter is α'/R^2, where R is the typical length scale of variation of any spacetime field, so this is known as the α' expansion.

It is important to realize that string theory in a general background has therefore a double expansion. First, there is the loop expansion in the string coupling constant, which corresponds to the genus expansion summing over worldsheet topologies. Second, for any given worldsheet topology, the computation of the path integral over the (interacting) 2d field theory is done as a loop expansion in the 2d world, the α' expansion.

Both expansion are typically very involved, and most results are known at one loop in either expansion. The issue of the α' expansion makes it very difficult to use string theory in regimes where very large curvatures of spacetime are present, like black hole or big-bang singularities.

This is a bit unfortunate, because α' mainly encode effects which encode the fact that the fundamental object in string theory is an extended object, rather than a point particle. For instance, the geometry seen by string theory, at scales around L_s, is different from the geometry a point particle would see. This new notion of geometry (which is still vague in many formal respects) is called stringy geometry (or quantum geometry, by B. Greene, because it corresponds to taking into account loops in α', in the 2d quantum field theory).

Happily, there still exist some simple enough situations where α' effects are tractable, and can be seen to be spectacular. For instance, the fact the complete equivalence of string theory on two different spacetime geometries, once stringy effects are taken into account (T-duality).

We conclude with an important issue. We have emphasized the importance of conformal invariance of the 2d worldsheet field theory in order to have a consistent string theory (with finite amplitudes, etc). Therefore, the interacting 2d field theory given by (2.14) should correspond to a conformal field theory. In general, this can be checked only order by order in the α' expansion, and in practice the results are known at leading order (one loop in α'). In perturbation theory in the 2d field theory, conformal invariance means that the (one-loop in α' beta functions for the coupling constants in the 2d field theory lagrangian) vanish.

Notice that in a sense, the background fields play the role of these coupling constants. The condition that their beta function equals zero amounts to the constraint that the background fields obey some differential equation. The amazing thing is that these differential equations are exactly the equations of motion that one obtains from the spacetime effective action for the spacetime fields (6.13). That is, string propagation is consistent (2d action is conformal field theory) exactly in background which obey the equations of motion from the spacetime effective action (derived from scattering amplitudes, etc, i.e. from a different method). I regard this as an amazing self-consistency property of string theory.

It should be pointed out that these statements remain valid for string theories beyond the closed bosonic theory we are studying for the moment.

A final comment concerns an alternative approach to study string theory beyond flat space. A whole lot is known about two-dimensional field theory which are exactly conformal [10]. Some of them can be solved exactly, namely one can give expression for any 2d correlator, exactly i.e. to all orders in the 2d loop expansion. One can then imagine using conformal field theories of this kind (so-called exactly solvable conformal field theories) to describe the string worldsheet. The question is then to identify what is the background where the string is propagating. In several cases this can be done and corresponds to very exotic possibilities, for instance Witten's black hole, compact curves spaces with very small size (or order the string length, etc. The importance of these models is that by construction all α' effects are included. Another motivation is that in this language it is clear that spacetime is in a sense a derived concept in string theory, and that the worldsheet theory is more fundamental (this view was dominant before '95, and is perhaps slightly changed nowadays; still it has a point).

2.1.7 Compactification

In this section we study a special and very important class of backgrounds, which lead, in the low-energy limit, to effective theories with smaller number of dimensions than the original one. We center on constructing models which produce four-dimensional physics, of course (although people often study e.g. six-dimensional models, etc).

The idea is to consider string propagation in a spacetime of the form

$$X_{26} = M_4 \times X_{\mathrm{comp.}} \tag{2.17}$$

where M_4 is 4d Minkowski and X_{comp} is a compact 22-dimensional manifold (with Euclidean signature), called the compactification manifold, or internal space.

The recipe to write the worldsheet action is as above. In general, it corresponds to a nonlinear sigma model, an interacting theory, and we can study it only in the α' expansion (and often at leading order). From the spacetime viewpoint, this means that we study the point particle limit, we use the effective field theory (6.13), which is basically 26d Einstein theory (plus other fields in this background). This approximation is good as long as the typical size of the compactification manifold is larger than the string scale. In this regime, our theory looks a standard Kaluza-Klein theory.

In very special cases (mainly when the compactification manifold is a torus) the sigma model reduces to a free field theory, which is solved exactly (in the sense of the α' expansion). In such cases, the theory can be studied reliably even for small sizes of the compactification manifold. When these

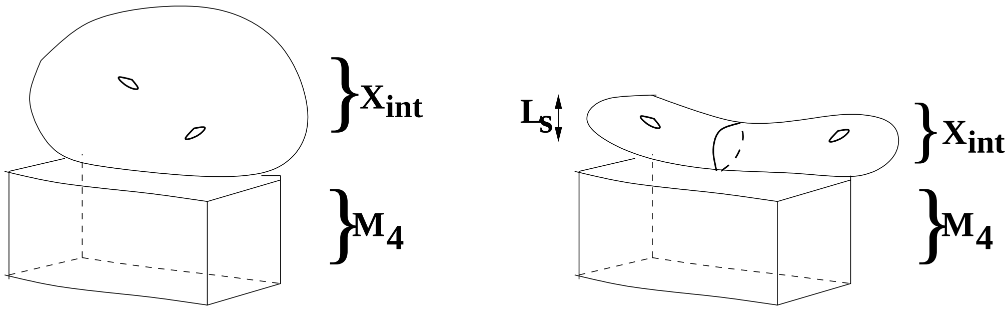


Figure 2.11: Picture of compactification spacetimes; thick small lines represent string states which are light in the corresponding configuration. When the internal manifold has size of the order of L_s, stringy effects (which do not exist in theories of point particles) become relevant; for instance, string winding modes (where a closed string winds around some internal dimension) may be light and appear in the low energy spectrum (even if they do not appear, they may modify importantly the low-energy effective action).

sizes are of the order of the string length, stringy effects become spectacular, and there happen things which are unconceivable in a theory of point particle. For instance, a typical stringy effect is having closed strings wrapping around the non-trivial curves in the internal space. For large volumes, these states are hugely massive, and do not affect much the low-energy physics. For stringy volumes, such states can be very light (as light as other 'point-particle' like modes, or even massless!) and do change the low-energy physics.

Let us first consider large volume compactifications for the moment (so we work in the effective field theory approach) and explain why the low-energy physics is four-dimensional. Consider first a toy model of a 5d spacetime of the form $X_5 = M_4 \times S^1$, on which a 5d massless scalar field $\varphi(x^0,\ldots,x^4)$ propagates with 5d action

$$S_{5d\varphi} = \int_{M_4\times S^1} d^5x\, \partial_M\varphi\partial^M\varphi \tag{2.18}$$

Since x^4 parametrizes a circle, it is periodic, and we can expand the x^4 dependence in Fourier modes

$$\varphi(x^0,\ldots,x^4) = \sum_{k\in\mathbf{Z}} e^{2\pi i k x^4/L}\,\varphi_k(x^0,\ldots,x^3) \tag{2.19}$$

where L is the lenght of $\S^1$.

From the 4d viewpoint, we see a bunch of 4d scalar fields $\phi_k(x^0,\ldots,x^4)$, labeled by the integer index k, the 5d momentum. The 4d spacetime mass of those fields increases with k^2. To see that, take the 5d mass-shell condition

$$P^2 = 0 \quad \text{that is } P_{4d}^2 + p_5^2 = 0 \tag{2.20}$$

For the field ϕ_k, we have

$$P_{4d}^2 + (k/L)^2 = 0 \tag{2.21}$$

which means that the 4d mass of the field ϕ_k is $m_k^2 = (k/L)^2$

At energies much lower than the compactification scale $M_c = 1/L$, $E \ll 1/L$, the only mode which is observable is the zero mode $\phi_0(x^0,\ldots,x^3)$. So we see just a single 4d field, with a 4d action, which is obtained by replacing $\phi(x^0,\ldots,x^4)$ in (2.18) by the only component we are able to excite $\phi_0(x^0,\ldots,x^3)$. The x^4 dependence drops and we get

$$S_{eff} = \int_{M_4} d^4x\, L\, \partial_\mu\varphi_0\partial^\mu\varphi_0 \tag{2.22}$$

So we recover 4d physics at energies below M_c. This is the Kaluza-Klein mechanism, or Kaluza-Klein reduction. The massive 4d fields ϕ_k are known as Kaluza-Klein (KK) excitations or KK replicas of ϕ_0.

As explained in the first lecture, the Kaluza-Klein reduction works for any higher dimensional field. An important new feature arises when the original higher dimensionl field has non-trivial Lorentz quantum numbers. The procedure is then to first decompose the representation of the $SO(d)$ higher-dimensional Lorentz group with respect to the 4d one $SO(4)$ (i.e. separate different components according to their behaviour under 4d Lorentz), and finally perform KK reduction for each piece independently. For instance, for a 5d graviton we have the KK reduction

$$\begin{aligned} G_{MN}(x^0,\ldots,x^4) \to\ & G_{\mu\nu}(x^0,\ldots,x^4) \to\ G_{\mu\nu}^{(0)}(x^0,\ldots,x^3) \\ & G_{\mu 4}(x^0,\ldots,x^4) \to\ G_{\mu 4}^{(0)}(x^0,\ldots,x^3) \\ & G_{44}(x^0,\ldots,x^4) \to\ G_{44}^{(0)}(x^0,\ldots,x^3) \end{aligned} \tag{2.23}$$

where the first step is just decomposition in components, and the second is KK reduction. We therefore obtain, at the massless level, a 4d graviton, a 4d

$U(1)$ gauge boson, and a 4d scalar. Recall that diffeormophism invariance in 5d implies gauge invariance of the 4d vector gauge boson. Also notice that the vev for the scalar field is G_{44}, which is related to the length of the internal circle. Therefore, it is not an external parameter, but the vev of a 4d dynamical scalar field. On the other hand, the compactification is consistent (solves the 5d equations of motion) no matter what circle radius we choose; this implies that in the 4d effective action there is no potential for the 4d scalar, it parametrizes what is called a flat direction of the potential, the field is called a modulus (and it is similar to the string theory dilaton in many respects).

Obs: If the higher-dimensional field theory contains massive fields with mass M, the 4d KK tower has masses $m_k^2 = M^2 + (k/L)^2$, so they will not be observable at energies below M.

The lesson learned here is very general, and can be applied to compactification of any theory on any internal manifold, and an arbitrary set of fields. In particular, it can be applied to string theory. Massless 26-dimensional string states will lead to massless 4d fields corresponding to the zero modes in the KK reduction. KK replicas are not visible at energies below M_c. Massive 26-dimensional string states give massive 4d states, with masses at least or order M_s, which is huge, and are not observable at low energies.

Let us skip the details of KK reduction in manifolds more general than tori, and simply say that in general the role played by the momentum k in toroidal directions is played by the eigenvalues of the laplace operator in the internal manifold (which are also quantized in units of $1/L$, where L is the typical length of the internal space).

2.2 Superstrings and Heterotic string phenomenology

2.2.1 Superstrings

Spacetime fermions vs worldsheet fermions

See discussion in sect 10 in [71].

In trying to connect string theory with the kind of physics observed in Nature, we have seen that compactification is able to solve the dimension problem of the bosonic string theory: how to get four-dimensional physics (at least at low energies) out of a theory which must propagate on a 26d spacetime.

A more difficult problem is that bosonic string theory does not contain spacetime fermions in its spectrum, and fermion fields are essential in our understanding of the real world. This (and also the closed string tachyon, etc) is enough to discard bosonic string theory as realized in Nature [1]

Happily there exist other string theories which are not the bosonic string theory. We are now advanced enough to understand that a string theory is basically defined by a 2d conformal field theory which provides the worldsheet action. What we are about to do is to construct a new kind of worldsheet theories, with Poincare invariance in d-dimensional spacetime, and which contain more fields than just the worldsheet scalars $X^\mu(\sigma,t)$. The resulting string theories have a spectrum of spacetime particles different from that in the bosonic string theory, and in particular they will turn out to contain spacetime fermions.

The basic idea is to supersymmetrize the 2d worldsheet theory. That is, we consider a 2d field theory with d worldsheet scalar fields $X^\mu(\sigma,t)$, and d worldsheet fermion fields $\psi^\mu(\sigma,t)$, which are their superpartners. In Polyakovs formulation one also has the worldsheet metric $g_{\alpha\beta}(\sigma,t)$ and now we also introduce its superpartner (which is a worldsheet gravitino). After gauge fixing these will disappear so we will not be very explicit about them.

String theories with worldsheet supersymmetry are called superstrings. They are just string theories with a different structure for the worldsheet action.

It is very important to notice that the 2d fields $\psi^\mu(\sigma,t)$ are fermions on the worldsheet (and so have anticommutation relations, etc in the 2d quantum field theory) but they transform as a vector under the d-dimensional spacetime Lorentz group, and so they behave as spacetime bosons. This makes sense because the Lorentz group is a global symmetry from the worldsheet viewpoint, and it commutes with the worldsheet supersymmetry, so 2d fields in the same supermultiplet must transform in the same way under the global symmetry.

So, the reason why superstrings contain spacetime fermions is not automatically because they contain fermions on the worldsheet. Indeed the connection is much more subtle and we will not study it until the detailed

[1] Leaving aside the speculative possibility that bosonic string theory may contain fermions in its non-perturbative spectrum.

lectures.

Something similar happens with spacetime supersymmetry. The fact that superstrings have worldsheet supersymmetry does not imply that the spacetime spectrum of particles is supersymmetric. Several superstring theories DO have a spectrum of particles which is spacetime supersymmetric, but the way this arises is very subtle and follows from the so-called GSO projection. These are the most studied superstring theories, because they are well-behaved, for instance do not contain tachyons in their spectrum, so are stable. There also exist some superstrings which have a supersymmetric worldsheet theory, but are not supersymmetric in spacetime; very often they contain tachyons in their spectrum, so are not so much in control.

A common feature of all superstrings (and one which distinguishes them from the bosonic theory) is that, since we have modified the content of fields of the 2d worldsheet theory, the conformal anomaly changes, and the constraint on the number of dimension changes. The number of dimensions on which superstrings consistently propagate is 10. As usual, one uses compactification to construct theories with four-dimensional physics at low energies.

The different ten-dimensional superstring theories

Skipping many important details to be studied in coming lectures, here we would like to briefly describe the structure of the five superstring theories, which are also supersymmetric in spacetime (have a supersymmetric spectrum of spacetime particles).

For references on the structure of susy and sugra multiplets, see [12].

• **Type IIA superstring**

This is a theory of closed oriented strings.

Type IIA string theory has $N = 2$ (local) supersymmetry in ten dimensions, i.e. it is invariant under two Majorana-Weyl supercharges (of different chirality).

Its massless sector contains the following 10d bosonic fields: The graviton G_{MN}, a 2-form B_{MN}, the dilaton scalar ϕ; A 1-form A_M and a 3-form C_{MNP}. Their supersymmetric partners are basically some $N = 2$ $D = 10$ gravitinos of opposite chirality (and spin 3/2) and two spin-1/2 fermions of opposite chiralities.

We would like to remark that the p-form fields C_p are gauge potentials, namely all their interactions and couplings are invariant under the gauge transformations with gauge parameter given by a $(p-1)$-form Λ_{p-1}

$$C_p \to C_p + d\Lambda_{p-1} \tag{2.24}$$

The gauge invariant field strengths are given by

$$H_{p+1} = dC_p \tag{2.25}$$

The above matter content is the gravity supermultiplet of non-chiral $N = 2$ $D = 10$ supergravity. Indeed the low energy effective action of type IIA string theory is that of non-chiral $N = 2$ $D = 10$ supergravity, and its form is uniquely fixed by supersymmetry. It contains the Einstein term, the kinetic term for the p-forms and the dilaton, and their supersymmetric completion involving the fermions.

It is also useful to know that the degrees of freedom in a p-form gauge potential C_p can be encoded in a dual $(8-p)$-form $\hat{C}_{8-p}$ by Hodge-duality of their field strengths

$$H_{p+1} = *_{10d}\hat{H}_{9-p} \tag{2.26}$$

So the 1-form has a 7-form dual, and the 3-form has a 5-form dual.

- **Type IIB superstring**

This is a theory of closed oriented strings.

Type IIB string theory has $N = 2$ (local) supersymmetry in ten dimensions, i.e. it is invariant under two Majorana-Weyl supercharges (of SAME chirality).

Its massless sector contains the following 10d bosonic fields: The graviton G_{MN}, a 2-form B_{MN}, the dilaton scalar ϕ; A 0-form a, a2-form $\tilde{B}_{MN}$ and a 4-form A_{MNPQ} of self-dual field strength. Their supersymmetric partners are basically some $N = 2$ $D = 10$ gravitinos of SAME chirality (and spin 3/2) and two spin-1/2 fermions of SAME chiralities. The p-forms are gauge potentials.

The above matter content is the gravity supermultiplet of CHIRAL $N = 2$ $D = 10$ supergravity. Indeed the low energy effective action of type IIB string theory is that of CHIRAL $N = 2$ $D = 10$ supergravity, and its form is uniquely fixed by supersymmetry. It contains the Einstein term, the kinetic term for the p-forms and the dilaton, and their supersymmetric completion involving the fermions.

An important observation is that the theory is chiral, so in principle it may be ill-defined at the quantum level due to gravitational anomalies (i.e. diffeomorphism invariance of the classical theory may be violated at the quantum level, leading to violations of unitarity, etc and rendering the theory inconsistent). Happily a detailed computation of the anomaly shows that it vanishes (in a very nontrivial way) [13].

• **The two versions of Heterotic string theory**

This is a theory of closed oriented strings.

Heterotic string theory has $N = 1$ (local) supersymmetry in ten dimensions, i.e. it is invariant under one Majorana-Weyl supercharge.

Its massless sector contains the following 10d fields: The graviton G_{MN}, a 2-form B_{MN}, the dilaton scalar ϕ, plus fermion superpartners. They fill out a graviton supermultiplet of $N = 1$ $D = 10$ supergravity. In addition there are 496 gauge bosons A_M^a associated to generators of a gauge group, which is either $E_8 \times E_8$ or $SO(32)$ (so there are two different versions of heterotic string theory). These gauge bosons have fermionic partners (in the adjoint representation of the gauge group, gauginos), filling out vector multiplets of $D = 10$ $N = 1$ supersymmetry.

The low energy effective action is that of $N = 1$ $D = 10$ supergravity, coupled to $E_8 \times E_8$ or $SO(32)$ gauge vector multiplets. The supersymmetry and gauge symmetry uniquely fixed the form of the effective action. It contains the Einstein term, the kinetic term for the 2-form and the dilaton, and Yang-Mills action for gauge bosons, and their supersymmetric completion involving the fermions.

An important observation is that the theory is chiral, so in principle it may be ill-defined at the quantum level due to gravitational and gauge anomalies. Happily a detailed computation of the anomaly shows that it vanishes (in a very nontrivial way), involving a novel mechanism (previously unknown in field theory), the so-called Green-Schwarz mechanism [14]. For the mechanism to work it is essential that the gauge group is one of the above mentioned.

• **Type I string theory**

This is a theory of closed and open unoriented strings. Unoriented means that the genus expansion includes non-orientable surfaces, like the Klein bottle or the Moebius strip, etc.

Type I string theory has $N = 1$ (local) supersymmetry in ten dimensions, i.e. it is invariant under one Majorana-Weyl supercharge.

Its massless sector contains the following 10d fields: The graviton G_{MN}, a 2-form B_{MN}, the dilaton scalar ϕ, plus fermion superpartners. They fill out a graviton supermultiplet of $N = 1$ $D = 10$ supergravity. In addition there are 496 gauge bosons A_M^a associated to generators of a gauge group, which $SO(32)$ (but NOT $E_8 \times E_8$). These gauge bosons have fermionic partners (in the adjoint representation of the gauge group, gauginos), filling out vector multiplets of $D = 10$ $N = 1$ supersymmetry.

The low energy effective action is that of $N = 1$ $D = 10$ supergravity, coupled to $SO(32)$ gauge vector multiplets. The supersymmetry and gauge symmetry uniquely fixed the form of the effective action. It contains the Einstein term, the kinetic term for the 2-form and the dilaton, and Yang-Mills action for gauge bosons, and their supersymmetric completion involving the fermions.

An important observation is that the theory is chiral, so in principle it may be ill-defined at the quantum level due to gravitational and gauge anomalies. Happily the anomaly cancels, also involving a version of the Green-Schwarz mechanism [15, 15].

This clearly shows that extra dimensions and supersymmetry and supergravity are ideas easily accommodated in the string theory setup. That (and the amazing self-consistency of the theory, namely the fact that it always leads to anomaly-free low-energy field theories) is the reason why lots of people got attracted into the study of these theories.

2.2.2 Heterotic string phenomenology

From the viewpoint of trying to reproduce the observed physics, many attempts were taken in the framework of Kaluza-Klein compactification in type II string theories. However, as discussed previously, it is difficult to reproduce chiral 4d fermions with the non-trivial gauge quantum numbers unless the original 10d theory contains elementary non-abelian gauge fields [16]. For that reason, compactification of other theories like type I or the heterotics is more promising.

In fact, most efforts centered in the study of heterotic theory. In a sense, if we study compactifications on curved spaces, where we use the low energy effective action, the type I theory looks very similar to the $SO(32)$ heterotic. Finally, there has been a traditional preference for the $E_8 \times E_8$ heterotic since it leads (in the simplest compactifications) to smaller gauge groups.

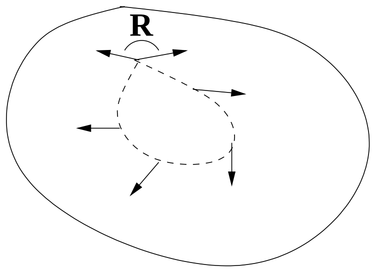


Figure 2.12: The holonomy group is given by the set of rotations R relating a vector and its image under parallel transport along a closed path, for all possible paths.

2.2.3 The picture of our world as a heterotic string compactification

Enough of a speculation! We would like to address what these constructions may have to do with the real world!

So, we conclude this brief review by describing the picture of our world as a heterotic string compactifications. This follows [61].

In order to obtain four-dimensional physics we need to take spacetime to be of the form $M_4 \times X_6$. The original 10d theory has a lot of supersymmetry: $D = 10$ $N = 2$ corresponds to 16 supercharges, the equivalent to $D = 4$ $N = 4$ supersymmetry. This amount of supersymmetry is too much to allow for 4d chiral fermions.

If X_6 is too simple, like a T^6, the supersymmetries are unbroken and we obtain a non-chiral theory. The reason why T^6 does not break supersymmetry is because it is flat, and has trivial holonomy group.

The holonomy group of a d-dimensional manifold (endowed with a metric) is defined by taking a vector, parallel-transporting it along a closed path, and finding the $SO(d)$ rotation relating the original vector and the final one. The set of all such roations for all possible closed paths is the holonomy group of the manifold (with the corresponding metric). For a torus, any vector comes back to itself (with no rotation at all) under parallel transport around any closed path. see figure 2.12.

For manifolds with non-trivial holonomy groups, there are topological obstructions to defining conserved supercharges globally [2], so the supersym-

[2]Similar to the impossibility of defining a global vector field in a 2-sphere, i.e. it is

metry observed at low energies correspons only to the supercharges which can be defined globally.

A generic 6-dimensional manifold has holonomy $SO(6)$ and breaks all supersymmetries. Manifolds with holonomy in a proper subgroup of the generic holonomy group are known as special holonomy manifolds. They break some supersymmetries, but preserve some.

For heterotic string theory, if X_6 is chosen to have $SU(3)$ holonomy, (which is a subgroup of $SO(6)$), then the low energy theory in 4d has only $N = 1$ supersymmetry. As discussed in the first lecture, this is a phenomenologically desirable feature. Spaces of $SU(3)$ holonomy are called Calabi-Yau spaces, and compactification on them is often called Calabi-Yau compactification.

On the other hand, the original gauge group in heterotic string theory is very large, it has 496 generators. We should think about some way of breaking it. Happily there is a way of doing it in the process of compactification.

Consider that, in the same way as we consider a non-trivial background for the internal metric (curved internal space), we consider turning on a non-trivial background for the internal components of the gauge potentials. That is, we turn on a nontrivial profile for the fiels A_i^a, with i polarized in the internal directions in X_6, and a associated to generators in a subgroup H of the original group, say in $E_8 \times E_8$. In fancy language, we are considering a non-trivial gauge bundle (with structure group H) over the manifold X_6.

This choice is consistent with Poincare invariance in four dimensions. However, since it priviledges some direction in gauge space, the gauge group observed at low energies is not the full $E_8 \times E_8$. In fact, the 4d gauge group is given by those gauge transformations which leave the gauge background invariant. This is the group generated by generators commuting with the generators of H, and is called in group theory the commutant of H in $E_8 \times E_8$.

Moreover, it can be seen that the consistency of a Calabi-Yau compactification requires SOME internal gauge background to be turned on. This is interesting, because it forces the gauge group to be broken, althouth consistency does not force on us any specific choice of the subroup H.

A very popular choice is the so-called standard embedding, which amounts to choosing $H = SU(3)$. More specifically, it corresponds to setting the internal gauge connection to be equal to the Riemannian connection on X_6. With this choice, the commutant of $SU(3)$ in $E_8 \times E_8$ is $E_6 \times E_8$. This is

impossible to comb a 2-sphere without leaving hair whirlpools.

a very exciting possibility, since E_6 has been considered as a possible group for grand unification. Taking slightly more involved choices for the gauge background it is possible to obtain even smaller groups, like $SU(5)$ of simply the Standard Model group.

The last ingredient that we would need is how to obtain chiral fermions charged under E_6 (or whatever other group we get in 4d). Amazingly the above ingredients (Calabi-Yau compactification and internal gauge bundle) are enough to provide chiral 4d fermions in the Kaluza-Klein reduction of the 10d gauginos. The resulting fermions transform naturally in the representation 27 of E_6 (or as $10 + \bar{5}$ of $SU(5)$, or standard fermion families of the standard model group).

The number of fermion families is given in terms of topological invariants of the internal manifold and the gauge bundle over it. For instance, for the standard embedding, it is given by the Euler number of X_6. The number of families is roughly speaking fixed by the number of (chiral) zero modes for a Dirac equation for the internal part of a 10d gaugino. So the different families are associated to different resonant modes of the 10d gaugino field in the internal X_6 space. B. Green describes this in a very poethic way [18]. It is possible (although not easy, it requires strong techniques in differential topology) to construct models where this number is 3.

The fact that the number of families is related to topological invariants is natural. In general one expects that, given a string compactification, the masses of light modes can vary if we make a small deformation of the configuration, like deforming the metric or the gauge background. However, the number of chiral families must be invariant under those deformations, because chirality protects fermions against getting Dirac masses. Hence, the number of chiral families is invariant under deformations of the metric or the gauge background, i.e. it is a topological invariant, which can be related to standard topological invariants of the manifold X_6 and the gauge bundle.

2.2.4 Phenomenological features and comparison with other proposals beyond the standard model

The lesson is that this picture, shown in figure A.2, provides four-dimensional theories which are extremely close to the Standard Model.

Moreover, the description includes some very interesting ingredients of physics beyond the standard model

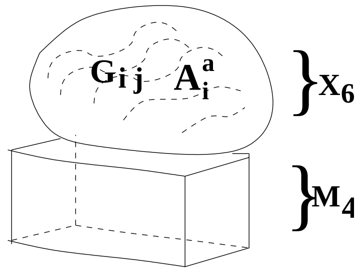


Figure 2.13: Schematic depiction of the compactification of heterotic string theory on a Calabi-Yau manifold (metric background) with non-trivial internal gauge bundle (gauge background).

- Unification: All interaction arise from $E_8 \times E_8$, so at high enough energies $E \sim M_c$, when we start to be able to resolve the internal space, the original 10d gauge symmetry is restored and all interactions are unified. Of course, there is also unification with gravity, as in all string theories. Heterotic string theory also predicts gauge coupling unification at a scale $\sim M_s$
- Supersymmetry: Is a basic ingredient in this construction. The issue of supersymmetry breaking remains an open question
- Hidden sector: One attractive possibility is to break supersymmetry by strong coupling dynamics (gaugino condensation) in the untouched E_8. This sector is decoupled from the Standard Model one, with which it communicates only via gravitational interactions, it is a hidden sector. So it implements the idea of supersymmetry breaking in a hidden sector.
- Extra dimensions. Also essential in the construction. Notice that both gauge and gravitational interaction propagate in 10d, so this constructino cannot be used to realize the brane-world scenario (other constructions, not based in heterotic, will be studied later on).

There also remain different open questions, whose answer is not clear for the moment. These are the main problems in string phenomenology, to be solved perhaps by next-generation students like you!

- How to break supersymmetry? There exist proposals like gaugino condensation, etc.

- The moduli problem: Or how to get rid of the large number of massless scalars which exist in many compactifications in string theory (and whose vevs encode the parameters of the underlying geometry and gauge bundle (like sizes of the internal manifold, etc)).
- The vaccum degeneracy problem: Or the enormous amount of consistent vacua which can be constructed, out of which only one (if any at all) is realized in the real world. Is this model preferred by some energetic, cosmological, anthropic criterion? Or is it all just a matter of chance?
- The cosmological constant problem, which in general is too large once we break supersymmetry. Does string theory say anything new about this old problem?

Chapter 3

Overview of string theory beyond perturbation theory

3.1 The problem

The prescription we have given to compute amplitudes in string theory in perturbation theory is well-defined and consistent. However, it is not the complete string theory, there are indications that there is plenty of non-perturbative structure missed by the prescription we have given.

Making an analogy with point particle physics, the perturbative prescription we have given is equivalent to giving the propagators for the different particles, and giving a set of interaction vertices. With both ingredients one can build the Feynmann diagrams of the theory and recover the complete perturbative expansion.

On the other hand, we know that in point particle physics there are plenty of non-perturbative effects (like non-perturbative states (solitons), instanton effects, etc) which are obtained only when we compute non-perturbatively (e.g. using lattice methods) the path integral over spacetime field configurations, using the spactime action of the theory.

Now in string theory we do NOT have a spacetime action for the spacetime fields configurations (we just have a worldsheet action, which is the analog of the worldline action in point particle physics, clearly not the same as a spacetime field action). Therefore we do not have a well-defined prescription to compute non-perturbatively the path integral over spacetime field configurations, and it is very likely that we are missing plenty of non-perturbative physics.

There exists an approach to string theory, dubbed string field theory, which introduces a string field $\Psi[X^\mu(\sigma,t)]$, which is a functional of the string configuration function $X^\mu(\sigma,t)$. It can be thought of as the spacetime wavefunction providing the quantum amplitude for a state to correspond to a string configuration given by $X^\mu(\sigma,t)$. Expanding in oscillator modes, the string field splits as an infinite set of spacetime (point particle) fields, each corresponding to a string oscillator mode (i.e. to a spacetime particle).

Subsequently, it is possible to build a spacetime action for the string field, such that the perturbative expansion reproduces exactly the perturbative string theory amplitudes computed with the above prescription.

On the other hand, one would expect that string field theory also encodes information about string theory beyond perturbation theory. For some reason, this last hope has not been quite fulfilled. String field theory is technically very involved, so not many solutions to the string field equations are known. In particular, string field theory has been unable to provide information about some string theory non-perturbative states found via other indirect methods (p-branes, D-branes) [1], so it is not clear that string field theory is the right tool to address non-perturbative dynamics in string theory (or else, perhaps is not the tool that we know how to handle). We will not discuss string field theory in these course.

In this lecture we discuss several other indirect methods which have uncovered part of the non-perturbative structure of string theory (although not to a complete microscopic definition of it).

One may wonder why, if there is no complete definition of string theory beyond perturbation theory, we still claim that it is a consistent, finite, theory of gravity at the quantum level, etc. This was only checked with the perturbative description. A related objection is why to bother about non-perturbative effects, and simply state that our theory is *defined* by the perturbative prescription. The objections are reasonable.

The reason why we need non-perturbative effects, and why we believe that they do not spoil (but rather improve) the good properties of string theory, is that there exist some very special, very singular, situations where perturbative string theory would break down, and certain computable non-perturbative effects make the physics non-singular and well-behaved.

[1]Nevertheless, string field theory has led to important results in the context of open string tachyon condensations, see [19]

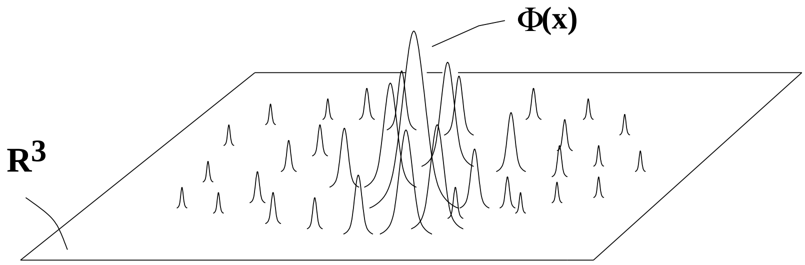


Figure 3.1: Artistic view of a soliton in a field theory.

So, our present understanding is that in smooth situations, the non-perturbative sectors do not spoil the good properties of perturbative string theory, they merely induce some small corrections. In other singular situations, however, the perturbative prescription would break down, and it is precisely the non-perturbative sector that saves the situation. We will see several examples of this phenomenon.

3.2 Non-perturbative states in string theory

A basic non-perturbative effect in string theory is the existence of states which are not seen in perturbation theory. That is, they do not appear in the Hilbert space of the quantized string. They are not modes of the fundamental string, so are not stringy in nature. They are more similar to solitons in field theories of point particles, which we now briefly review.

3.2.1 Non-perturbative states in field theory

An excellent discussion can be found in [72]. See also [73].

A soliton in a (to start with, classical) field theory is a finite energy solution to the equations of motion which is localized in some spatial dimensions, and is static in time. For instance, it the solution is localized (i.e. vanishes or goes to the trivial vacuum solution quickly outside a sphere of characteristic size R (the size of the soliton)) in three spatial directions in a four-dimensional field theory, then the solition looks like a 'fat' particle propagating in time. See picture 3.1.

There are explicit examples of such solitions. The simplest is the 't Hooft - Polyakov monopole [28], which we describe briefly.

The 't Hooft - Polyakov monopole

Consider the Georgi-Glashow model. It is an $SO(3)$ (or $SU(2)$) gauge field theory in four dimensions, with a complex scalar field (Higgs) charged in the adjoint representation (**3** of $SO(3)$). We denote it by $\vec{\phi}$, with the vector notation refering to the internal $SO(3)$. Let us take the scalar potential to have a minimum at $|\vec{\phi}|^2 = v^2$ [2]

The action is roughly speaking

$$S_{GG} = \int d^4x \, \frac{1}{g^2} \left[\text{tr} \, F^a_{\mu\nu} F^{\mu\nu}_a + D_\mu \vec{\phi} \cdot D^\mu \vec{\phi} \right] + V(\phi) \tag{3.1}$$

with

$$D_\mu \phi_i = \partial_\mu \phi_a + A^a_\mu \, (T_a)_{ij} \phi_j \tag{3.2}$$

Different vacua $|\vec{\phi}|^2 = v^2$ are related by $SO(3)$, so we may pick $\vec{\phi} = (v, 0, 0)$. The gauge group is spontaneously broken to $SO(2)$, equivalentely $U(1)$. This is the structure of the vacuum. Perturbative states of the theory are obtained by expanding the fields around the vacuum configuration, and contain the massive Higgs field, the massive vector bosons, etc. These generate different states in the quantum theory.

Now there also exist some finite energy configurations, which are therefore states in the quantum theory, which do not correspond to the above perturbative states. Consider a configuration where asymptotically in space $\mathbf{R^3}$ the field $\vec{\phi}(x)$ points (in the internal $SO(3)$) in the direction specified by the location $\vec{x}$ (in the space $\mathbf{R^3}$ $SO(3)$. Namely, for very large $r = |\vec{x}|$

$$\begin{aligned} \phi^a(\vec{x}, t) &\rightarrow \frac{v}{r} \, x^a \, + \mathcal{O}(1/r^2) \\ A^a(\vec{x}, t) &\rightarrow \frac{1}{r^2} \, x^a \, + \mathcal{O}(1/r^2) \end{aligned} \tag{3.3}$$

This is the so-called hedgehog configuration, shown in figure B.4.

[2] In many situations, for instance in supersymmetric models, the scalar potential is identically zero, and the vev for $\vec{\phi}$ is undetermined. Any vev defines a possible vacuum of the theory, the set of all possible vevs (up to gauge transformations) is called the moduli space of the theory. Notice that the name 'moduli' is associated to fields with no potentials, either in the string theory context (like the dilaton, or the compactficiation radii moduli) or in the field theory context. For each vev condition $|\vec{\phi}|^2 = v^2$ one may repeat the argument below.

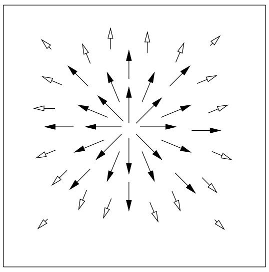


Figure 3.2: Picture of the hedgehog configuration for the Higgs field in the Georgi-Glashow model.

Since asymptotically $|\vec{\phi}| \to v$, the potential energy vanishes at infinity. The kinetic energy also vanishes asymptotically because we choose a gauge background which makes the covariant derivative vanish. Statitc solutions (solitons) with those asymptotics exist, and therefore have finite energy. They represent lumps of energy localized in the three spatial directions, i.e. particle-like states.

Their main properties are: their mass (energy of the configuration) is of the order of v/g^2, and so they are very heavy at weak coupling, and non-perturbative in nature. They are magnetically charged under the surviving $U(1)$ gauge group, i.e. taking the gauge field configuration in the soliton background, and integrating the field strength of the $U(1)$ part $F = F^a\phi^a$ around a large $\mathbf{S^2}$ in $\mathbf{R^3}$ we get

$$\int_{\mathbf{S^2}} F = 1 \tag{3.4}$$

These solitons are therefore called magnetic monopoles (in fact, magnetic monopoles in more realistic models, like grand unified theories, are constructed similarly). Since the charge they carry arises from the topology of the background (notice that the quantity (3.4) is topological, it is independent of the spacetime metric), they are also called topological defects.

Notice that if we had started with a higher dimensional theory, say in $D{+}1$ dimensions, one can still pick a particular $\mathbf{R^3}$ and construct the above soliton background. It is still localized in three dimensions, but the configuration is now Poincare invariant under the spectator $D-2$ dimensions. The soliton now represents an extended object with $D-3$ spatial dimensions. It is still charged magnetically with respect to the unbroken $U(1)$. The volume swept

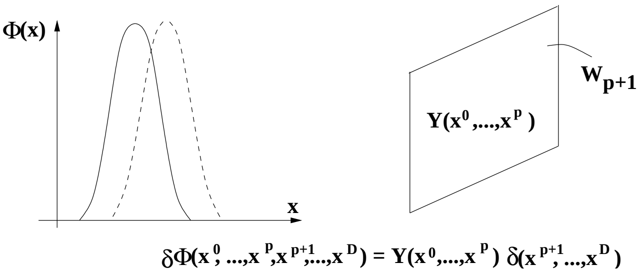


Figure 3.3: Picture of the zero modes of a soliton.

out by the soliton core as it moves in time is called the soliton world-volume (generalizing the ideas of worldlines and worldsheets).

Collective coordinates

It is interesting to see what the theory looks like around the soliton background. This is done by expanding the fields as background plus fluctuations, and substituting into the field theory action to obtain a field theory for the fluctuation fields. An interesting subset of fluctuations are zero modes, which correspond to fluctuations which are massless in the background of the soliton. They parametrize changes in the fields which do not change the energy of the soliton.

For instance, it is clear that applying translations $\phi_{x_0}(x) = \phi(x - x_0)$, one can construct solitons centered not at $\vec{x} = 0$ but at any $\vec{x} = \vec{x}_0$. The difference between two configurations $Y^i = \phi_0$ and $\phi_{\delta x^i}$ is a zero mode fluctuation. Notice tht both configurations are equal almost everywhere, so the fluctuation is localized on the volume of the soliton [3]. So, it can be roughly written as a field depending on the $p+1$ worldvolume coordinates (for a soliton with p spatial extended dimensions) $Y^i(x^0, \ldots, x^p)$, with $i = p+1, \ldots, D+1$. See picture 3.3 below.

In fact, the zero mode fluctuations describe dynamics of the soliton (and not dynamics of the underlying vacuum), they are sometimes called collective coordinates of the configuration. Very often they are associated to symmetries of the vacuum which are broken by the presence of the soliton (just like the above translational symmetries). So these massless fluctuations can

[3]Beyond those three translational collective coordinates, there is a fourth one associated to gauge transformations which do not vanish at infinity and therefore related different configurations which are not gauge equivalent. We will skip this mode in our discussion.

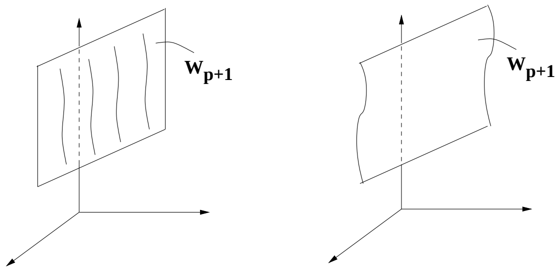


Figure 3.4: A nontrivial configuration for one of the worldvolume translational zero modes corresponds to a non-trivial embedding of the soliton worldvolume in spacetime.

be understood as Goldstone bosons of the symmetries broken in the soliton background.

Their vevs parametrize the possible configurations of the soliton background with the same energy; i.e. the set of soliton solutions of the same kind, e.g. location of solition worldvolume

$$\langle\, Y^i(x^0,\ldots,x^p)\,\rangle = a^i \tag{3.5}$$

The set of such vevs, the set of soliton configurations, is called the moduli space of solitons of that particluar kind (magnetic monopole moduli space in this case). Non-trivial configurations for these fields $Y^i(x^0,\ldots,x^p)$ describe excitations of the solition background; for instance a non-trivial profile for some of the translational zero modes corresponds to a non-flat soliton worldvolume (an energetically costly configuration). See picture A.7

It is possible to write down a worldvolume effective action for these worldvolume fields, which describes the dynamics of the soliton. We will not do so for the field theory example, but we will come back to this point when we look at non-perturbative states in string theory.

Beyond the classical approximation, the quantum behaviour of the soliton is obtained by expanding the classical theory around the soliton background, and quantizing the fluctuations. Concerning the subsector of the zero modes, this corresponds to promoting the worldvolume field theory to a quantum field theory in $p+1$ dimensions. And corresponds to quantizing the soliton state.

Many of these properties will have analogs in non-perturbative states in string theory, and that is why we discussed them in some detail.

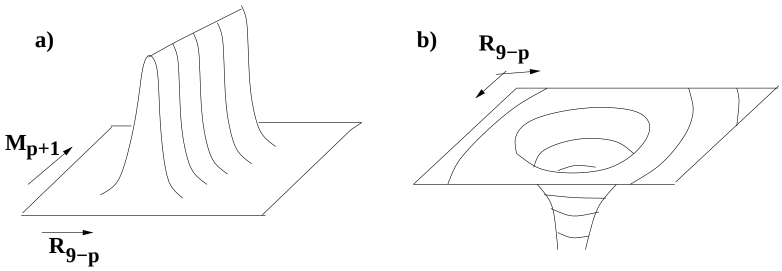


Figure 3.5: Two pictures of the p-brane as a lump of energy. The second picture shows only the transverse directions, where the p-brane looks like point-like.

3.2.2 Non-perturbative p-brane states in string theory

In order to try to find similar non-perturbative states in string theory, the only spacetime action that we can use to find spacetime field configurations is the low-energy effective action for the light modes of string theory (the graviton, dilaton, antisymmetric tensor fields, etc). It is important to realize that this is only the low-energy approximation to string theory, and it is questionable if any solution to its equation of motion is really a solution of full string theory. This issue will be settled for a particular class of solutions, as we will see below.

The approach is remarkably successful. Taking the different low-energy effective actions for the different superstrings (which correspond to different ten-dimensional supergravity theories), it is possible to find finite energy solutions (which are of a special kind (1/2 BPS) see below) to the equations of motion, which look like lumps of energy localized in some directions and extended in p spatial directions. They are known as p-branes; they have Poincare invariance in $p+1$ dimensions, and the core of the non-perturbative lump is called the p-brane world-volume. See 14.1 for a picture

To give one example, the supergravity solution for a 3-brane (with N units of charge) in type IIB theory is given by

$$\begin{aligned} ds^2 &= f(r)^{-1/2}\,[(dx^0)^2+\ldots+(dx^3)^2]\;+\;f(r)^{1/2}\,[(dx^4)^2+\ldots+(dx^9)^2] \\ f(r) &= 1+\frac{4\pi g_s\alpha'^2 N}{r^4} \quad ; \quad r=[(x^4)^2+\ldots+(x^9)^2] \\ F_5 &\simeq d(Vol)_{\mathbf{S^5}} \end{aligned} \tag{3.6}$$

where the field strength 5-form is proportional to the volume form of the angular 5-sphere in the transverse six-dimensional space.

The main properties of these solutions are

- For a given string theory, there exist p-brane solutions for values of p for which there exists a $(p+1)$-form field in the (perturbative) massless spectrum of the string. See table 3.1

- The energy per unit volume of these branes is of order $1/g_s$ or $1/g_s^2$ in string units $M_s = 1$. So they are intrinsecally non-perturbative

- p-branes are charged electrically under the $(p+1)$-forms; conversely, they are charge magnetically under the dual $(7-p)$-forms, namely

$$\int_{\mathbf{S^{8-p}}} H_{8-p} = 1 \tag{3.7}$$

  where H_{8-p} is the field strength for the $(7-p)$-form, and we integrate over a $(8-p)$-sphere in the transverse $\mathbf{R^{9-p}}$.

- The solutions are invariant under half of the supersymmetric transformations of the vacuum theory. The solutions are said to be 1/2 BPS. This is the key property that makes these solutions special, and reliable beyond the supergravity approximation.

- We will not discuss these theories in detail, but the worldvolume field theories for these p-branes are known. They contain $9-p$ real scalar fields, Goldstone bosons of the broken translational symmetries, and some fermions, which can be understood as Goldstinos of the supersymmetries broken by the background. These (and other) fields group together in multiplets of the unbroken supersymmetries, and define a supersrymmetric field theory in $p+1$ dimensions.

We turn to the issue of why the existence of these non-perturbative states should be trusted in the full string theory. After all, we found them as solutions of a truncated theory, the supergravity effective action describing the $\alpha' = 0$ regime.

The key feature is that BPS states are remarkably stable under smooth deformations of the theory (like for instance, turning on α' i.e. including more and more stringy corrections until we eventually reach full string theory). The argument proceeds through various steps

String theory	Branes	$(p+1)$-form	Tension
Type IIA	F1, NS5	B_2, $\hat{B}_6$	$\simeq 1/g_s^2$
	D0, D2, D4, D6, D8	C_1, C_3 $\hat{C}_5$, $\hat{C}_7$	$\simeq 1/g_s$
Type IIB	F1, NS5	B_2, $\hat{B}_6$	$\simeq 1/g_s^2$
	D(-1), D1, D3, D5, D7	a, $\tilde{B}_2$, C_4, $\hat{C}_6$, $\hat{C}_8$	$\simeq 1/g_s$
Heterotic	F1, NS5	B_2, $\hat{B}_6$	$\simeq 1/g_s^2$
Type I	D1, D5	$\tilde{B}_2$, $\hat{C}_6$	$\simeq 1/g_s$

Table 3.1: Partial list of the spectrum of p-branes in the different string theories.

i) Recall how one builds supersymmetric multiplets of states in a supersymmetric theory. One separates the supergenerators of the theory, in two sets (creators and annihilators), and defines the ground state of the multiplet as annihilated by annihilators. The rest of the multiplet is obtaine by applying creators to the ground state and using the algebra.

A 1/2 BPS state is invariant under half of the supersymmetries, so the ground state of the supermultiplet is annihilated by the creator operators of the corresponding susys. This means that this kind of multiplet contains half the number of states as a generic multiplet. Consequently, multiplets are called short and long, according to the number of states they contain.

To give a toy description, consider four supercharges, separated as two annihilators Q_1, Q_2 and their adjoints the creators $Q_1^\dagger$, $Q_2^\dagger$. A generic multiplet, constructed form a ground state $|st.|rangle$ satisfying $Q_i|st.\rangle = 0$, is given by

$$|st.\rangle \quad , \quad Q_1^\dagger|st.\rangle \quad , \quad Q_2^\dagger|st.\rangle \quad , \quad Q_1^\dagger Q_2^\dagger|st.\rangle \tag{3.8}$$

A 1/2 BPS multiplet is built out of a ground state which in addition satisfies $Q_2^\dagger|st.\rangle$, so the multiplet contains

$$|st.\rangle \quad , \quad Q_1^\dagger|st.\rangle \tag{3.9}$$

Namely contains half the number of states.

ii) Since the number of states in short and long multiplets is different, it is not possible that a BPS state becomes non-BPS upon a continuous change of parameters of the system. In particular, BPS states remain BPS upon turning on α'.

iii) The supersymmetry algebra in the presence of p-form charges is modified by the inclusion of central charges $Z(\phi)$ (operators that commute with all supergenerators and the hamiltonian, and appear in the susy algebra). They are related to the charges of the configurations, and are known functions of the moduli. The susy algebra looks like

$$\{Q_\alpha^A, Q_{\dot{\alpha}}^B\} = \delta_{AB}\,(\sigma^\mu)_{\alpha\dot{\alpha}} P_\mu + Z_{\alpha\dot{\alpha}}^{AB}(\phi) \tag{3.10}$$

Applying the algebra to the ground state of the BPS multiplet for the choice of Q^B that annihilates it, the left hand side gives zero. On the right hand side, in the rest frame of the brane, the momentum operator looks like $(M, 0, \ldots, 0)$ with M the mass or tension of the object, while Z gives its charge. Roughly speaking we get a relation $M = Q$, namely the tension of the BPS object is determined in terms of its charge.

iv) Since charges are quantized, they cannot change as we change parameters continuously. Since BPS states remain BPS upon such changes, their tension remains determined by their charges, so it is possible to determine them exactly even after all α' corrections are included.

This concludes the argument. If we find a BPS state in the supergravity approximation and compute its properties (charge, tension), there will exist a BPS state (a stringy improved version of the original one) with the same properties in the full string theory. The tension of the object is determine from its charge as dictated by the central extension of the susy algebra, so they can be reliably followed as moduli change (for instance, as the coupling gets strong).

BPS states are a subsector of the theory which is protected by supersymmetry, so it can be reliably studied in some simpler approximation schemes, like low-energy effective supergravity.

3.2.3 Duality in string theory

p-brane democracy

We start this section by pointing out a remarkable fact. Some of the p-branes that we have discussed above carry the same charges as the string, namely they have electric coupling to the (NS-NS) 2-form in the massless sector, just like string. In fact, the corresponding supergravity solution corresponds to the background created by a macroscopic, infinitely extended, string. But which is not essentially different from the basic string of the theory. For this reason, such 1-branes are known as fundamental string solutions and denoted F1-branes.

The fact that the fundamental string arises, in this sense, in the same way as other p-branes, suggests the idea that perhaps all p-brane solutions should be treated on an equal footing. This is also suggested by the fact that different brane solutions are often related by symmetries in supergravity, called U-duality symmetries (a discrete subgroup of which is realized in full-fledged string theory. This idea that different branes are on an equal footing is called p-brane democracy [22].

Of course, we have learned that in perturbation theory the fundamental string is more fundamental than any other object in the theory. In particular, a large part of the spectrum of the theory is obtained by quantizing the oscillation modes of the fundamental string. The p-brane democracy idea proposes that this is just and artifact of the perturbative description.

The idea is that there is a unique underlying theory with a bunch of BPS states. As one moves to a particular limit (like weak coupling) some of these states look more fundamental than others, and the light spectrum in that limit can be computed by quantizing these fundamental objects. In particular, it is conceivable that there exist other limits where other BPS states are fundamental and are more useful to describe the physics of the system.

This is the picture underlying the proposal of string duality.

String duality

Indeed this idea is realized in many string configurations. The simplest case is that of the ten-dimensional superstrings. There exists a perturbative limit where the theory is described in terms of weakly interacting strings and one recovers the perturbation theory we have described in previous lectures. As one moves to the non-perturbative regime, the different branes look really democratical. In the limit of infinite coupling the theory again simplifies and becomes a weakly interacting theory, but where the fundamental degrees of freedom correspond to originally non-perturbative states. The situation is shown in picture 14.3. Notice that the tensions of the objects can be realiably followed as a function of the moduli (the dilaton vev, string coupling) thanks to the fact that these states are BPS.

Thus, roughly speaking, the strong coupling limit of a string theory can be described as a weak coupling limit of a dual string theory (which may be or

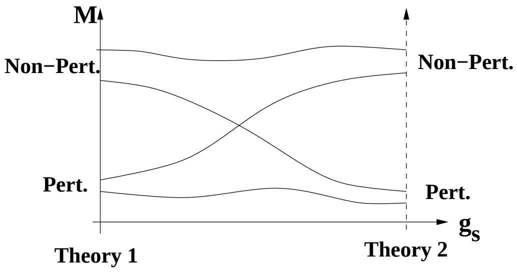


Figure 3.6: As a modulus (the dilaton vev) is changed, the original weakly coupled string theory becomes strongly interacting, and at infinite coupling it can be described as a weakly interacting *dual* theory. Perturbative and non-perturbative states are reshuffled in this interpolation.

not of the same kind). Perturbative and non-perturbative states are reshuffled as one changes the vev of the dilaton modulus to interpolate between them. We will see explicit examples below

We now explain the dual theories describing the strong coupling regime of the ten-dimensional superstrings. The original reference for these results is [23]

Duality for ten-dimensional superstrings

Type IIB self-duality

The limit of strong coupling of type IIB string theory is described by a different type IIB string theory, with weak coupling. The string couplings in the two theories are related by

$$(g_s)_1 = 1/(g_s)_2 \tag{3.11}$$

The basic mapping of branes are as follows

Type IIB	$\leftrightarrow$	Type IIB
F1, NS5		D1, D5
D3		D3

The mapping of massless fields is easy as well, roughly speaking

Type IIB	$\leftrightarrow$	Type IIB
$\tau = a + ie^{-\phi}$		$-1/\tau$
G_{MN}		G_{MN}
B_2		$\tilde{B}_2$
$\tilde{B}_2$		B_2
C_4		C_4

The transformation $g_s \to 1/g_s$ is a transformation that maps type IIB string theory to itself. In particular, it is a subgroup of an exact $SL(2, \mathbf{Z})$ symmetry of type IIB theory. This symmetry group is a particular case of U-duality, which encodes duality properties of the theories upon compactification, and can be used to find dual description in other limits. See [24].

$SO(32)$ **heterotic - Type I duality**

The strong coupling limit of the $SO(32)$ heterotic string is described by a dual weakly coupled type I theory, and viceversa. The mapping of branes is

$SO(32)$ Heterotic	$\leftrightarrow$	Type I
F1, NS5		D1, D5

The mapping of fields is: the string coupling is inverted, the 2-forms are exchanged, the metric and the $SO(32)$ gauge fields are invariant.

Notice that the relation implies a mapping between the low-energy supergravity theories, written in terms of heterotic and type I variables. This is possible because both sugra theories are $d = 10$ $N = 1$ sugra coupled to $SO(32)$ gauge multiplets.

Type IIA - M-theory duality

As the coupling constant of type IIA theory gets stronger, the strong coupling limit is not described by a dual string theory, but rather in terms of a far more mysterious theory called M-theory. The argument is as follows.

Type IIA theory contains non-perturbative particle-like D0-branes, with masses given by k/g_s, where k is the D0-brane charge under C_1. In the strong coupling limit, all these states are becoming massless, so the strong coupling limit is a theory with an infinite tower of states becoming massless.

The idea is to propose that type IIA theory has a dual description as an 11d theory compactified on a circle, with radius related to the string coupling as $R = g_s$. The states with mass k/g_s correspond in the dual picture to the Kaluza-Klein replicas of the 11d graviton multiplet. Type IIA theory at extreme strong coupling corresponds to the decompactification limit of this theory.

There is a supergravity theory in 11d which under compactification on a circle reduces to $d = 10$ $N = 2$ non-chiral supergravity. It contains and 11d gravition, a 3-form C_3 (and its dual $\tilde{C}_6$), plus gravitino etc superpartners. In particular, it does not contain a dilaton field, so it does not have a coupling constant. This theory is however ill-defined in the UV (non-renormalizable, etc), so should be regarded as an effective description of an underlying quantum theory, which for the moment is completely unknown . So the natural proposal is that the strong coupling limit of type IIA theory corresponds to a quantum theory, called M-theory, whose low energy limit is given by 11d supergravity.

This is a nice result, and explains the role of 11d sugra in string theory (previously this sugra was unrelated to string theory, in contrast with its 10d cousins). Understanding the microscopic degrees of freedom of M-theor, the theory underlying 11d sugra, in one of the main challenges in string theory today.

M-theory also contains p-brane states, which are found as BPS solutions to 11d sugra, which therefore must exist in the full theory (since they are BPS). They correspond to a 2-brane and a 5-brane, denoted M2-, M5-branes, resp. The mapping of fields between Type IIA and M-theory is

M-theory	$\leftrightarrow$	Type IIA
G_{MN}	$\rightarrow$	$G_{\mu\nu}$
		$A_\mu = G_{\mu,10}$
		$\phi = G_{10,10}$
C_{MNP}	$\rightarrow$	$B_{\mu\nu} = C_{\mu\nu,10}$
		$C_{\mu\nu\rho}$

On the other hand, Type IIA D0-branes are KK replicas of the 11d fields, the D2-brane is an M2-brane transverse to the M-theory $\mathbf{S^1}$, the F1 is an M2 wrapped on the $\mathbf{S^1}$, the D4 is an M5 wrapped on $\mathbf{S^1}$, the NS5 is an unwrapped M5. Finally the D6-brane corresponds to a purely gravitational background in M-theory known as Taub-NUT metric.

$E_8 \times E_8$ **heterotic - Horava-Witten duality**

The strong coupling limit of the $E_8 \times E_8$ heterotic is also not a string theory, but is related to a compactification of M-theory. Heterotic theory has less supersymmetry than M-theory, so we need to break some of the supersymmetry in the compactification. The compactification is taken to be not on a circle $\mathbf{S^1}$, but on the quotient of a circle by the $\mathbf{Z_2}$ symmetry corresponding to reflection with respect to one of its diameters, and simultaneously mapping C_3 to $-C_3$. This is equivalent to compactification on an interval, see picture B.4 This compactification of M-theory is known as

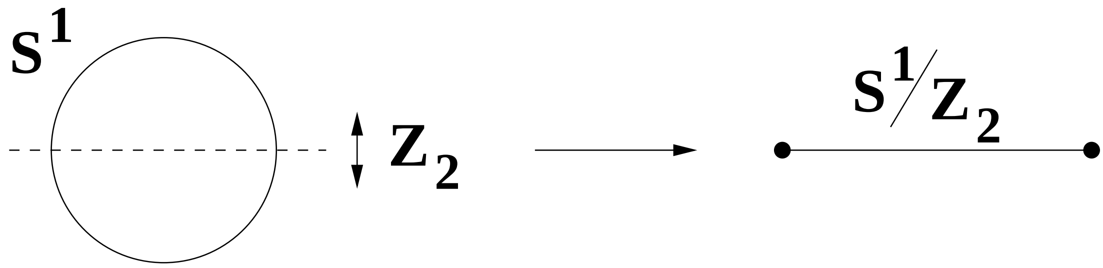


Figure 3.7: The quotient of a circle by a reflection under a diameter is an interval $I = \mathbf{S^1}/\mathbf{Z_2}$.

Horava-Witten theory [82].

The $E_8 \times E_8$ heterotic at string coupling g_s is proposed to be equivalent to the compactification of M-theory on the interval of radius $R = g_s$. Again, the heterotic strong coupling limit corresponds to the decompactification limit.

The mapping of fields is as follows. The $N = 1$ $d = 1$ supergravity multiplet of the heterotic theory is mapped to the sector of 11d supergravity which is invariant under the $\mathbf{Z_2}$ symmetry. On the other hand, the E_8 gauge multiplets must necessarily arise at the fixed points of the $\mathbf{Z_2}$ action, so they are localized at the ten-dimensional boundaries of the spacetime $M_{10} \times I$. Each E_8 gauge multiplets propagates at one of the boundary points of I times M_{10}, and does not propagate in the M-theory direction. This is our first example of gauge interactions localized on a submanifold of spacetime. see figure B.5.

The duality web

As one compactifies the 10d theories, more moduli appear, associated to the geometry of the compactification space. Then there are more limits that can be taken, for instance, strong coupling and small radii, with fixed ratios. In this situation more duality relations appear; These dualities involve non-perturbative as well as perturbative dualities, like T-duality. To give just one example, compactification of M-theory on a two-torus is dual or equivalent to compactification of type IIB theory on a circle, etc. This can be understood by taking M-theory reducing to IIA on a circle, then reducing on a second circle, and T-dualizing to type IIB theory.

Different compactifications of the different superstrings and M-theory are related by an intricate duality web. We will not describe any more dualities in this lecture. But they suggest a nice picture that we would like to discuss

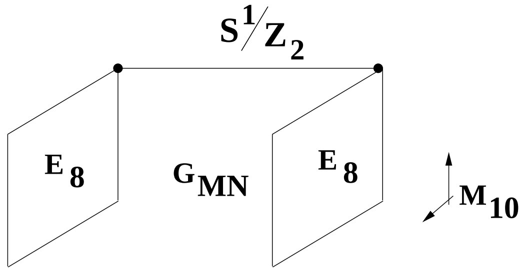


Figure 3.8: The strong coupling description of $E_8 \times E_8$ heterotic involves the compatification of M-theory on a space with two 10d boundaries. Gravity propagates in 11d, while gauge interactions are localized on the 10d subpaces at the boundaries.

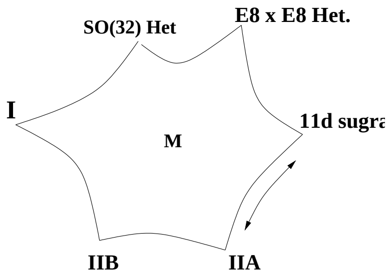


Figure 3.9: Map of the moduli space of the underlying theory and its different known limits.

The picture that emerges is that in a sense there is a unique theory, which describes all kinds of extended BPS objects, and which in different limits reduces to perturbatives string theories (where strings are the fundamental objects) or to other more exotic theories (like M-thoery, which is not a string theory). This picture has become popular in the pictorial representation 3.9. By abuse of language, the underlying theory is often called M-theory as well.

Surprisingly enough, string theory is NOT just a theory of strings!! It is a huge challenge to really understand what string theory is about, once we are far from any perturbative regime.

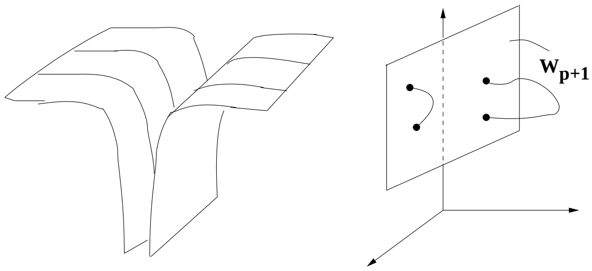


Figure 3.10: Fluctuations of the theory around a Dp-brane sugra solution can be described in stringy language as open strings with ends on a $(p+1)$-dimensional surface, located at the core of the topological defect.

3.3 D-branes

We conclude this lecture with a brief review of a very simple description of some p-brane states in type II and type I theories, the Dp-branes.

3.3.1 What are D-branes

Given a p-brane state, one is interested in the spectrum of the theory when expanded around this state. In general, this can be computed only in the supergravity approximation, by expanding the sugra fields in background plus fluctuations and computing the action for fluctuations by substitution in the sugra action. This is extremely involved, and moreover suffers from plenty of corrections.

The remarkable insight by Polchinski [26] is that he gave a completely stringy proposal to obtain the spectrum of fluctuations of string theory around certain p-brane states, the Dp-branes mentioned above. In fact, it is a stringy definition of such p-brane states.

The proposal is to replace the p-brane soliton core by a $(p+1)$ dimensional hypersurface in flat space. The fluctuations of the theory around the p-brane background correspond to open strings with ends on this hypersurface. The spectrum of fluctuations of the theory around the p-brane background can be obtained by simply quantizing such open strings. The hyperplane is known as Dp-brane. The situation is shown in figure 3.10.

Notice that the Dp-brane, as a state, is non-perturbative, it does not appear as an oscillator state of the string. On the other hand, what we have

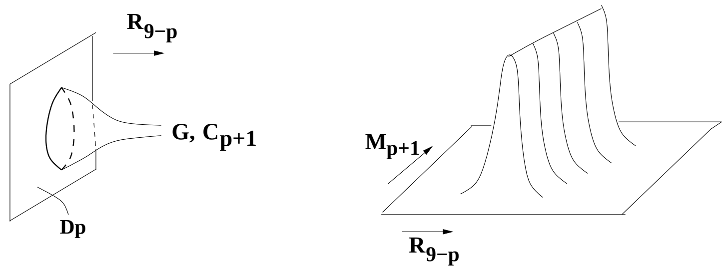


Figure 3.11: D-branes interact with closed string modes, and in particular couple to the bulk graviton and $(p+1)$-form fields, i.e. they have tension (of order $1/g_s$ in string units) and carry charge. Their backreaction on the background curves and deforms it into the p-brane solution seen in the supergravity regime.

provided is a stringy description of the spectrum of fluctuations of the theory around the p-brane state, in terms of oscillation modes of open strings with ends on the Dp-brane worldvolume.

Properties

This surprising proposal works. The Dp-brane interacts with closed string via diagrams with the topology of a disk, as in figure A.4.

In particular, they can be seen to carry tension and charge, which matches the tension and charge of the p-brane solutions in supergravity. This suggests that the Dp-branes described as subspaces where open strings can end is a stringy version of the fat p-brane solutions of supergravity. The back-reaction of the Dp-brane on the flat background curves and modifies it to the full sugra solution.

Moreover, it can be seen that the Dp-branes described in this way break half of the supersymmetries, so they are BPS states of the theory.

It is important to notice that NOT all p-branes in string theory are Dp-branes. For the NS5-branes and others, there is no simple stringy description for their spectrum of fluctuations. So the study of the dynamics of these objects is much more complicated than for D-branes.

It is also important to realize that NOT all superstring theories contain D-branes. Namely, the p-branes in heterotic string theories are not Dp-branes, so there are no D-branes in heterotic theories. Type IIB theory contains D$(2p+1)$-branes, while IIA contains D$2p$-branes, and type I contains D1, D5 and D9-branes.

3.3.2 Worldvolume theory

The quantization of open strings leads to a stringy tower of modes. The lightest of these are massless and correspond to the zero modes of the topological defect as introduced above. Consider a Dp-brane with $(p+1)$-dimensional worldvolume extended along the directons $x^0, \ldots, x^p$, in flat 10d spacetime. Consider an open string with both endpoints on the Dp-brane. The lightest oscillation states of this string correspond to gauge bosons, A_μ, $9-p$ scalars Y^i (Goldstone bosons of the translational symmetries of the vacuum, broken by the Dp-brane), and some fermions λ^a (Goldstinos of the supersymmetries of the vacuum which are broken by the Dp-brane). Notice that since the open string endpoint must be on the D-brane worldvolume, these fields are naturally localized on the D-brane worldvolume. They define a $(p+1)$-dimensional field theory, which describes the dynamics of the Dp-brane. For instance, for a D3-brane in type IIB theory, the massless modes of an open string with ends on the D3-brane correspond to a $U(1)$ vector boson, six real scalar fields, and four Majorana fermions, all neutral under the $U(1)$ group.

An important feature of Dp-branes (and p-branes) in general, is that the BPS property implies that several parallel Dp-branes of the same kind do not suffer net attraction or repulsion. The equality of tension and charge for BPS branes guarantees that the gravitational attraction is cancelled by the repulsion due to their equal charges. So it is possible to consider configurations with several parallel Dp-branes at arbitrary points in the transverse space.

In particular, several of these Dp-branes may coincide at the same point. This is an interesting configuration, so let us consider n coincident Dp-branes in flat 10d space. Without going into much details, it is possible to understand that now there are n^2 possible open strings, depending on on which brane the string is starting (out of the n possible ones) and on which it is ending (out of the n possible ones). It is important to recall that we work with oriented open strings. The situation is shown in figure 3.12. The spectrum in each sector is similar, so the total open string sector, for D3-branes for instance, contains n^2 4d gauge bosons, which can be seen to organize into an $U(n)$ gauge group, six 4d real scalars, with transform in the adjoint representation (of dimension n^2), and four 4d Majorana fermions, also in the adjoint.

If the D-branes are slightly separated, the stretching of the open string means that some of the fields are slightly massive, with mass given by the

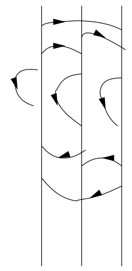


Figure 3.12: Open string stretched within a stack of 3 overlapping D-branes. They are shown as separated for the sake of clarity.

string tension times the D-brane separations. The above modes are massless for overlapping D-branes, and have small masses $\ll M_s$ if the inter-D-brane distance is much smaller than the string length.

The interpretation of these modes is trickier than for just one brane. In general, we may say that the eigenvalues of the scalars vevs (which are matrices in the adjoint) correspond to the positions of the D-branes in transverse space. However, there is an intriguing underlying matrix structure, which leads some researchers to the idea that spacetime positions, coordinates, should become matrices at length scales much smaller than the string length. This idea underlies some of the most advanced proposals to understands string theory, M-theory, and the structure of spacetime, like the M(atrix) theory proposal [29].

The effective action for light modes of the open strings can be obtained by computing their scattering amplitudes using the rules in the previous sections, and cooking up an effective action reproducing them. Alternatively, one can consider turning on a background for these fields (for instance, for the D-brane gauge fields), writing a 2d action for the worldsheet in the presence of these fields, and imposing that the worldsheet theory is conformally invariant. The coupling of gauge fields to the worldsheet is described by adding to the usual Polyakov action the boundary action

$$S_{bdry} = \int_{\partial\Sigma} d\xi^a \partial_a X^\mu(\sigma, t) A_\mu(X(\sigma, t)) \tag{3.12}$$

where $\partial\Sigma$ is the boundary of the wordsheet Σ. It amounts to taking the 1-form A_1 on the D-brane worldvolume, and integrating it along the 1d worldsheet boundary, i.e. $S_{bdry} = \int_{\partial\Sigma} A_1$. This shows that the string endpoints are charged with respect to the worldvolume 1-form gauge field.

By either method, one obtains a $(p+1)$-dimensional effective action for the worldvolume massless modes, which looks like (a supersymmetrization with respect to the 16 unbroken supercharges, in type II D-branes)

$$S_{Dp} = T_{Dp} \int d^{p+1}x \, [\, -det(G + B + \alpha' F)\,]^{1/2} \tag{3.13}$$

plus some topological terms (Wess-Zumino terms) which will not interest us for the moment. This is the so called Dirac-Born-Infeld action (DBI). Here G and B are the induced metric and 2-form induced on the worldvolume from the 10d ones, and F is the worldvolume field strenght. The leading order of this action is just the string tension times the worldvolume volume; next order in F is the Yang-Mills action for the worldvolume gauge bosons [4] So the vector bosons A_μ are indeed gauge bosons.

So this is a second situation where we find that gauge interactions can be consistently localized to subspaces of spacetime, while gravity propagates in full spacetime. These gauge interactions are therefore qualitatively different from those in heterotic string theory.

A last comment is that considering a non-trivial background for the worldvolume scalar fields $Y^i(x^0, \ldots, x^p)$ amounts to considering a curved Dp-brane worldvolume. Dp-brane can therefore do all kinds of things, like wrap a non-trivial cycle in a topologically non-trivial spacetime (for example, wrap around a circle in the internal space in a $M_4 \times T^6$ compactification).

3.3.3 D-branes in string theory

Here we review some of the main applications where D-branes are important in string theory

Theories with open strings

Some string theories, like type I, contain open strings already in their vacuum state. D-branes have become so useful and popular, that now any theory with open strings is rephrased in D-brane language. Using the above rules, the space where open strings are allowed to end IS a D-brane, which is present in

[4] In fact the DBI action is valid just for $U(1)$, the generalization to the non-abelian case is not known.

the vacuum of the theory (so in the present context should not be regarded as a soliton-like excited state!).

For instance, type I theory contains open strings already in its vacuum, so contains a number of D-branes in its vacuum. Since the endpoints of type I open string can be anywhere in 10d space, the D-branes in the vacuum of type I theory have a 10d worldvolume, which fills 10d spacetime completely, namely they are D9-branes. The gauge bosons in type I theory can be regarded as the gauge bosons on the worldvolume of these D-branes. There are 32 D9-branes in type I theory, so the gauge group in the open string sector would be $U(32)$, but the fact that the open strings are unoriented reduces the group to $SO(32)$. We will construct this theory in more detail in subsequent lectures.

Non-perturbative effects and D-branes

Effects of non-perturbative states in string theory can be very important. Here we would like to review a situation where the perturbative description of string theory breaks down and give singular answers for some quantities; happily, non-perturbative effects come to the rescue precisely in this situation and make physics of string theory smooth.

Strominger's conifold

In the study of the compactification of type IIB theory on Calabi-Yau spaces, one realizes that the effective action becomes singular at a point in the moduli space of Calabi-Yau geometries. This means that the perturbative prescription for computing amplitudes is giving some infinite answers, which appear as a singular behaviour in the dependence of the string action on moduli vevs.

This seemingly ill behaviour of string theory puzzled experts for many years. The issue was solved in a beautiful paper [90], which realized there is a non-perturbative state playing a key role in this situation.

It can be seen that the singular behaviour appears precisely at the point in moduli space where one submanifold of the Calabi-Yau, a 3-cycle, degenerates to zero size. The geometry of the Calabi-Yau near this 3-cycle can be locally described by the set of points in $\mathbf{C^4}$ satisfying the equation

$$z_1^2 + z_2^2 + z_3^2 + z_4^2 = \epsilon \tag{3.14}$$

and ϵ is the vev of a modulus field in 4d, which controls the size of the 3-cycle (for instance, if ϵ is real, the above CY contais a 3-sphere of radius

$\epsilon^{1/2}$, obtained by restricting to real z_i). This geometry is called the conifold singularity, and is very popular in the string theory community (it is the most generic singularity in Calabi-Yau spaces).

Strominger's insight was to realize that there exist a non-perturbative state which corresponds to a D3-brane wrapped on this 3-sphere, so which looks like a particle-like state in 4d. Its mass is the D3-brane tension times the 3-sphere volume

$$M_{D3} = T_{D3} V_{\mathbf{S^3}} \tag{3.15}$$

so the particle is becoming massess as $\epsilon \to 0$. Therefore, the dynamics of this state is exteremly relevant, precisely at the point at which the perturbative effective action is becoming singular. Strominger moreover provided quantitative arguments showing that including the additional light state into the effective action makes it smooth and well behaved. And integrating it out in the smooth effective action leads to the singularity observed using just the perturbative prescription.

In fact, the theory has 4d $N = 2$ susy, so its action is completely determined once the spectrum is known. The relevant piece of the spectrum is an $N = 2$ $U(1)$ vector multiplet, whose gauge boson arises from the IIB 4-form with three indices along the 3-cycle; and one $N = 2$ hypermultiplet, given by the D3-brane state, charged under the vector multiplet. The effective action is just an $N = 2$ $U(1)$ gauge field theory with one charged hypermultiplet. Completely standard and completely smooth!

Notice that the result is present no matter how small the string coupling is. Here non-perturbative effects are crucial even in the perturbative regime.

Notice also that the result is amazing from the string theory perspective. Here we have a light particle, which is not describe as an oscillation mode of the string. It is however natural from the viewpoint of non-perturbative string theory, where objects with different string or brane nature are on an equal footing.

There are many other examples of this kind of behaviour. As usual, string theory is clever enough to give finite answers even in the most singular situations. The theory has an incredible amount of self-consistency.

Topology change

Further investigation of the conifold non-perturbative states led to a fantastic effect [91]. Non-perturbative states can mediate phase transitions

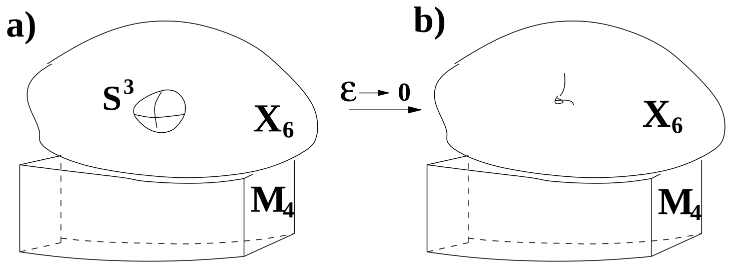


Figure 3.13: Tuning a modulus in the Calabi-Yau geometry, a 3-cycle shrinks and the geometry develops a conifold singularity.

where the topology of the internal space (and so, of spacetime, changes). Taking a Calabi-Yau with two conifold singularities (with homologically related $\mathbf{S^3}$'s), and shrinking the corresponding 3-cycles, one finds that at the singular point in moduli space the low energy field theory is $N = 2$ $U(1)$ gauge theory with two charged massless hypermultiplets, H_a. This theory has a Higgs branch, where these hypermultiplets (which have non-perturbative origin!) acquire and expectation value along a flat direction of the scalar potential. The flat direction is parametrized by a field with no potential, a modulus. It has a geometric interpretation, which corresponds to parametrizing the size of 2-spheres which resolve the conifold singularities. This is schematically shown in fig C.3.

In the process of sending $\epsilon \to 0$ and going to the Higgs branch the topology has changed, we have replaced an $\mathbf{S^3}$ by and $\mathbf{S^2}$. The transition is codified in a picture like C.4

This fact is remarkably important. The fact that string theory can smoothly interpolate between compactification spaces of different topology means that the choice of compactification manifold is in a sense dynamical, and determined by vevs of dynamical fields of the theory. All moduli spaces of different compactifications are connected into a huge universal moduli space.

3.3.4 D-branes as probes of spacetime

As already mentioned, vevs of worldvolume massless scalar fields correspond to coordinates of the brane in transveser space. This means that the moduli space of vacua of the field theory on the volume of a D-brane is the geometry of the space transverse to the D-brane. In this sense, spacetime can be

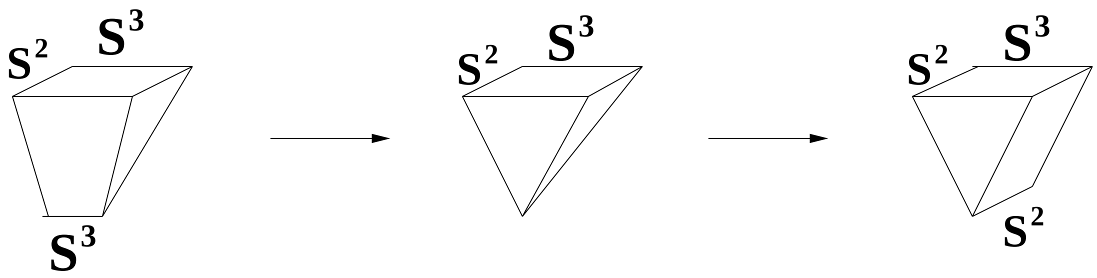


Figure 3.14: Topology change in the neighbourhood of a conifold singularity. Starting with a finite size $\mathbf{S^3}$ we tune a modulus to shrink it; at this stage massless state appear; a vev for them parametrizes growing an $\mathbf{S^2}$ out of the conifold singualrity.

considered a concept derived from more fundamental entities, like the field theory on the D-branes. This proposal generlizes to more general and less supersymmetric situations (like D-branes at singularities [32]).

This idea lies at the heart of some proposals like M(atrix) theory, which attempts at providing a microscopic definition of 11d M-theory [29]. The fundamental concept in M(atrix) theory is the worldvolume (worldline) action on a bunch of n type IIA D0-branes, in the limit $n \to \infty$. This is given by the dimensional reduction to $0+1$ dimensions of $d = 10$ $N = 1$ $U(n)$ super Yang-Mills.

In this approach, spacetime is obtained as the moduli space of the D0-brane gauge theory. Moreover, it is possible to reproduce supergravity interactions between objects by considering the dynamics of the $0+1$ gauge theory on configurations with slowly varying backgrounds for scalar fields (i.e. wavepackets slowly moving in spacetime). The arbitrariness in the number of D0-branes allows to explore arbitrarily high momentum in the M-theory dimension, and to recover 11d physics of M-theory.

Other applications of D-branes as probes includes throwing D-branes to diverse singularities of spacetime to see whether string theory can make sense of them. This approach has been successful in some cases, and has led to the understanding of certain naked singus in spacetime [33].

3.3.5 D-branes and gauge field theories

It is possible to take a low-energy limit in string theory in the presence of D-branes, which keeps all physical quantities of the worldvolume gauge field theory finite. In this limit the dynamics reduces to a quantum gauge field theory in $p+1$ dimensions, with gravity decoupled from it. Knowledge about perturbative and non-perturbative dynamics of string theory and D-branes can be used to explore or reproduce the dynamics of quantum gauge field theories. There are several examples of this, let us review two prototypical cases.

Montonen-Olive duality

One can use dualities of string theory to derive dualities in quantum gauge field theories. For instance, consider the 4d $N = 1$ supersymmetric $U(n)$ gauge theory obtained in the low-energy limit on a stack of n overlapping Type D3-branes. Gauge bosons and superspartners are obtained from open strings stretched between the different D3-branes. The gauge coupling is fixed by the string coupling $(g_{YM})^2 = g_s$.

Type IIB theory has a dual description in terms of another type IIB theory with string coupling $1/g_s$. In the dual theory, our configuration is given by n D3-branes, so it is a $U(n)$ gauge theory but now with gauge coupling $g'_{YM} = 1/g_{YM}$. The original perturbative states, open strings between the original D3-branes, are mapped to D1-branes stretched between D3-branes; it is possible to see that they correspond to ‘tHooft Polyakov monopoles of the dual theory.

Hence $N = 4$ $U(n)$ super Yang-Mills has a strong-weak duality relating the theory with coupling g_{YM} and $1/g_{YM}$, and exchanging fundamental and solitonic degrees of freedom. This duality had been previously proposed from purely field theoretical considerations [34], but we see here that it follows easily from the conjectured self-duality of type IIB string theory.

AdS/CFT correspondence (Maldecena conjecture)

We have proposed two different descriptions for the same object, the Dp-brane; one in terms of a solution to the sugra equations of motion, the other in terms of open strings ending on a $(p + 1)$-dimensional hyperplane. In principle both describe the same dynamics.

The Maldecan conjecture proposes to take a low energy limit in these two descriptions and match the result. On one side, we recover 4d $N = 4$ super Yang-Mills, decoupled from gravity; on the supergravity side, the 3-brane solution becomes and $\mathrm{AdS}_5 \times S^5$ geometry. So the proposal by Maldacena [96] is that $N = 4$ $U(n)$ super Yang-Mills is completely equivalent to type IIB string theory in $\mathrm{AdS}_5 \times S^5$.

This is a striking statement, that a string theory is completely equivalent to a gauge field theory! In fact, a subtle features makes this statement less striking. String theory on the curved space $\mathrm{AdS}_5 \times S^5$ does not have an exactly solvable worldsheet theory, so we can study it only in the supergravity approximation, valid for small curvatures. This regime corresponds, in the language of the dual field theory, to the limit of large $\lambda = g_{YM}^2 N$, this is a strongly coupled regime; λ is known as the 't Hooft coupling, and 't Hooft indeed proposed that in the large λ regime gauge field theory should be described as a string theory [36]. Hence the tractable regime in string theory is mapped to an untractable regime in gauge theory (because of the strong coupling). On the other hand, the tractable regime in gauge theory (small N) maps to string theory in spaces with string scale curvatures, which is completely untractable. So no paradox arises in relating a gauge field theory and a full-fledged string theory.

This conjecture has led to many important insights into gauge field theories in the large N limit, using the dual supergravity as a computational tool. In cases with less susy than $N = 4$ one can show at a qualitative level some features of strongly coupled gauge theories like confinement, chiral symmetry breaking, etc.

3.4 Our world as a brane-world model

We conclude this discussion by mentioning what applications all these non-perturbative objects may have in constructing phenomenological models of our world. The main motivation is that branes provide us with a mechanism to generate non-abelian gauge symmetries very different from that in heterotic theory. In particular, it is possible to localize gauge interactions in a subspace of spacetime, while gravity is still able to feel full spacetime.

The brane world idea is that it may be possible to construct string/M theory models where all or some of the particles of the standard model are part of the gauge sector of some branes, and hence are unable to propagate in some directions transverse to the brane. On the other hand, gravity would still be able to propagate on such directions.

There are basically two scenarios where this can be realized in string theory.

Horava-Witten phenomenology

The first is the Horava-Witten theory, which already before compactification has E_8 gauge interactions localized on 10d subspaces in an 11d world.

In order to build a phenomenological model, one may operate in a manner similar to that in the weakly coupled heterotic. Namely, compactify six of the ten dimensions in a Calabi-Yau manifold, endowed with some internal background for some of the E_8 gauge bosons. This configuration leads to 4d gravitational interactions and gauge interactions (with a gauge group determined by the internal gauge background), plus several families of charged chiral fermions.

Most of the phenomneology is similar to that in weakly coupled heterotic theory, except for hte choice of fundamental scale. As we discuss later on, the existence of one direction transverse to all gauge interactions allows to lower the fundamental scale below the 4d Planck scale. A nice choice in this context is to take the fundamental scale (11d Planck length to be around the gut scale 10^{16} GeV). This scenario was proposed in [37], and explored in many subsequent papers.

D-brane worlds

This possibility has been considered in [106] and many subsequent papers. It corresponds to considering compactifications of type II or type I theories on say a Calabi-Yau manifold X_6, with D-branes spanning four-dimensional Minkowski space and wrapped on a submanifold of X_6.

The simplest possibility would be to consider the standard model to be embedded in the volume of a D3-brane sitting at a point in X_6. Other possibilities would be to consider it to be embedded in a D5-brane whose worldovlume spans 4d Minkowski space and wraps a 2-cycle in X_6. The situation is shown in fig 3.15. In general Dp-brane leads to a 4d gauge sector if it wraps a $(p-3)$-dimensional submanifold Σ of X_6.

In principle, compactification in X_6 leads to 4d gravity; on the other hand, the gauge sector on the D-brane is also compactified on Σ and leads to 4d gauge sector. One has to work rather hard to construct configurations of D-branes whose open string sector leads to something like the standard model, but this has been achieved in several ways. We will skip these details here.

This kind of construction allows to build models where the fundamental string scale is not of the order of the 4d Planck mass, and can in fact be

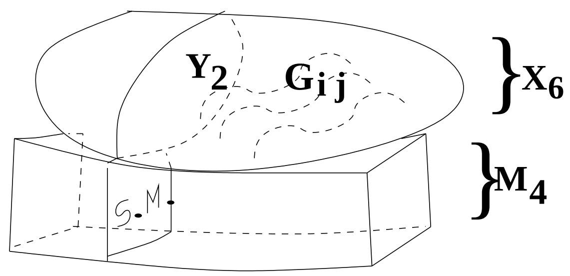


Figure 3.15: Schematic picture of a brane-world construction, with the Standard Model localized on the volume of e.g D5-branes with worldvolume $M_4 \times Y_2$, with Y_2 a compact submanifold of X_6.

much lower (in order to be consistent with experiment, it cannot be lower than a few TeV. The largeness of M_P can be generated if the compactification manifold is very large, so that gravity gets diluted. On the other hand, we should keep the internal directions along the brane of small to avoid too light KK replicas of Standard Model particles ($M_c \leq$ TeV along directions in Y_2 in the picture). However, constraints on the size of the directions in X_6 transverse to the brane (which are felt only gravitaionaly) are very mild, and such size can be as large as 0.1 mm.

More quantitatively, before compactification gravitational and gauge interactions are described by an effective action

$$\int d^{10}x \, \frac{M_s^8}{g_s^2} \, R_{10d} \, + \int d^{p+1}x \, \frac{M_s^{p-3}}{g_s} \, F_{(p+1)d}^2 \tag{3.16}$$

where the powers of g_s follow from the Euler characteristic of the worldsheet which produces the propagator of gravitons (sphere) and gauge bosons (disk), while the powers of M_s are fixed by dimensional analysis.

Upon compactification, the 4d action picks us a volume factor, as we saw in the discussion of KK compactification, and reads

$$\int d^4x \, \frac{M_s^8 V_{X_6}}{g_s^2} \, R_{4d} \, + \int d^4x \, \frac{M_s^{p-3} V_\Sigma}{g_s} \, F_{4d}^2 \tag{3.17}$$

This allows to read off the 4d Planck mass and gauge coupling, which are experimentally measured.

$$M_P^2 \;\; = \;\; \frac{M_s^8 V_{X_6}}{g_s^2} \simeq 10^{19}\,\mathrm{GeV}$$

$$1/g_{YM}^2 \quad = \quad \frac{M_s^{p-3} V_\Sigma}{g_s} \simeq 0.1 \tag{3.18}$$

If the geometry is factorizable, we can split $V_{X_6} = V_\Sigma V_{trans}$, with V_{trans} the transverse volume. One therefore obtains

$$M_P^2 \, g_{YM}^2 \; = \; \frac{M_s^{11-p} V_{trans}}{g_s} \tag{3.19}$$

This shows that it is possible to generate a large Planck mass in 4d with a low string scale, by simply increasing the volume transverse to the brane.

This allows to rephrase the hierarchy problem in geometric terms. The fundamental string scale could be close to the weak scale, around a few TeV, and the 4d Planck scale could be a derived scale arising from a large transversal volume.

It is important however, that having a low string scale is a possibility, not a necessity, in the brane world picture. However, it is an exciting possibility to provide new realizations of theories similar to our standard model within the framework of string theory.

Whether it is heterotic string theory or a brane-world scenario the way in which string theory is realized in Nature (if any of these mechanisms, there may be other ways not known to us for the moment), it is matter of experiment for coming generations of experiments. For the moment, we should be happy enough with the possibility of realizing such rich theories into a beatiful structure such as string theory.

Chapter 4

Quantization of the closed bosonic string

In this lecture we obtain the spectrum of oscillations of the closed bosonic string.

4.1 Worldsheet action

For this discussion I closely follow section 1.2 of [55]

As a string evolves in time, it sweeps out a two-dimensional surface in spacetime Σ, known as the worldsheet, and which is the analog of the worldline of a point particle in spacetime. Closed string correspond to worldsheets with no boundary, while open string sweep out worldsheets with boundaries. Any point in the worldsheet is labeled by two coordinates, t the 'time' coordinate just as for the pointparticle worldline, and σ, which parametrizes the extended spatial dimension of the string at fixed t. We denote σ, t collectively as ξ^a, $a = 1, 2$.

Our pupose is to write down the action for a string configuration in flat D-dimensional Minkowski space. For the bosonic string, such configurations are in principle described by D embedding functions $X^\mu(\sigma, t)$, with $\mu = 0, \ldots, D-1$, which can be regarded as 2d fields on the worldsheet.

4.1.1 The Nambu-Goto action

The natural action for a string configuration is the integral of the area element on the worldsheet, in principle measured with the metric inherited from the ambient metric in M_D. The ambient metric is computed as follows

$$\begin{aligned} ds^a &= h_{ab}\, d\xi^a\, d\xi^b \\ ds^2 &= \eta_{\mu\nu}\, dX^\mu\, dX^\nu \;=\; \eta_{\mu\nu}\, \frac{\partial X^\mu}{\partial \xi^a}\, \frac{\partial X^\nu}{\partial \xi^b}\, d\xi^a\, d\xi^b \end{aligned} \tag{4.1}$$

hence

$$h_{ab} \;=\; \eta_{\mu\nu}\, \frac{\partial X^\mu}{\partial \xi^a}\, \frac{\partial X^\nu}{\partial \xi^b} \tag{4.2}$$

The Nambu-Goto action is

$$S_{\rm NG}[X(\xi)] \;=\; -\frac{1}{2\pi\alpha'} \int_\Sigma d^2\xi\, (-h)^{1/2} \tag{4.3}$$

where $h = det(h_{ab})$ and α' is related to the string tension $T = \frac{1}{2\pi\alpha'}$.

4.1.2 The Polyakov action

The Nambu-Goto action is not very convenient for quantizing the worldsheet theory. So we are going to replace it by another action, which is classically equivalent, but which is much more convenient for quantization, the Polyakov action.

To do that we introduce another degree of freedom on the worldsheet, a worldsheet metric $g_{ab}(\xi)$ which is in principle independent of the induced metric h_{ab}. The natural action on the worldsheet is then

$$S_{\rm P} \;=\; -\frac{1}{4\pi\alpha'} \int_\Sigma d^2\xi\, (-g)^{1/2}\, g^{ab}(\sigma, t)\, \partial_a X^\mu\, \partial_b X^\nu \eta_{\mu\nu} \tag{4.4}$$

with $g = det(g_{ab})$

Classical equivalence with the Nambu-Goto action follows from solving the equations of motion for g_{ab}, namely $\delta S/\delta g_{ab} = 0$. Using

$$\delta g \;=\; -g\, g_{ab}\, \delta g^{ab} \tag{4.5}$$

one gets

$$\begin{aligned}\delta_g S_P &= -\frac{1}{4\pi\alpha'}\int_\Sigma d^2\xi\,(-g)^{1/2}\,\delta g^{ab}\,[\,-\frac{1}{2}\,g_{ab}\,g^{cd}\partial_c X^\mu\,\partial_d X_\mu\,+\partial_a X^\mu\,\partial_b X_\mu\,] = \\ &= -\frac{1}{4\pi\alpha'}\int_\Sigma d^2\xi\,(-g)^{1/2}\,\delta g^{ab}\,[\,-\frac{1}{2}\,g_{ab}\,g^{cd}\,h_{cd}\,+\,h_{ab}\,] \qquad (4.6)\end{aligned}$$

The equations of motion read

$$h_{ab} = \frac{1}{2}\,g_{ab}\,g^{cd}\,h_{cd} \qquad (4.7)$$

Taking determinant

$$(-h)^{1/2} = \frac{1}{2}\,(-g)^{1/2}\,g^{cd}\,h_{cd} \qquad (4.8)$$

and replacing into (9.1) we get

$$S_P[X(\xi), g_{\text{clas}}(\xi)] = -\frac{1}{2\pi\alpha'}\int_\Sigma d^2\xi\,(-h)^{1/2} = S_{NG}[X(\xi)] \qquad (4.9)$$

4.1.3 Symmetries of Polyakov action

The action (9.1) has some important symmetries which we now discuss

1. D-dimensional Poincaré invariance.

$$\begin{aligned}X'^\mu(\xi) &= \Lambda^\mu_\nu\,X^\nu(\xi)\,+a^\mu \\ g'_{ab}(\xi) &= g_{ab}(\xi) \qquad (4.10)\end{aligned}$$

   It is a global symmetry from the worldsheet viewpoint.

2. Two-dimensional diffeomorphism invariance, namele coordinate reparametrization of the worldsheet.

$$\begin{aligned}\xi'^a &= \xi'^a(\xi) \\ X'^\mu(\xi') &= X^\mu(\xi) \\ g'_{ab}(\xi') &= \frac{\partial\xi^c}{\partial\xi'^a}\,\frac{\partial\xi^d}{\partial\xi'^b}\,g_{cd}(\xi) \qquad (4.11)\end{aligned}$$

   It is a local (i.e. ξ dependent) symmetry. The 2d fields $X^\mu(\xi)$ behave as scalars while $g_{ab}(\xi)$ is a 2-index tensor (metric).

3. Two-dimensional Weyl invariance

$$\begin{aligned} X'^{\mu}(\xi) &= X^{\mu}(\xi) \\ g'_{ab}(\xi) &= \Omega(\xi)\, g_{ab}(\xi) \end{aligned} \tag{4.12}$$

It is a local symmetry.

Weyl-related string configurations correspond to the *same* embedding of the world-sheet in spacetime. So this is an extra redundancy in the Polyakov description, not present in the Nambu-Goto description.

It is convenient to emphasize at this point that a commonly mentioned symmetry, conformal invariance, is a subset of these symmetries. In particular, in covariant quantization one fixes the so-called conformal gauge, which amounts to using diff and Weyl invariances to set $g_{ab} = \eta_{ab}$. There is then a left-over local symmetry which is the set of coordinate transformations, whose effect on the metric can be undone with a Weyl transformation (so that the gauge fixed flat metric is preserved). This set of transformations is the 2d conformal group, which is extremely important in string theory. However, we will quantize the string in a different gauge, and conformal symmetry will not be manifest.

4.2 Light-cone quantization

For this section, we follow the computations in sections 1.3 and 1.4 of [55]. A more detailed treatment, using the formalism of quantization of constrained systems can be found in [39].

In quantizing the 2d field theory, we need to fix the gauge freedom. The light-cone gauge is the simplest one, and the most convenient to obtain the spectrum. This is because the final states will be the physical states of the theory, and in particular spacetime gauge particles will arise in the unitary gauge (namely, we will obtain only the two physical polarization modes of massless gravitons or gauge particles, and no spacetime spurious gauge degrees of freedom).

4.2.1 Light-cone gauge fixing

Define the light-cone coordinates

$$X^{\pm} = \frac{1}{\sqrt{2}}\,(X^0 \pm X^1)$$

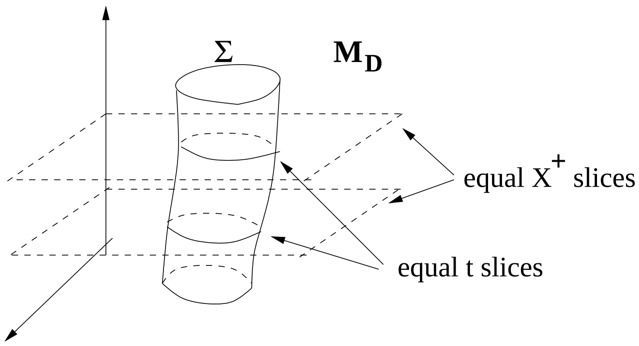


Figure 4.1: The light cone condition defines equat t slices on the worldsheet in terms of equal X^+ slices on spacetime.

$$X^i \qquad\qquad i = 2, \ldots, D-1 \tag{4.13}$$

The metric (scalar product) in M_D then reads

$$A^\mu B_\mu = -A^+ B^- - A^- B^+ + A^i B^i \tag{4.14}$$

so

$$A_- = -A^+ \quad , \quad A_+ = -A^- \quad , \quad A_i = A^i \tag{4.15}$$

The gauge fixing proceeds through several steps

1. Reparametrization of t

Fix the t reparametrization freedom by setting the so-called light-cone condition

$$X^+(\sigma, t) = t \tag{4.16}$$

see figure 4.1. So X^+ will play the role of worldsheet time, and its conjugate variable $P_+ = -P^-$ will play the role of worlsheet energy (2d hamiltonian).

2. Reference line in σ

Choose a line on the worldsheet $\sigma_0(t)$ intersecting all constant t slices orthogonally (w.r.t. the 2d metric g). Namely

$$g_{t\sigma}(\sigma, t) = 0 \qquad \text{at } \sigma = \sigma_0(t) \tag{4.17}$$

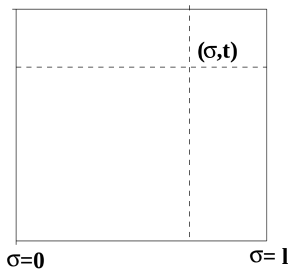


Figure 4.2: The coordinate t on the worldsheet corresponds to the coordinate X^+ of the spacetime point where it is embedded. The coordinate σ is defined as the invariant distance, to a reference line $\sigma = 0$, along fixed t slices. The total string length is fixed to be ℓ.

Notice that this still leaves the freedom of an overall motion of the reference line. This will be important as an additional constraint on the final spectrum (see (4.43)).

3. Reparametrization of σ

For slices of constant t, define a new spatial coordinate σ' for each point of the slice. σ' is defined as the (diffeomorphism and Weyl) invariant distance to the reference line along the slice

$$\sigma' = c(t) \int_{\sigma_0}^{\sigma} f(\sigma, t)\, d\sigma \tag{4.18}$$

where

$$f(\sigma) = (-g)^{-1/2}\, g_{\sigma\sigma}(\sigma, t) \tag{4.19}$$

and $c(t)$ is a σ independent coefficient used to impose that the total length of the string is fixed, a constant in t which we call ℓ. The situation is shown in figure 4.2.

In the new coordinates, $f(\sigma')$ is σ' independent. In the following we will only use this coordinatization, and we drop the prime. So we write

$$\partial_\sigma f(\sigma, t) = 0 \tag{4.20}$$

4. Weyl invariance

Now we use Weyl invariance to impose that

$$g = -1 \qquad \forall \sigma, t \tag{4.21}$$

Since $f(\sigma)$ is Weyl-invariant, it still satisfies $\partial_\sigma f(\sigma,t)=0$. Using the definition of f, we get

$$\partial_\sigma g_{\sigma\sigma} = 0 \tag{4.22}$$

This concludes the gauge fixing. The metric and inverse metric read

$$(g_{ab}) = \begin{pmatrix} g_{\sigma\sigma}(t)^{-1}[-1+g_{t\sigma}(\sigma,t)^2] & g_{t\sigma}(\sigma,t) \\ g_{t\sigma}(\sigma,t) & g_{\sigma\sigma}(t) \end{pmatrix} \; ; \; (g^{ab}) = \begin{pmatrix} -g_{\sigma\sigma}(t) & g_{t\sigma}(\sigma,t) \\ g_{t\sigma}(\sigma,t) & g_{\sigma\sigma}(t)^{-1}[1-g_{t\sigma}(\sigma,t)^2] \end{pmatrix}$$

4.2.2 Gauge-fixed Polyakov action, Hamiltonian

The Polyakov lagrangian in light-cone coordinates reads

$$\begin{aligned} L &= -\tfrac{1}{4\pi\alpha'}\textstyle\int_0^\ell d\sigma \quad [-2\,g^{tt}\partial_t X^+\,\partial_t X^- + g^{tt}\partial_t X^i\,\partial_t X^i - 2\,g^{\sigma t}\,\partial_t X^+\,\partial_\sigma X^- + \\ &\qquad +2\,g^{\sigma t}\,\partial_\sigma X^i\,\partial_t X^i + g^{\sigma\sigma}\,\partial_\sigma X^i\,\partial_\sigma X^i\,] = \\ &= -\tfrac{1}{4\pi\alpha'}\textstyle\int_0^\ell d\sigma \quad [\,g_{\sigma\sigma}\,(2\,\partial_t X^- - \partial_t X^i\,\partial_t X^i) - 2\,g_{\sigma t}\,(\partial_\sigma X^- - \partial_\sigma X^i\,\partial_t X^i) + \\ &\qquad g_{\sigma\sigma}^{-1}\,(1-g_{\sigma t}^2)\,\partial_\sigma X^i\,\partial_\sigma X^i\,] \end{aligned} \tag{4.23}$$

Defining the center of mass and relative coordinates $x^-(t)$, $Y^-(\sigma,t)$

$$\begin{aligned} x^-(t) &= \frac{1}{\ell}\int_0^\ell d\sigma\, X^-(\sigma,t) \\ X^-(\sigma,t) &= x^-(t) + Y^-(\sigma,t) \end{aligned} \tag{4.24}$$

we obtain

$$\begin{aligned} L &= -\frac{\ell}{2\pi\alpha'}\,g_{\sigma\sigma}\,\partial_t x^-(t) - \frac{1}{4\pi\alpha'}\int_0^\ell d\sigma\,[-g_{\sigma\sigma}\,\partial_t X^i\,\partial_t X^i + \\ &\quad -2\,g^{\sigma t}\,(\partial_\sigma Y^- - \partial_\sigma X^i\,\partial_t X^i) + g_{\sigma\sigma}^{-1}\,(1-g_{\sigma t}^2)\,\partial_\sigma X^i\,\partial_\sigma X^i\,] \end{aligned} \tag{4.25}$$

The $Y^-(\sigma,t)$ does not have time derivatives in this lagrangian, so it acts as a Lagrange multiplier imposing

$$\partial_\sigma g_{\sigma,t}(\sigma,t) = 0 \quad \forall \sigma, t \tag{4.26}$$

Since we have $g_{\sigma t}(\sigma=0,t)=0$ due to (4.17), we get

$$g_{\sigma,t}(\sigma,t) = 0 \quad \forall \sigma, t \tag{4.27}$$

The lagrangian becomes

$$L \;=\; -\frac{\ell}{2\pi\alpha'}\, g_{\sigma\sigma}\, \partial_t x^-(t) \;+\; \frac{1}{4\pi\alpha'} \int_0^\ell d\sigma\, [\, g_{\sigma\sigma}\, \partial_t X^i\, \partial_t X^i \;-\; g_{\sigma\sigma}^{-1}\, \partial_\sigma X^i\, \partial_\sigma X^i\,]$$

The momentum conjugate to $x^-(t)$ is

$$p_- \;=\; -p^+ = \frac{\partial L}{\partial(\partial_t x^-)} \;=\; -\frac{\ell}{2\pi\alpha'}\, g_{\sigma\sigma} \tag{4.28}$$

so $g_{\sigma\sigma}$ is not really an independent coordinate variable, but a momentum variable.

The momenta conjugate to $X^i(\sigma, t)$ are

$$\Pi^i(\sigma,t) \;=\; \frac{\partial \mathcal{L}}{\partial(\partial_t X^i)} \;=\; \frac{1}{2\pi\alpha'}\, g_{\sigma\sigma}\, \partial_t X^i(\sigma,t) \;=\; \frac{p^+}{\ell}\, \partial_t X^i(\sigma,t) \tag{4.29}$$

We can construct the Hamiltonian

$$\begin{aligned}
H \;&=\; p_- \partial_t x^-(t) \;+\; \int_0^\ell d\sigma\, \Pi_i(\sigma,t)\, \partial_t X^i(\sigma,t) \;-\; L \;= \\
&=\; -\frac{\ell}{2\pi\alpha'}\, g_{\sigma\sigma}\, \partial_t x^-(t) \;+\; \int_0^\ell d\sigma\, \frac{1}{2\pi\alpha'}\, g_{\sigma\sigma}\, \partial_t X^i(\sigma,t)\, \partial_t X^i(\sigma,t) \;+ \\
&\quad +\frac{\ell}{2\pi\alpha'}\, g_{\sigma\sigma}\, \partial_t x^-(t) \;-\; \frac{1}{4\pi\alpha'} \int_0^\ell d\sigma\, [\, g_{\sigma\sigma}\, \partial_t X^i\, \partial_t X^i \;-\; g_{\sigma\sigma}^{-1}\, \partial_\sigma X^i\, \partial_\sigma X^i\,] \;= \\
&=\; \frac{1}{4\pi\alpha'} \int_0^\ell d\sigma\, [\, g_{\sigma\sigma}\, \partial_t X^i\, \partial_t X^i \;+\; g_{\sigma\sigma}^{-1}\, \partial_\sigma X^i\, \partial_\sigma X^i\,] \;=
\end{aligned} \tag{4.30}$$

In terms of momenta

$$H \;=\; \frac{\ell}{4\pi\alpha' p^+} \int_0^\ell d\sigma\, [\, 2\pi\alpha'\, \Pi_i\, \Pi_i \;+\; \frac{1}{2\pi\alpha'}\, \partial_\sigma X^i\, \partial_\sigma X^i\,] \tag{4.31}$$

The equations of motion for x^-, $p_- = p^+$ are

$$\begin{aligned}
\partial_t x^-(t) \;&=\; \frac{\partial H}{\partial p_-} \;=\; -\frac{\partial H}{\partial p^+} \;=\; \frac{H}{p^+} \\
\partial_t p^+(t) \;&=\; -\frac{\partial H}{\partial x^-} \;=\; 0
\end{aligned} \tag{4.32}$$

so p^+ is conserved, and x^- is linear in t and has trivial dynamics.

The equations of motion for X^i, Π_i are

$$\begin{aligned} \partial_t X^i(\sigma,t) &= \frac{\delta H}{\delta \Pi_i} = c\, 2\pi\alpha' \,\Pi_i \\ \partial_t \Pi_i(\sigma,t) &= -\frac{\delta H}{\delta X^i} = \frac{c}{2\pi\alpha'}\,\partial_\sigma^2 X^i \end{aligned} \tag{4.33}$$

with $c = \ell/(2\pi\alpha' p^+)$ So we get

$$\partial_t^2 X^i = c^2\,\partial_\sigma^2 X^i \tag{4.34}$$

the wave equation for two-dimensional fields $X^i(\sigma,t)$. Indeed, for fixed (because it is conserved) p^+, we see that H is the hamiltonian for $D-2$ *free* bosons in 2d [1].

It is useful to set $\ell = 2\pi\alpha' p^+$, and so $c = 1$.

4.2.3 Oscillator expansions

The general solution to the equations of motion is a superposition of left- and right-moving waves

$$X^i(\sigma,t) = X_L^i(\sigma+t) + X_R^i(\sigma-t) \tag{4.35}$$

For closed strings, we need to impose boundary conditions, periodicity in σ

$$X^i(\sigma+\ell,t) = X^i(\sigma,t) \tag{4.36}$$

The general form of X_L, X_R with those boundary conditions is

$$\begin{aligned} X_L^i(\sigma+t) &= \frac{x^i}{2} + \frac{p_i}{2p^+}(t+\sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbf{Z}-\{\mathbf{0}\}} \frac{\alpha_n^i}{n}\, e^{-2\pi i\, n\,(\sigma+t)/\ell} \\ X_R^i(\sigma-t) &= \frac{x^i}{2} + \frac{p_i}{2p^+}(t-\sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbf{Z}-\{\mathbf{0}\}} \frac{\tilde{\alpha}_n^i}{n}\, e^{2\pi i\, n\,(\sigma-t)/\ell} \end{aligned} \tag{4.37}$$

The coefficients x^i, p_i denote the center of mass coordinate and momentum, while the two infinite sets of coeffients α_n^i, $\tilde{\alpha}_n^i$ denote the amplitudes of the momentum n mode for left and right movers.

[1] Recalling our discussion about the α' expansion, this means that we can quantize the theory exactly in α'.

Promoting the worldsheet degrees of freedom $x^-(t)$, p^+, X^i, Π^i to operators, with canonical commutators, we obtain the commutation relations

$$\begin{aligned} \left[x^-, p^+\right] &= -i \\ \left[x^i, p_j\right] &= i\delta_{ij} \\ \left[\alpha^i_m, \alpha^j_n\right] &= \left[\tilde{\alpha}^i_m, \tilde{\alpha}^j_n\right] = m\,\delta_{ij}\,\delta_{m,-n} \\ \left[\alpha^i_m, \tilde{\alpha}^j_n\right] &= 0 \end{aligned} \tag{4.38}$$

We can obtain the hamiltonian in terms of these

$$\begin{aligned} H &= \frac{1}{2}\int_0^\ell d\sigma\,[\,2\pi\alpha'\,\Pi^i\,\Pi^i + \frac{1}{2\pi\alpha'}\,\partial_\sigma X^i\,\partial_\sigma X^i\,] = \\ &= \frac{p_i p_i}{2p^+} + \frac{1}{\alpha' p^+}\left[\sum_{n>0}[\,\alpha^i_{-n}\alpha^i_n + \tilde{\alpha}^i_{-n}\tilde{\alpha}^i_n\,] + E_0 + \tilde{E}_0\right] \end{aligned} \tag{4.39}$$

We get the quantum mechanics of the center of mass motion and two infinite sets of decoupled harmonic oscillators. Here we have normal-ordered the creation and annihilation modes and E_0, $\tilde{E}_0$ are the corresponding zero point energies, to be discussed below.

The Hilbert space of string states is obtained by defining a vacuum $|k\rangle = |k_-, k_i\rangle$ by

$$p^+|k\rangle = k_-|k\rangle \quad , \quad p_i|k\rangle = k_i|k\rangle \quad , \quad \alpha^i_n|k\rangle = \tilde{\alpha}^i_n|k\rangle = 0 \quad \forall n > 0 \tag{4.40}$$

and acting on it with the creation ladder operators α^i_{-n}, $\tilde{\alpha}^i_{-n}$, with $n > 0$, in an arbitrary way (almost, see later for an additional constraint).

As discussed in the overview lectures, each oscillation state of the string is observed as a particle from the spacetime viewpoint, with spacetime mass

$$M^2 = -p^2 = 2p^+p^- - p_i p_i \tag{4.41}$$

Notice that p^- corresponds to ∂_{x^+}, which in light cone gauge is ∂_t, which corresponds to the 2d hamiltonian H, so $p^- = H$, and $M^2 = 2p^+H - p_i p_i$. We have

$$\alpha'\,M^2 = N + \tilde{N} + E_0 + \tilde{E}_0 \tag{4.42}$$

with $N = \sum_{n>0}\alpha^i_{-n}\alpha^i_n$ the total left oscillator number (analogously for $\tilde{N}$). It is important to recall from the commutation relations, that a single mode α^i_n or $\tilde{\alpha}^i_n$ contributes n to the oscillator number.

Hence the masses of spacetime particles increase with the number of oscillators in the corresponding string state.

There is one further constraint we must impose on the spectrum. Recall that after gauge fixing we still had the freedom to perform an overall translation of the reference line $\sigma = 0$ by a t independent amount. This forces to restrict the spectrum to the subsector invariant under translations in σ. This amounts to requiring the net 2d momentum along σ to vanish, namely the left- and right-moving operators in a state should carry the same total momentum. Recalling that a mode n carries momentum n, the constraint is

$$N = \tilde{N} \tag{4.43}$$

the so-called level matching constraint. It is an important fact that the quantization procedure can be performed independently for left- and right- movers (e.g. defining left- and right-moving hamiltonians, and mass operators, etc) and they only talk to each other at the level of building the physical spectrum via the constraint (4.43).

Finally, we need to compute the zero point energies $E_0 = \tilde{E}_0$. Formally, for each i

$$E_0^i = \frac{1}{2} \sum_{n=1}^{\infty} n \tag{4.44}$$

This is infinite so we compute it with a regularization prescription, i.e. as the limit $\epsilon \to 0$ of the non-singular part of

$$Z(\epsilon) = \frac{1}{2} \sum_{n=1}^{\infty} n\, e^{-n\epsilon} \tag{4.45}$$

After some massage

$$Z(\epsilon) = \frac{1}{2} \sum_{n=1}^{\infty} n\, e^{-n\epsilon} = -\frac{1}{2} \frac{d}{d\epsilon} \sum_{n=1}^{\infty} e^{-n\epsilon} = -\frac{1}{2} \frac{d}{d\epsilon} \frac{1}{1 - e^{-\epsilon}} \tag{4.46}$$

Since

$$\begin{aligned} \frac{1}{1 - e^{-\epsilon}} &= \frac{1}{\epsilon} \frac{1}{1 - \epsilon/2 + \epsilon^2/6 + \mathcal{O}(\epsilon^3)} = \frac{1}{\epsilon} [1 + \epsilon/2 - \epsilon^2/6 + \epsilon^2/4 + \mathcal{O}(\epsilon^3)] = \\ &= \frac{1}{\epsilon} + \frac{1}{2} + \frac{1}{12}\epsilon + \mathcal{O}(\epsilon^2) \end{aligned} \tag{4.47}$$

we get

$$Z(\epsilon) \;=\; -\frac{1}{2}\,[\,-\frac{1}{\epsilon^2}\,+\,\frac{1}{12}\,+\mathcal{O}(\epsilon)\,] \tag{4.48}$$

Dropping the infinite part and letting $\epsilon \to 0$, the zero point energy for a single 2d free boson is

$$E_0^i = \tilde{E}_0^i = -\frac{1}{24} \tag{4.49}$$

So for $D-2$ we have $E_0 = \tilde{E}_0 = -(D-2)/24$

$$\alpha'\,M^2 \;=\; N \,+\, \tilde{N} \,-\, 2\,\frac{D-2}{24} \tag{4.50}$$

Dropping the infinity amounts to redefining the vacuum energy. One might think that this is not possible because the Polyakov action includes a worldsheet metric (i.e. gravity). However, this is not present in our gauge fixing and the problem is avoided. It is important to emphasize that this infinity is not present in other gauge fixings (like the conformal gauge), so the infinity is an artifact of our gauge fixing. However, the zero point energy we have computed has physical consequences, like fixing the dimension of spacetime to be 26. In the light-cone gauge, which is not manifestly Lorentz invariant, it appears when we require the spectrum to be Lorentz invariant, as we motivate below. In other gauges, the condition appears in other ways. For instance, in the conformal gauge fixing, as the cancellation of the conformal anomaly.

For $D=26$ we have

$$\alpha'\,M^2 \;=\; N \,+\, \tilde{N} \,-\, 2 \tag{4.51}$$

4.2.4 Light spectrum

It is now time to obtain the lightest particles in the spectrum of the string. The states with smallest number of oscillators that we can construct satisfying (4.43) are

$$\begin{array}{lcl} N=\tilde{N}=0 & |k\rangle & \alpha' M^2 = -2 \\ N=\tilde{N}=1 & \alpha_{-1}^i\alpha_{-1}^j|k\rangle & \alpha' M^2 = 0 \end{array} \tag{4.52}$$

The closed string groundstate is a spacetime tachyon. This field is troublesome, and it is thought to signal an instability of the theory. The result of this instability is not known.

The second states transform as a two-index tensor with respect to the $SO(D-2)$ subgroup of the Lorentz group manifest in the light-cone gauge.

One should recall that in a Lorentz invariant theory in D dimensions, physical states of fields belong to representations of the so-called little group (subgroup of Lorentz group which leaves invariant the D-momentum of the particle). For massive particles, the D-momentum can be brought to the form $P = (M, 0, \ldots, 0)$ in the particle's rest frame, so the little group is $SO(D-1)$. For massless particles, the D-momentum can be brought to the form $(M, M, 0, \ldots)$, so the little group is $SO(D-2)$.

Our particles in the first excited sector are clearly not enough to fill out a representation of $SO(D-1)$, so to have Lorentz invariance it is crucial that they are massless. Notice that this is so only because we have imposed $D = 26$, so this is a derivation of the dimension of spacetime in which string theory can propagate consistently. Indeed, it is possible to construct the Lorentz generators in terms of the oscillator numbers etc and check that the Lorentz algebra is recovered only if $D = 26$. We skip this computation which can however be found in standard textbooks, like [39]

Let us also point out that massive states in the theory *do* fill out representations of $SO(D-1) = SO(25)$, altough only $SO(24)$ is manifest.

The massless two-index tensor can be split in irreducible representations of $SO(24)$, by taking its trace (which is a 26d scalar particle, the dilaton ϕ), its antisymmetric part (which is a 26d 2-form field $B_{\mu\nu}$) and its symmetric traceless part (which is a 26d symmetric tensor field $G_{\mu\nu}$).

4.2.5 Lessons

The result of light cone quantization for the bosonic string can be phrased in terms of the following recipe, which will be valid for other string theories as well

- The only relevant degrees of freedom left are the center of mass and $D-2$ transverse coordinates $X^i(\sigma, t)$, $i = 2\ldots, D-1$

- For closed string theories the 2d theory splits into two sectors, left- and right-movers, which can be quantized independently. The only relation between them appears at the final stage, when imposing the level matching condition on the physical spectrum.

- The spacetime $(\text{mass})^2$ operator (on each sector) is given by the oscillator numbers plus the zero point energy, which should be computed using the $e^{-\epsilon n}$ regularization.

4.2.6 Final comments

Upon studying interactions of these 26d fields one concludes that $G_{\mu\nu}$ is a 26d graviton and $B_{\mu\nu}$ is a gauge potential. So 26d interactions between these fields are invariant under 26d coordinate reparametrization and gauge transformations for B

$$B_{\mu\nu} \to B_{\mu\nu} + \partial_{[\mu}\Lambda_{\nu]}(X) \tag{4.53}$$

The 26d low energy effective action for these modes was described in the overview lessons. In the string frame

$$S_{\text{eff.}} = \frac{1}{2k^2}\int d^{26}X\,(-\tilde{G})^{1/2}\,\{\,\tilde{R} + \frac{1}{12}e^{-\tilde{\phi}/12}\,H_{\mu\nu\rho}\,H^{\mu\nu\rho} - \frac{1}{6}\partial_\mu\tilde{\phi}\partial^\mu\tilde{\phi}\,\} + \mathcal{O}(\alpha') \tag{4.54}$$

We emphasize again that the dilaton vev fixes the string interaction coupling constant in the 26d theory. So the string interaction coupling constant is not an arbitrary external parameter, but the vacuum expectation value of a spacetime dynamical scalar field in the theory. Instead of a continuum of different string theories, labeled by the value of the coupling constant, we have a unique string theory with a continuous set of vacua parametrized by the vev for a scalar field with flat potential $V(\phi) \equiv 0$. Fields with flat potential are called moduli, and the set of vacua is called the moduli space of the theory.

Chapter 5

Modular invariance

5.1 Generalities

In this Section we mainly follow the line of thought of section 7.3 in [55]. Our computation is however done in the light-cone gauge.

In this lecture we discuss the simplest case where we can witness the remarkable finiteness properties of string theory. The example is provided by the 1-loop vacuum amplitude. It corresponds to a worldsheet diagram for a closed string moving in a circle and closing onto itself, so it has the topology of a two-torus with no insertions of external lines. It represents the 1-loop amplitude of the vacuum going to vacuum process (in spacetime). See figure 5.1

We know from the overview lectures that the amplitude is obtained by summing over *all possible inequivalent* worldsheet geometries with two-torus topology.

It is crucial to incorporate all possible geometries, and not to double-count equivalent geometries. Concerning this, it is extremely important to realize that a given geometry can receive two different interpretations. A diagram corresponding to a two-torus with circle lengths ℓ_1 and ℓ_2 can be regarded as

1) A closed string of length ℓ_1 propagating over a distance ℓ_2

2) A closed string of length ℓ_2 propagating over a distance ℓ_1

The two processes, although look different, correspond to the same geometry, so should be counted only *once*. This will be crucial later on.

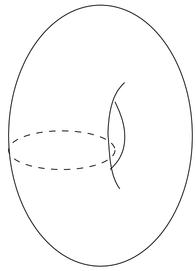


Figure 5.1: One-loop diagram for the vacuum going to vacuum process.

5.2 Worldsheet coordinatization in light-cone gauge

Recall our recipe to compute amplitudes. First we sum over geometries of an abstract worldsheet Σ with two-torus topology. Second, for each such geometry we sum over possible configurations of the 2d dynamical fields in Σ (in the light cone gauge, the transverse fluctuations $X^i(\sigma, t)$).

Recall that in the light cone gauge we have **1)** a coordinate σ which parametrizes a direction of fixed length ℓ; **2)** a coordinate t which is locally orthogonal to σ at every point; **3)** a Hamiltonian for the physical degrees of freedom, generating evolution in t for the 2d system. In terms of oscillator and center of mass momentum

$$H = \frac{\sum_i p_i^2}{2p^+} + \frac{1}{\alpha' p^+} \left[L_0 + \tilde{L}_0 \right] \tag{5.1}$$

with

$$L_0 = \sum_i \left[\sum_{n>0} \alpha_{-n}^i \, \alpha_n^i + E_o^i \right] \quad , \quad \text{and } E_0^i = -\frac{1}{24} \tag{5.2}$$

and similarly for $\tilde{L}_0$.

A two-torus can be described as the two-dimensional real plane, modded out by translations by vectors in a two-dimensional lattice, see figure 13.3

There is a more or less obvious set of worldsheet geometries which we should consider. It is shown in figure 5.3a), and corresponds to a closed string (of σ-length ℓ) evolving for $t = \tau_2 \ell$ (for $\tau_2 > 00$ and closing back onto itself.

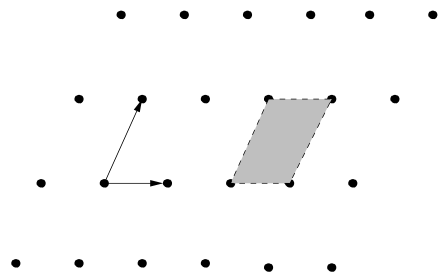


Figure 5.2: A two-torus can be constructed by modding out the two-dimensional plane by translations in a two-dimensional lattice. The unit cell is a parallelogram with sides identified. Each vector corresponds to a non-contractible cycle in the two-torus

Denoting $z = \sigma + i\,t$, the two-torus is defined by the identifications $z \equiv z + \ell$, $z \equiv z + \tau_2 \ell$.

However, there are more general possibilities, as shown in figure 5.3b), corresponding to a closed string of length ℓ evolving for $t = \tau_2 \ell$, and gluing back to the original state up to a change in the reference line $\sigma = 0$ (given by a translation by $\tau_1 \ell$ in the σ-direction). Since there is no preferred choice of the reference line, as discussed in the previous lecture, this is an allowed possibility. The geometry corresponds to a two-torus defined by the identifications $z \equiv z + \ell$ and $z \equiv z + \tau \ell$, with $\tau = \tau_1 + i\tau_2$. The parameter τ is called the complex structure of the two-torus, for reasons not very relevant here.

5.3 The computation

5.3.1 Structure of the amplitude in operator formalism

We have to sum over all possible configurations of 2d physical fields $X^i(\sigma, t)$ for a given 2d geometry. In operator formalism, this amounts to considering the complete set of quantum 2d states at a given time (i.e. the Hilbert space of the 2d theory), apply evolution in t for a total time of $t = \tau_2 \ell$ and glue the resulting state to the initial one (modulo a σ-translation by $\tau_1 \ell$). The amplitude for two-torus geometry corresponding to τ is therefore

$$Z(\tau) \,=\, \sum_{\text{states}} \langle \text{st.}|\, e^{-\tau_2 \ell H}\, e^{i\tau_1 \ell P}\, |\text{st.}\rangle \qquad (5.3)$$

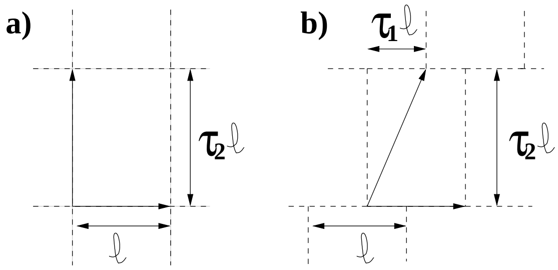


Figure 5.3: Figure a) shows an obvious class of worldsheet geometries with two-torus geometries, a closed string of length ℓ evolves for some time $t = \tau_2\ell$ and closes back to the initial state. Figure b) shows the more general class, where the closed string is glued back to the original state modulo a change in the reference line in σ.

where P is the generator of translations along σ

$$P \;=\; \int_0^\ell d\sigma\,\Pi_i\,\partial_\sigma X^i \;=\; \frac{2\pi}{\ell}\,(L_0 - \tilde{L}_0) \qquad (5.4)$$

(namely $\partial_\sigma X^i$ gives the amoung of X shift induced by the σ-translation, and Π_i implements the effect of the X shift on the Hilbert space).

The amplitude hence corresponds to taking a trace over the Hilbert space $\mathcal{H}_{\rm cl.}$ of the closed string 2d theory

$$\begin{aligned} Z(\tau) \;&=\; \mathrm{tr}_{\mathcal{H}_{\rm cl.}}\left(e^{-\tau_2\ell H}e^{i\tau_1\ell P}\right) \;= \\ &=\; \mathrm{tr}_{\mathcal{H}_{\rm cl.}}\left(\exp[-\tau_2 2\pi\alpha' p^+\,[\,\frac{\sum p_i^2}{2p^+} \;+\; \frac{1}{\alpha' p^+}(L_0+\tilde{L}_0)\,]\,\exp[2\pi i\tau_1(L_0-\tilde{L}_0)]\right) \;= \\ &=\; \mathrm{tr}_{\mathcal{H}_{\rm cl.}}\left(\exp[-\tau_2\pi\alpha'\sum p_i^2]\,\exp[2\pi i(\tau_1+i\tau_2)L_0]\,\exp[2\pi i(\tau_1-i\tau_2)\tilde{L}_0]\right) \;= \end{aligned} \qquad (5.5)$$

Defining $q = e^{2\pi i\tau}$, we have

$$Z(\tau) \;=\; \mathrm{tr}_{\mathcal{H}_{\rm cl.}}\left(\exp[-\tau_2\pi\alpha'\sum p_i^2]\,q^{L_0}\,\overline{q}^{\tilde{L}_0}\right) \qquad (5.6)$$

Then we should sum over geometries, i.e. integrate over τ. Notice that when we integrate over τ_1 the level-matching constraint $L_0 = \tilde{L}_0$ is automatically implemented

$$\int d\tau_1\,e^{2\pi i\tau_1(L_0-\tilde{L}_0)} \simeq \delta_{L_0,\tilde{L}_0} \qquad (5.7)$$

Hence, we can take the trace over the unconstrained set states constructed by applying arbitrary numbers of all possible left and right oscillators to the vacuum. Subsequently the sum over geometries will implement that only physical states, satisfying the level matching constraint, propagate.

Hence the general structure of the states we are tracing over is

$$\prod_{n,i} (\alpha^i_{-n})^{K_{n,i}} \prod_{m,j} (\tilde{\alpha}^j_{-m})^{\tilde{K}_{m,j}} \, |p_-, p_i\rangle \tag{5.8}$$

That is, the Hilbert space is given by a set of momentum states, on which we apply an arbitrary number of times K, $\tilde{K}$ oscillator creation operators out of an infinite set labeled by n, i, m, j.

5.3.2 The momentum piece

The trace over center of mass degrees of freedom give an overall factor independent of the oscillator ocuppation numbers $K_{n,i}$, $K_{m,j}$. Moreover, the center of mass trace factorizes as product of traces over different directions

$$tr_{\text{c.m.}} \, e^{-\tau_2 \pi \alpha' \sum_i p_i^2} \; = \; (\text{tr}_{\,\text{c.m.1d}} \, e^{-\tau_2 \pi \alpha' p^2})^{24} \tag{5.9}$$

For each direction, we can take the trace by summing over (center of mass) position eigenstates

$$\begin{aligned} \text{tr}_{\,\text{c.m.1d}} \, e^{-\tau_2 \pi \alpha' p^2} \; &= \; \textstyle\int dx \, \langle x | \, e^{-\tau_2 \pi \alpha' p^2} \, | x \rangle \; = \\ \textstyle\int dx \, \int \frac{dp}{2\pi} \langle x | p \rangle \, \langle p | \, e^{-\tau_2 \pi \alpha' p^2} \, | p \rangle \, \langle p | x \rangle \; &= \; (\textstyle\int dx) \, (4\pi^2 \alpha' \tau_2)^{-1/2} \end{aligned} \tag{5.10}$$

Hence

$$tr_{\text{c.m.}} \, e^{-\tau_2 \pi \alpha' \sum_i p_i^2} \; = \; V_{24} \, (4\pi^2 \alpha' \tau_2)^{-12} \tag{5.11}$$

where V_{24} is a regularized volume of the 24d transverse space.

5.3.3 The oscillator piece

The oscillator creation operators can be applied independently, so the trace factorizes in traces over the Hilbert space of each independent oscillator.

For a single oscillator, the trace over states $(\alpha^i_{-n})^K |0\rangle$ goes like

$$\begin{aligned} \text{tr} \, q^{\hat{N} + E_0} \; &= \; q^{E_0} \textstyle\sum_{K=0}^{\infty} \langle 0 | \, (\alpha^i_n)^K \, q^{\hat{N}} (\alpha^i_{-n})^K \, | 0 \rangle \; = \\ &= q^{E_0} \, \textstyle\sum_{K=0}^{\infty} q^{Kn} \; = \; q^{-1/24} \, \frac{1}{1 - q^n} \end{aligned} \tag{5.12}$$

For two oscillators, the trace over states $(\alpha_{-n_1})^{K_1}(\alpha_{-n_2})^{K_2}|0\rangle$ is

$$\begin{aligned}
\mathrm{tr}\; q^{\hat{N}+E_0} &= q^{-2/24} \sum_{K_1,K_2=0}^{\infty} \langle 0|\,(\alpha_{n_1})^{K_1}\,(\alpha_{n_2})^{K_2}\,q^{\hat{N}_1+\hat{N}_2}\,(\alpha_{-n_1})^{K_1}\,(\alpha_{-n_2})^{K_2}\,|0\rangle = \\
&= q^{-2/24} \sum_{K_1,K_2=0}^{\infty} \langle 0|\,(\alpha_{n_1})^{K_1}\,q^{\hat{N}_1}\,(\alpha_{-n_1})^{K_1}\;(\alpha_{n_2})^{K_2}\,q^{\hat{N}_2}\,(\alpha_{-n_2})^{K_2}\,|0\rangle = \\
&= q^{-2/24}\,(1-q^{n_1})^{-1}\,(1-q^{n_2})^{-1} \qquad (5.13)
\end{aligned}$$

So for the infinite set of left and right oscillators

$$\mathrm{Tr}\; q^{L_0}\,\overline{q}^{\tilde{L}_0} = q^{E_0}\,\overline{q}^{\tilde{E}_0} \prod_{i=2}^{26}\prod_{n=1}^{\infty}(1-q^n)^{-1} \prod_{j=2}^{26}\prod_{m=1}^{\infty}(1-q^m)^{-1} = \left| q^{1/24} \prod_{n=1}^{\infty}(1-q^n)\right|^{-48} \quad (5.14)$$

Using the definition of the Dedekind eta function (A.2)

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty}(1-q^n) \qquad (5.15)$$

the complete partition function, for fixed τ, is

$$Z(\tau) = V_{24}\,(4\pi^2\alpha'\tau_2)^{-12}\,|\eta(\tau)|^{-48} \qquad (5.16)$$

5.4 Modular invariance

5.4.1 Modular group of $\mathbf{T}^2$

To obtain the complete partition function we should sum over all inequivalent geometries. As we have discussed, it is crucial not to overcount geometries. Since we have characterized the worldsheet geometry in terms of τ, it is crucial to realize that there exist different values of τ which nevertheless correspond to the *same* geometry.

i) For instance, as shown in figure 5.4, two two-tori corresponding to τ and $\tau+1$ are defined by the same lattice on the 2-plane, hence correspond to the same two-torus geometry.

ii) A slightly trickier equivalence is that of two two-tori with complex structure parameters τ and $-1/\tau$. Let us verify this in the simpler case of $\tau_1 = 0$; in this case we have the equivalence of τ_2 and $1/\tau_2$. This is shown in figure 5.5: the two-torus with parameter i/τ_2 is equivalent to that with

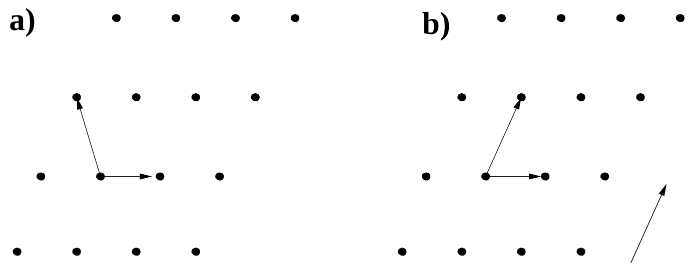


Figure 5.4: The two-tori corresponding to τ and $\tau + 1$ correspond to the same two-dimensional lattice of translation, hence are the same two-torus.

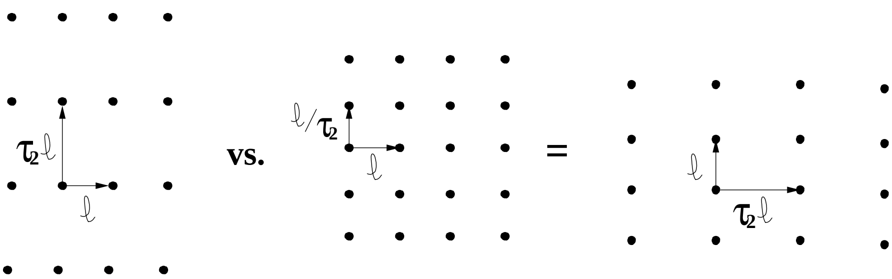


Figure 5.5: The geometry of two two-tori with parameters $i\tau_2$ and i/τ_2 is the same, as can be seen by exchanging the roles of σ and t and performing a rescaling of coordinates.

parameter $i\tau_2$, up to an exchange of the roles of σ and t, and a rescaling to ensure that the total length of the new σ coordinate is ℓ.

Two two-tori with parameters τ and $-1/\tau$ are simply related by the exchange of the roles of the two basis vectors generating the two-dimensional lattice.

In other words, there exist different choices of τ which lead to the same geometry, namely two two-tori which can be related by coordinate changes on the worldsheet

Denoting $z = \sigma + i\,t$, the two torus geometrical parameter τ is specified by the periodic identifications

a) $\sigma \to \sigma + \ell$, $t \to t$ which gives $z \to z + \ell$

b) $\sigma \to \sigma + \tau_1\ell$, $t \to t + \tau_2\ell$ which gives $z \to z + \tau\ell$

Performing a change of variables

$$\sigma' = \sigma + t/\tau_2 \quad ; \quad t' = t \tag{5.17}$$

The two-torus is defined in terms of the identifications

a) $\sigma \to \sigma + \ell$, $t \to t$, which gives $\sigma' \to \sigma' + \ell$, $t' \to t'$, namely $z' \to z' + \ell$

b) $\sigma \to \sigma + \tau_1\ell$, $t \to t + \tau_2\ell$, which gives $\sigma' \to \sigma' + (\tau_1 + 1)\ell$, $t' \to t' + \tau_2\ell$, namely $z' \to z' + (\tau + 1)$

So in these coordinates the two-torus has parameter $\tau + 1$.

Performing instead a change of variables

$$\sigma' = \frac{\tau_2 t + \tau_1\sigma}{\tau_1^2 + \tau_2^2} \quad ; \quad t' = \frac{\tau_1 t - \tau_2\sigma}{\tau_1^2 + \tau_2^2} \tag{5.18}$$

the two-torus is defined in terms of the identifications

a) $\sigma \to \sigma - \ell$, $t \to t$, which gives $\sigma' \to \sigma' + \tau_1'\ell$, $t' \to t' + \tau_2'\ell$, namely $z' \to z' + \tau'\ell$ with $\tau' = -1/\tau$

b) $\sigma \to \sigma + \tau_1\ell$, $t \to t + \tau_2\ell$, which gives $\sigma' \to \sigma' + \ell$, $t' \to t'$, namely $z' \to z' + \ell$.

So in these coordinates the two-torus has parameter $-1/\tau$.

This shows that the geometries corresponding to values of τ related by the transformations $\tau \to \tau + 1$, $\tau \to -1/\tau$ are equivalent up to coordinate changes, diffeomorphisms. It is important to notice that the diffeormorphisms involved are 'large', that is they are not continuously connected to the identity (they involve drastic things like exchanging the roles of σ, t; however, they are simply coordinate changes).

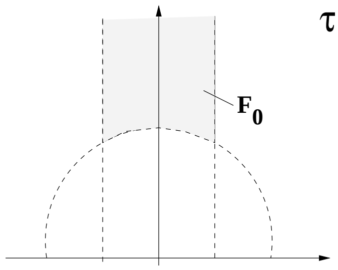


Figure 5.6: Fundamental domain of τ. Any point in the upper half plane can be mapped to some point in F_0 using the basic modular transformations $\tau \to \tau + 1$, $\tau \to -1/\tau$.

The set of transformations of τ which leaves the geometry invariant has the structure of a group, called the modular group of the two-torus. By composing the transformations $\tau \to \tau + 1$ and $\tau \to -1/\tau$, the most general tranformation is of the form

$$\tau \to \frac{a\tau + b}{c\tau + d} \quad \text{with } a, b, c, d \in \mathbf{Z} \quad \text{and } \mathbf{ad} - \mathbf{bc} = \mathbf{1} \tag{5.19}$$

The parameters a, b, c, d can be written as a 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ of integer entries and unit determinant. The group is therefore $SL(2, \mathbf{Z})$.

The set of inequivalent geometries is therefore characterized by the parameter τ in the upper half complex plane (recall we had $\tau_2 > 0$, modulo $SL(2, \mathbf{Z})$ transformations. A choice of fundamental domain of τ is shown in figure 5.6

$$-1/2 \leq \tau_1 < 1/2 \quad , \quad |\tau| \leq 1 \tag{5.20}$$

The set of points in F_0 correspond to the set of all possible two-torus geometries. Integrating τ over F_0 corresponds to summing over two-torus geometries with no overcounting.

5.4.2 Modular invariance of the partition function

The closed bosonic string partition function $Z(\tau)$ should be the same for equivalent tori, since it should be invariant under reparametrizations of the worldsheet. So $Z(\tau)$ should be modular invariant, i.e. $SL(2,\mathbf{Z})$ invariant. This is not completely obviour, since the diffeomorphisms involved in reparametrizations changing τ by modular transformations are not small, so in principle our gauge fixing procedure (good for 'small' diffeomorphism, continuously connected to the identity) is not good enough to take care of them [1].

Happily, using the modular tranformation properties of Dedekind's eta function (A.3), we find that

$$Z(\tau) \simeq \tau_2^{-12}\,|\eta(\tau)|^{-48} \stackrel{\tau\to\tau+1}{\longrightarrow} \tau_2^{-12}\,|\eta(\tau)|^{-48} \qquad (5.21)$$

$$Z(\tau) \simeq \tau_2^{-12}\,|\eta(\tau)|^{-48} \stackrel{\tau\to -1/\tau}{\longrightarrow} \frac{(\tau_1^2+\tau_2^2)^{12}}{\tau_2^{12}}\,\frac{1}{|\tau|^{24}\,|\eta(\tau)|^{48}} = \tau_2^{-12}\,|\eta(\tau)|^{-48}$$

It is modular invariant! From the viewpoint of the way we computed $Z(\tau)$, invariance under e.g. $\tau\to -1/\tau$ is remarkable: The sum over all states of a string along σ propagating in t is the same as the sum over all states of the string in the dual channel, a string along t and propagating in σ. Strinking conspiracy of the sum over the string tower... From another viewpoint, it is just a simple consequence of the geometry of the worldsheet. The amplitude is a function of the worldsheet geometry, and gives the same number for different values of τ that correspond to the same intrinsic geometry.

The complete vacuum amplitude is obtained by summing over inequivalent geometries, that is restricting to integrating τ over F_0

$$Z = \int_{F_0} \frac{d^2\tau}{4\tau_2}\,(4\pi^2\alpha'\tau_2)^{-12}\,|\eta(\tau)|^{-48} \qquad (5.22)$$

where $d^2\tau/(4\tau_2)$ is an $SL(2,\mathbf{Z})$ invariant measure in the space of two-tori geometries (the so-called Teichmuller space). It is easy to check this invariance by hand.

5.4.3 UV behaviour of the string amplitude

It is now time to study the UV behaviour of this amplitude. To understand better the nice UV properties of string theory, it is useful to obtain the

[1] We may say that, since even within our gauge fixing we still encounter the same geometry for different values of τ, our gauge fixing slices are passing through each gauge orbit more than once. If the value of Z is the same in each such point, we may by hand just keep one of them. If not, then the theory is not invariant under large diffeomorphisms, it does not have a consistent worldsheet geometry.

vacuum to vacuum amplitude in a theory of point particles. In a theory of one point particle of mass m in D dimensions, the amplitude of a diagram given by a circular worldline of length l is

$$Z_m = V_d \int \frac{d^D k}{(2\pi)^D} \int_0^\infty \frac{dl}{2l}\, e^{-(k^2+m^2)l/2} \tag{5.23}$$

with $(k^2+m^2)/2$ the worldline hamiltonian, and $dl/(2l)$ the measure in the space of circle geometries, with the denominator $2l$ removing the freedom of translation plus inversions of the circle. We have

$$Z_m = iV_d \int_0^\infty \frac{dl}{2l}\,(2\pi l)^{-D/2}\, e^{-m^2 l/2} \tag{5.24}$$

For any $D > 0$ this amplitude is divergent in the UV, as $l \to 0$. On the other hand, it is IR convergent if $m^2 > 0$.

One could imagine that string theory is just a theory with an infinite number of particles in spacetime. That is not really true, in a very subtle way which we will see below. If that were true, then the vacuum to vacuum amplitude in string theory would be just the sum of contributions like (5.24) for all particles in the string tower. Using that the mass of a string state is given by $m^2 = 2/\alpha'(L_0 + \tilde{L}_0)$ we have

$$Z' = iV_d \int_0^\infty \frac{dl}{2l}\,(2\pi l)^{-D/2}\, \mathrm{tr}_{\mathcal{H}}\, e^{-l/\alpha'(L_0+\tilde{L}_0)} \tag{5.25}$$

We prefer to sum over the extended Hilbert space of the theory by not requiring directly $L_0 = \tilde{L}_0$, and rather imposing this constraint by hand via a delta function

$$\delta_{L_0,\tilde{L}_0} = \int_{-\pi/2}^{\pi/2} \frac{d\theta}{2\pi}\, e^{i(L_0-\tilde{L}_0)\theta} \tag{5.26}$$

to get

$$Z' = iV_d \int_0^\infty \frac{dl}{2l} \int_{-\pi/2}^{\pi/2} \frac{d\theta}{2\pi}\,(2\pi l)^{-D/2}\, \mathrm{tr}_{\mathcal{H}}\, e^{-l/\alpha'(L_0+\tilde{L}_0)}\, e^{i(L_0-\tilde{L}_0)\theta} \tag{5.27}$$

and introducing $\tau = \frac{\theta}{2\pi} + i\frac{l}{\alpha'}$

$$Z' = iV_d \int_R \frac{d^2\tau}{4\tau_2}\,(4\pi^2\alpha'\tau_2)^{-D/2}\, \mathrm{tr}_{\mathcal{H}}\, q^{L_0}\, \overline{q}^{\tilde{L}_0} \tag{5.28}$$

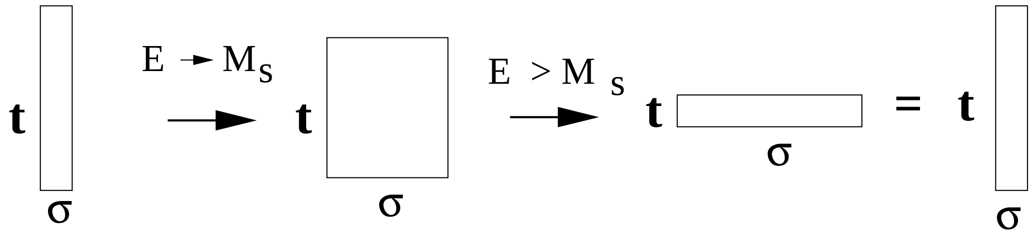


Figure 5.7: As the energy in the internal loop increases, longer strings run through it. The UV limit is geometrically equivalent to some infrared contribution, which has been already counted.

with R the region $\tau_2 > 0$, $-1/2 \leq \tau_1 < 1/2$.

This is the same as the true string amplitude, except for the crucial difference of the intergration region, $R \neq F_0$. Indeed if (5.28) were the true string amplitude we would obtain the same UV divergences at $\tau_2 \to 0$ as for a theory of point particles. On the other hand, in the true string amplitude (5.22), the UV divergent region $\tau_2 \to 0$ is simply *absent*!

To understand a bit better where the UV region has gone, let us consider summing over two-torus worldsheets as the energy of the intermediate states increases, see figure 5.7. As the energy increases, longer and longer strings are exchanged for a shorter and shorter time. For $E \gg M_s$ the diagram of very long strings propagating over a very short time has the same geometry as and IR contribution (by exchange of the roles of σ, t), so it has been already counted. Notice that very remarkably the sum of the UV behaviours of all the states in the string tower resums into an infrared behaviour, which is typically convergent [2]

Notice that to get this result it was crucial not to overcount the worldsheet geometries. Worldsheet geometry provides an extremely clever cutoff, which makes string theory quite different from just a field theory with an infinite number of fields.

Let us comment that this feature that any UV divergent region is absent in string theory is completely general, and valid for other diagrams, with more handles and with external insertions. For instance see figure 5.8. Just

[2] In the closed bosonic string theory, the IR is divergent due to the existence of a tachyonic state. The IR is well-behaved in other theories with no spacetime tachyons, like the superstrings.

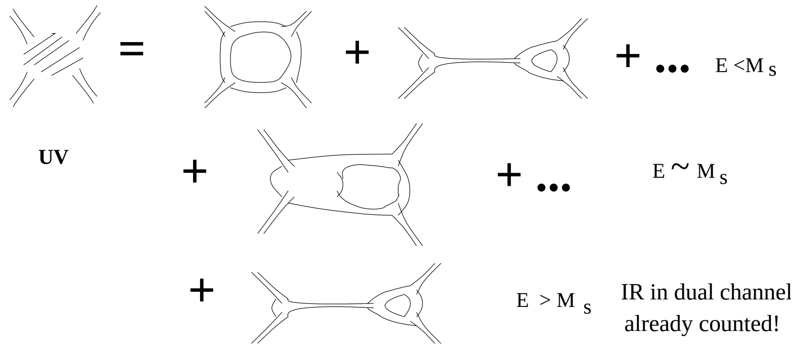


Figure 5.8: The contribution to a 1-loop four-string scattering amplitudes. The first line shows some low-energy contributions; the second line shows the first contributions for higher energy, with longer strings being exchanged in one internal leg. The third line shows the same diagram for energies much larger than M_S; this seemingly UV regime in geometrically the same as one of the IR contributions, so it has been already counted and should not be included again.

as above, the UV behaviour of the complete tower of string states resums into and IR contribution in a dual channel, which is a non-divergent contribution.

Let us conclude by pointing out that the low energy contribution to the partition function, the vacuum to vacuum amplitude is divergent in the bosonic string theory. This is because the IR contribution is dominated by the lightest mode, which is a tachyon with $m^2 = -4/\alpha'$. In the IR $\tau_2 \to i\infty$ the string partition function reduces to the point particle one with m given by the lightest state mass; one clearly gets an exponential $e^{+\tau_2}$ which diverges. In theories with no spacetime tachyon, the IR limits are however well-behaved, so the finiteness of string theory works as discussed above.

Concerning the IR divergence found above, one may wonder whether it is a physical infinity. It is easy to show that the vacuum to vacuum amplitude is related to the vacuum energy density, namely to the cosmological constant in spacetime. Since the spacetime theory is coupled to gravity, it is indeed a physical observable, and the infinity is physical. So the theory is to some extent sick.

There is a lot of speculation about the meaning of the tachyon in bosonic string theory. Our present idea is that it signals an instability of the vacuum of the theory, rather than an essential inconsistency of the theory; the problem is that we have no idea which is the correct vacuum, around which there would be not spacetime tachyons.

Chapter 6

Toroidal compactification of closed bosonic string theory

6.1 Motivation

As discussed in the overview lectures, a canonical mechanism to obtain four-dimensional physics at low energies out of a theory with $D > 4$ is to consider the theory in a curved background of the form $M_4 \times X_{D-4}$, with X_{D-4} a $(D-4)$-dimensional *compact* manifold, called the internal space. At energies $E \ll 1/L$, where L is the typical size of the dimensions in X_{D-4}, the physics is essentially 4d, we do not have enough resolution to see the internal space. This is called compactification of the theory.

One of the simplest possibilities is to consider the internal space to be a $(D-4)$-torus. In this section we are interested in exploring this possibility in string theory. Happily, the most interesting phenomena are already present in we compactify just one dimension on a circle, and reduce the 26d bosonic string theory to a 25d theory at low energies.

We start with a discussion of compactification in field theory. As we know,this provides a good approximation to the dynamics of string theory when α' corrections are negligible [1]. That is, when the internal space radius is much larger than the string length scale. Even in this regime there are interesting phenomena, like the Kaluza-Klein mechanism to generate gauge vector bosons out of the higher dimensional metric.

Next we turn to the explicit discussion of compactification in full-fledged

[1]Recall the picture 6.1.

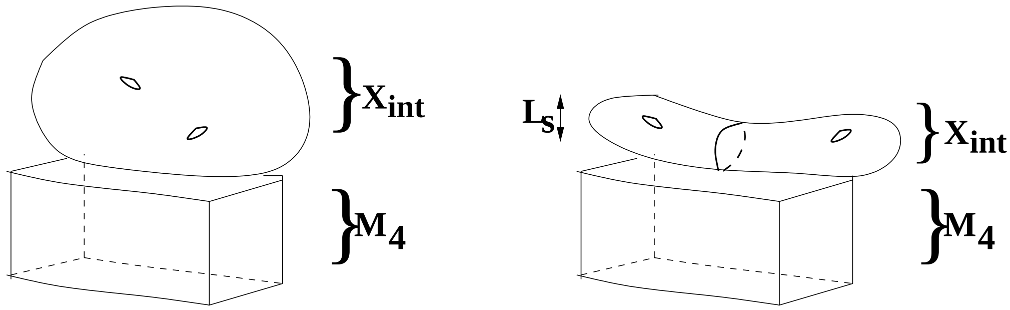


Figure 6.1: Picture of compactification spacetimes; thick small lines represent string states which are light in the corresponding configuration. When the internal manifold has size of the order of L_s, stringy effects (which do not exist in theories of point particles) become relevant; for instance, string winding modes (where a closed string winds around some internal dimension) may become light.

string theory. This can be carried out for toroidal compactification because it is described by a free worldsheet theory, which can be quantized exactly in the sense of the α' expansion. This means that for compactification on circles of radius comparable or smaller than the string length, string theory may (and does) differ from field theory.

Among the most surprising effects, we will find i) new light (and even massless) particles arising from closed string winding around the internal circle, and ii) T-duality, a complete physical equivalence of two theories living in different spacetimes.

Results in this section are useful in discussing toroidal compactifications in other string theories, like superstrings. Also, they will be useful in the construction of 10d heterotic string theories.

6.2 Toroidal compactification in field theory

Here we roughly follow ideas in section 8.1 of [55]. Our discussion is sketchy and provides most results without their detailed derivation.

Let us first consider circle compactification in field theory, which is a good approximation to the situation in string theory for circle radius much larger than the string length, so that α' effects (which are the ones related to the fact that the string is an extended object) are negligible.

So we consider field theories in D-dimensions, propagating on a background spacetime of the form $M_d \times \mathbf{S^1}$, with $D = d + 1$. To explain why the low-energy physics is d-dimensional, consider first a toy model of a D-dimensional massless scalar field $\varphi(x^0, \ldots, x^{D-1})$ propagates with D-dimensional action

$$S_{5d\varphi} = \int_{M_d \times \mathbf{S^1}} d^D x \, \Lambda^{D-4} \partial_M \varphi \partial^M \varphi \tag{6.1}$$

with $M = 0, \ldots, D-1$ and where Λ is some scale which we have introduced for dimensional reasons.

Since x^{D-1} parametrizes a circle, it is periodic, and we can expand the x^{D-1} dependence in Fourier modes

$$\varphi(x^0, \ldots, x^{D-1}) = \sum_{k \in \mathbf{Z}} e^{2\pi i k x^{D-1}/L} \, \varphi_k(x^0, \ldots, x^{d-1}) \tag{6.2}$$

where $L = 2\pi R$ is the length of $\S^1$.

From the d-dimensional viewpoint, we see a bunch of d-dimensional scalar fields $\varphi_k(x^0, \ldots, x^{d-1})$, labeled by the integer index k, which defines the momentum in the extra dimension $p_{D-1} = k/R$. The d-dimensional spacetime mass of those fields increases with k^2. To see that, take the D-dimensional mass-shell condition

$$P^2 = 0 \, , \quad \text{that is } P^2_{M_d} + p^2_{D-1} = 0 \tag{6.3}$$

For the field φ_k, we have

$$P^2_{M_d} + (k/R)^2 = 0 \tag{6.4}$$

which means that the d-dimensional mass of the field φ_k is

$$m_k^2 = (k/R)^2 \tag{6.5}$$

Equivalentely, we may obtain this result from the d-dimensional wave equation for the field φ_k

$$\partial_M \varphi \partial^M \varphi = 0 \quad \to \quad \partial_\mu \varphi_k \partial^\mu \varphi_k + (k/R)^2 = 0 \tag{6.6}$$

where $\mu = 0, \ldots, d-1$. And we recover (6.5).

At energies much lower than the compactification scale $M_c = 1/R$, $E \ll 1/R$, the only mode which is observable is the zero mode $\varphi_0(x^0, \ldots, x^{d-1})$. So we see just a single d-dimensional field, with a d-dimensional action, which is obtained by replacing $\varphi(x^0,\ldots,x^{D-1})$ in (6.1) by the only component we are able to excite $\varphi_0(x^0,\ldots,x^{d-1})$. The x^{D-1} dependence drops and we get

$$S_{eff} \,=\, \int_{M_d} d^d x\, \frac{L}{\Lambda^{D-4}}\, \partial_\mu \varphi_0 \partial^\mu \varphi_0 \tag{6.7}$$

So we recover d-dimensional physics at energies below M_c. This is the Kaluza-Klein mechanism, or Kaluza-Klein reduction. The massive d-dimensional fields φ_k are known as Kaluza-Klein (KK) excitations or KK replicas of φ_0.

Obs: If the higher-dimensional field theory contains massive fields with mass M, the 4d KK tower has masses $m_k^2 = M^2 + (k/R)^2$, so they will not be observable at energies below M.

The Kaluza-Klein reduction works for any higher dimensional field. An important new feature arises when the original higher dimensionl field has non-trivial Lorentz quantum numbers. The procedure is then to first decompose the representation of the $SO(D)$ higher-dimensional Lorentz group with respect to the lower-dimensional one $SO(d)$ (i.e. separate different components according to their behaviour under d-dimensional Lorentz), and finally perform KK reduction for each piece independently. For instance, for a D-dimensional graviton we have the KK reduction on $\mathbf{S^1}$

$$\begin{array}{lrl} G_{MN}(x^0,\ldots,x^{D-1}) \to & G_{\mu\nu}(x^0,\ldots,x^{D-1}) \to & G_{\mu\nu}^{(0)}(x^0,\ldots,x^{d-1}) \\ & G_{\mu,D-1}(x^0,\ldots,x^{D-1}) \to & G_{\mu 4}^{(0)}(x^0,\ldots,x^{d-1}) \\ & G_{D-1,D-1}(x^0,\ldots,x^{D-1}) \to & G_{44}^{(0)}(x^0,\ldots,x^{d-1}) \end{array} \tag{6.8}$$

where the first step is just decomposition in components, and the second is KK reduction. We therefore obtain, at the massless level, a d-dimensional graviton, a d-dimensional $U(1)$ gauge boson, and a d-dimensional scalar.

To be more specific, the only piece of the D-dimensional metric which is visible from the low-energy d-dimensional viewpoint is

$$ds^2 \,=\, G_{\mu\nu}\, dx^\mu\, dx^\nu \,+\, G_{dd}\,(dx^d + A_\mu dx^\mu)^2 \tag{6.9}$$

where the fields $G_{\mu\nu}$, G_{dd}, A_μ, are already taken to be the zero modes of the KK tower, and so depend only on the non-compact coordinates $x^0,\ldots,x^{d-1}$.

The original D-dimensional invariance under diffeomorphism has a remnant in this truncation of the theory. In particular, it is clear that we have d-dimensional diffeomorphism invariance acting on $x^0,\ldots,x^{d-1}$ (for which

$G_{\mu\nu}$ is the graviton). There is an additional freedom to reparametrize the internal coordinate as

$$x'^d = x^d + \lambda(x^\mu) \tag{6.10}$$

The effect of this tranformation is to change the d-dimensional vector boson

$$A'_\mu = A_\mu - \partial_\mu \lambda \tag{6.11}$$

So gauge transformations of this vector boson follow from coordinate reparametrization in the internal dimension. This remarkable result (gauge invariance from diffeomorphism invariance in higher dimensions) was the original motivation for the Kaluza-Klein program of unification of interactions, which has motivated much of the modern research in extra dimensions.

Another field whose KK reduction we will be interested in is a D-dimensional 2-form B_{MN}. By an argument similar to the above one for the graviton, the result is a d-dimensional theory with a d-dimensional 2-form $B_{\mu\nu}$ and a $U(1)$ gauge boson $\hat{A}_\mu$. Just as above, gauge invariance of the D-dimensional 2-form inplies invariance of the d-dimensional 2-form under

$$B_{\mu\nu} \rightarrow B_{\mu\nu}\, \partial_{[\mu}\, \Lambda_{\nu]}(x^\lambda) \tag{6.12}$$

We will be interested in performing the KK reduction of the effective field theory for the light modes of the closed bosonic string. This includes a 26d graviton G_{MN}, a 26d scalar dilaton ϕ, and a 26d 2-form field B_{MN}.

As discussed in the overview lectures, the original action is

$$S_{\rm eff.} = \frac{1}{2k_0^2} \int d^{26}X\, (-G)^{1/2}\, e^{-2\phi}\, \{\, R - \frac{1}{12}\, H_{MNP}\, H^{MNP} + 4\partial_M\phi\partial^M\phi \,\} + \mathcal{O}(\alpha') \tag{6.13}$$

where $H_{MNP} = \partial_{[M} B_{NP]}$.

Substitution of the 26d fields by the 25d zero modes of the KK tower, leads to the 25d effective action for the latter. Defining $G_{25,25} = e^{2\sigma}$, it is given by [2]

$$\begin{aligned} S_{25d} &= \tfrac{2\pi R}{2k_0^2} \int d^{25}X \quad (-G)^{1/2}\, e^{-2\phi+\sigma}\, [\, R - 4\,\partial_\mu\phi\partial^\mu\sigma + 4\partial_\mu\phi\partial^\mu\phi + \\ &\qquad - \frac{1}{4}\, e^{2\sigma}\, F_{\mu\nu}F^{\mu\nu} - \frac{1}{12}\, H_{\mu\nu\lambda}H^{\mu\nu\lambda} - \frac{1}{4}\, e^{2\sigma}\, \hat{F}_{\mu\nu}\hat{F}^{\mu\nu} \Big] = \\ &= \tfrac{2\pi R}{2k_0^2} \int d^{25}X \quad (-G)^{1/2}\, e^{-2\phi_{25d}}\, [\, R - 4\,\partial_\mu\sigma\partial^\mu\sigma + 4\partial_\mu\phi\partial^\mu\phi + \\ &\qquad - \frac{1}{4}\, e^{2\sigma}\, F_{\mu\nu}F^{\mu\nu} - \frac{1}{12}\, H_{\mu\nu\lambda}H^{\mu\nu\lambda} - \frac{1}{4}\, e^{2\sigma}\, \hat{F}_{\mu\nu}\hat{F}^{\mu\nu} \Big] \end{aligned} \tag{6.14}$$

[2] This combines eqs (8.1.9) and (8.1.13) in [55].

where $H_{\mu\nu\lambda} = \partial_{[\mu} B_{\nu\lambda]} - A_{[\mu} \hat{F}_{\nu\lambda]}$, and where we have defined $\phi_{25d} = \phi - \sigma/2$, the effective 25d dilaton, which fixes the 25d interaction strength.

Notice that the vev for the scalar field $G_{25,25}$ is related to the radius of the internal circle. In fact, only the combination $\rho = Re^{\sigma}$ labels inequivalent theories. Therefore, the radius is not an external parameter, but the vev of a 4d dynamical scalar field. On the other hand, the compactification background is consistent (solves the D-dimensional equations of motion) no matter what circle radius we choose; this implies that in the d-dimensional effective action there is no potential for this scalar, it parametrizes what is called a flat direction of the potential. The field is called a modulus, and its vev parametrizes inequivalent vacua of the theory. The set of vevs for this modulus is called the moduli space (of circle compactifications).

A last important comment. It is interesting to notice that states carrying momentum in the circle direction are charged with respect to A_μ. This is because the global version of the corresponding gauge symmetry is a translation along x^d, hence the corresponding charge is internal momentum. This is a lower-dimensional remnant of the fact that the higher dimensional gravition couples to the energy momentum tensor. On the other hand, the original field theory did not have states charged under the 2-form field, hence the lower-dimensional theory does not have any states charged under the gauge boson $\hat{A}_\mu$. Later on we will see that string theory does contain such charged states.

6.3 Toroidal compactification in string theory

Let us discuss the circle compactification of the closed bosonic string in string theory language. Naively, to do that, we need to specify the worlsheet action for a string propagating [3] in $M_{25} \times \mathbf{S^1}$, by replacing the Minkowski metric in M_{26} in the Polyakov action by the metric in $M_{25} \times \mathbf{S^1}$. The puzzling feature is that the latter metric is also flat, locally a Minkowski metric as well, so the worldsheet action is still

$$S_{\rm P} \;=\; -\frac{1}{4\pi\alpha'} \int_\Sigma d^2\xi \, (-g)^{1/2} \, g^{ab}(\sigma, t) \, \partial_a X^\mu \, \partial_b X^\nu \eta_{\mu\nu} \tag{6.15}$$

[3] It is possible to work in general and finally show that consistency requires the total dimension of spacetime to be $D = 26$ so we settle this from the start.

The difference between $M_{25} \times \mathbf{S^1}$ and M_{26} is a global effect, they have different topology although the local metric is the same for both. The effects of the compactification will arise not at the level of the local structure of the worldsheet, but in the boundary conditions we have to impose on the 2d worldsheet fields.

6.3.1 Quantization and spectrum

Indeed, the light-cone quantization can be carried out without change as in the uncompactified theory until we reach the hamiltonian

$$H \,=\, \frac{\ell}{4\pi\alpha' p^+} \int_0^\ell d\sigma \, [\, 2\pi\alpha' \, \Pi_i \, \Pi_i \,+\, \frac{1}{2\pi\alpha'} \, \partial_\sigma X^i \, \partial_\sigma X^i \,] \tag{6.16}$$

In order to rewrite it in terms of oscillator modes, etc, we need to specify the boundary conditions obeyed by the 2d physical fields $X^i(\sigma, t)$. For X^i, $i = 1, \ldots, 24$, we need to impose

$$X^i(\sigma + \ell, t) \,=\, X^i(\sigma, t) \quad \text{for} \quad i = 1, \ldots, 24 \tag{6.17}$$

as usual. However, the fact that X^{25} parametrizes a circle of radius R means that X^{25} and $X^{25} + 2\pi R$ correspond to the same point in spacetime. Hence, the following boundary condition defines a consistent closed string

$$X^{25}(\sigma + \ell, t) \,=\, X^{25}(\sigma, t) \,+\, 2\pi R\, w, \quad , \quad w \in \mathbf{Z} \tag{6.18}$$

It corresponds to a closed string winding around the internal circle a number of times given by w, which is called the winding number, see fig 6.2. [4] . Each value of w corresponds to a different closed string sector. The complete spacetime 25d spectrum is given by the set of states of closed string in all possible winding sectors.

---

[4] It is amusing to notice that, from the viewpoint of the 2d theory, configurations of fields $X^i(\sigma, t)$ satisfying boundary conditions with non-zero winding correspond to solitonic states of the 2d field theory. The topological quantity associated to these solitons is the spatial integral of the derivative of the 2d field, namely $\int_0^\ell \partial_\sigma X^{25} = 2\pi R w$. As usual, solitons of a field theory are associated to non-trivial topology of the target space where the fields take values (recall that in the 't Hooft-Polyakov monopole, the existence of a soliton in the 4d theory was associated to the non-trivial topology of the space of vacua, namely the space where the Higgs field takes values). Please recall that here we are talking about solitons on the worldsheet, and have no relation at all with spacetime solitons.

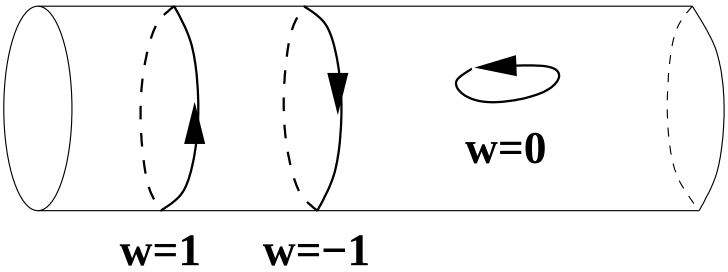


Figure 6.2: States representing closed strings winding around the compact dimension.

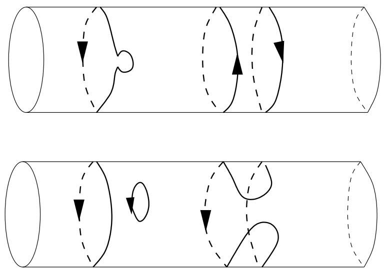


Figure 6.3: String interactions conserve winding number.

The existence of winding is possible only because strings are extended objects. The sector $w = 0$ corresponds to taking strings which are already closed without the compactification. These are the fields that appear in the approximation of compactifying the effective 26d field theory. We will see that for large radius states in non-zero winding sectors are very heavy, and this is a good approximation. For small radius, non-zero winding state lead to very interesting surprises!

Winding number is conserved in string interactions, see figure 6.3

Since the X^i, $i = 2, \ldots, 24$ behave as usual, we only center on the analysis of X^{25}. The mode expansion for the boundary conditions (6.2) are

$$X^{25}(\sigma, t) = x^{25} + \frac{p_{25}}{p^+} t + \frac{2\pi R w}{\ell} \sigma + i\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbf{Z}-\{\mathbf{0}\}} \left[\frac{\alpha_n^i}{n} e^{-2\pi i n(\sigma+t)/\ell} + \frac{\tilde{\alpha}_n^i}{n} e^{2\pi i n(\sigma-t)/\ell} \right] (6.1$$

Notice that the momentum must be quantized $p_{25} = k/R$, with $k \in \mathbf{Z}$ just like in the field theory discussion.

For future convenience, we may recast the expansion in terms of left and right movers $X^{25}(\sigma,t) = X^{25}_L(\sigma+t) + X^{25}_R(\sigma-t)$

$$\begin{aligned} X^{25}_L(\sigma+t) &= \tfrac{x^{25}}{2} + \tfrac{p_L}{2p^+}(t+\sigma) + i\sqrt{\tfrac{\alpha'}{2}}\sum_{n\in\mathbf{Z}-\{\mathbf{0}\}} \tfrac{\alpha^i_n}{n}\, e^{-2\pi i\, n\,(\sigma+t)/\ell} \\ X^{25}_R(\sigma-t) &= \tfrac{x^{25}}{2} + \tfrac{p_R}{2p^+}(t-\sigma) + i\sqrt{\tfrac{\alpha'}{2}}\sum_{n\in\mathbf{Z}-\{\mathbf{0}\}} \tfrac{\tilde{\alpha}^i_n}{n}\, e^{2\pi i\, n\,(\sigma-t)/\ell} \end{aligned} \quad (6.20)$$

with

$$p_L = \frac{k}{R} + \frac{wR}{\alpha'} \quad ; \quad p_R = \frac{k}{R} - \frac{wR}{\alpha'} \quad (6.21)$$

These will be called left and right moving momenta (although notice that each is a combination of the real spacetime momentum and winding).

The hamiltonian differs from the one in the non-compact situation only in the new contributions of winding terms to $\partial_\sigma X^{25}$. In terms of modes, etc, we obtain

$$\begin{aligned} H &= H_{w=0} + \frac{\ell}{4\pi\alpha' p^+}\int_0^\ell d\sigma\, \frac{1}{2\pi\alpha'}\,(\frac{2\pi R\, w}{\ell})^2 = \\ &= \sum_{i=2}^{24}\frac{p_i^2}{2p^+} + \frac{(k/R)^2}{2p^+} + \frac{R^2w^2}{2\alpha'^2p^+} + \frac{1}{\alpha' p^+}\,(N+\tilde{N}-2) \end{aligned} \quad (6.22)$$

where $H_{w=0}$ is the usual hamiltonian in the non-compact case. As usual, we build the Hilbert space of the theory by taking oscillator groundstates (each one labeled by a 25d momentum, a quantized momentum $k\in\mathbf{Z}$ in the circle, and a winding number) and applying oscillator creation operators to it.

The level matching constraint is $P=0$ with

$$\begin{aligned} P &= \int_0^\ell d\sigma\, \Pi_i\, \partial_\sigma X^i = \frac{p^+}{\ell}\int_0^\ell d\sigma\, \partial_t X^i\, \partial_\sigma X^i = \\ &= P_{w=0} + \frac{p^+}{\ell}\,\ell\,\frac{k/R}{p^+}\,\frac{2\pi R w}{\ell} = \frac{2\pi}{\ell}\,(N-\tilde{N}+kw) \end{aligned} \quad (6.23)$$

Each state corresponds to a particle in 25d spacetime. The 25d mass of the corresponding state is given by

$$M^2_{25d} = 2p^+H - \sum_{i=2}^{24} p_i^2 \quad (6.24)$$

We obtain

$$M_{25d}^2 = \frac{k^2}{R^2} + \frac{R^2}{\alpha'^2}w^2 + \frac{2}{\alpha'}(N + \tilde{N} - 2) \tag{6.25}$$

As mentioned above, for large R^2/α', the states with non-zero winding have large $\alpha' M^2$ and decouple. For not so large R^2/α', effects of winding states are very relevant and we cannot trust results obtained from the field theory approximation (namely, the physics obtained only from the $w = 0$ sector). Winding states, equivalentely α' effects, lead to important modifications of the physics, which can be regarded as important modification to how string theory feels the geometry when curvature lengths are as small as the string length (this is called stringy geometry for instance in the book by B. Greene).

For future convenience, we split the hamiltonian and mass in left and right handed pieces. We have $H = H_L + H_R$ with

$$\begin{aligned} H_L &= \frac{1}{4p^+}\left[\sum_{i=1}^{24} p_i^2 + p_L^2\right] + \frac{1}{\alpha' p^+}(N + E_0) \\ H_R &= \frac{1}{4p^+}\left[\sum_{i=1}^{24} p_i^2 + p_R^2\right] + \frac{1}{\alpha' p^+}(\tilde{N} + \tilde{E}_0) \end{aligned} \tag{6.26}$$

and $M^2 = M_L^2 + M_R^2$ with

$$\begin{aligned} M_L^2 &= \frac{p_L^2}{2} + \frac{2}{\alpha'}(N - 1) \\ M_R^2 &= \frac{p_R^2}{2} + \frac{2}{\alpha'}(\tilde{N} - 1) \end{aligned} \tag{6.27}$$

We see that one may carry out the quantization of the left and right moving coordinates independently, reach a mass formular for each side, and finally combine things together (satisfying the level matching constraint) at the end. This is only to re-emphasize the fact that in 2d the field theory of purely left-moving and purely right-moving fields make sense independently [5]. At a last stage, states of both theories are combined together to give physical states.

The level-matching constraint is

$$M_L^2 = M_R^2 \tag{6.28}$$

[5]This observation will be crucial in the construction of heterotic string theorires.

It is an easy exercise to obtain the one-loop partition function for this theory. For a two-torus worldsheet with geometry specified by τ_1, τ_2, we have

$$Z(\tau) = \mathrm{tr}_{\mathcal{H}_{\rm closed}} [\, e^{-\tau_2 \ell H}\, e^{i\tau_1 \ell P}\,] = $$
$$= \sum_{k,w=-\infty}^{\infty} \mathrm{tr}_{\mathcal{H}_{k,w}} [\, e^{-\tau_2\pi\alpha' \sum_{i=1}^{24} p_i^2}\, e^{-\tau_2\pi\alpha'(k/R)^2}\, e^{-\tau_2\pi R^2 w^2/\alpha'}\, e^{-2\pi\tau_2(N+\tilde{N}-2)}\, e^{2\pi i\tau_1(N-\tilde{N})}\, e^{2\pi i\tau_1\, kw}\,]$$

Here $\mathcal{H}_{k,w}$ is the closed string sector with momentum k and winding number w. Most of this computation is already familiar, the only new piece is the contribution over discrete momenta and the windings. We get

$$Z(\tau) \;=\; |\eta(\tau)|^{-48}\,(2\pi\alpha'\tau_2)^{-23/2} \sum_{k,w=-\infty}^{\infty} \exp\left[\, -\pi\tau_2 \left(\frac{\alpha' k^2}{R^2} + \frac{R^2 w^2}{\alpha'} + 2\pi i\tau_1\, kw \right) \right] \qquad (6.29)$$

This expression is modular invariant. Invariance under $\tau \to \tau+1$ is obvious, whereas invariance under $\tau \to -1/\tau$ can be shown by using Poisson resummation formula

$$\sum_{n\in\mathbf{Z}} \exp[\, -\pi A(n+\theta)^2 + 2\pi i\,(n+\theta)\phi\,] \;=\; A^{-1/2} \sum_{k\in\mathbf{Z}} \exp[\, -\pi A^{-1}(k+\phi)^2 - 2\pi i k\phi\,] \qquad (6.30)$$

on both sums over k and w. It is interesting to point out that the sum over winding and momenta is almost invariant under $\tau \to -1/\tau$, except for picking up a factor of $(\tau\overline{\tau})^{1/2}$ which compensates for the lack of invariance of $|\eta(\tau)|^{-48}(\tau_2)^{-23/2}$.

It is important to point out that in string theory compactified on a circle, winding states are crucial in obtaining a modular invariant partition partition function. One intuitive way to argue about this is as follows. Consider starting with the partition function of the uncompactified theory

$$Z_{\rm uncomp.} \;=\; \mathrm{tr}_{\mathcal{H}_{\rm uncomp.}} [\, e^{-\tau_2\ell H}\, e^{i\tau_1\ell P}\,] \qquad (6.31)$$

In order to describe the theory compactified on a circle, we may do by explicitly forcing that the only states that propagate are those invariant under translations of $2\pi R$ in X^{25}, by inserting the projector

$$\Pi \;=\; \sum_{w\in\mathbf{Z}} e^{iw2\pi R\Pi_{25}} \qquad (6.32)$$

in the trace. Here Π_{25} is the momentum operator, and $T_w = e^{i2\pi wR\Pi_{25}}$ translates X^{25} by $2\pi Rw$. The partition function is

$$Z_{\rm comp.} \;=\; \sum_{w\in\mathbf{Z}} \mathrm{tr}_{\mathcal{H}_{\rm uncomp.}} [\, e^{-\tau_2\ell H}\, e^{i\tau_1\ell P}\, T_w\,] \qquad (6.33)$$

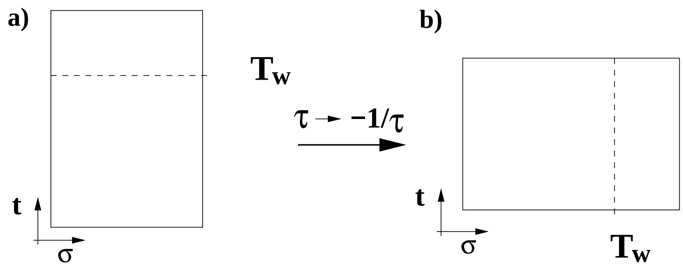


Figure 6.4: Under the modular transformation $\tau \to -1/\tau$, the roles of σ and t are exchanged. An insertion of T_w in the t (appearing from the insertion of the projector onto states invariant under discrete X^{25} translations) is mapped to an insertion of T_w in the σ direction, implying that we obtaine string closed up to translation in X^{25}, namely strings with winding w. Recall that sides of the rectangle are identified to make the worldsheet a two-torus.

This can be shown pictorially as in figure 6.4a. As the closed string propagates along the t direction, it crosses a cut along which the field $X^{25}(\sigma, t)$ jumps an amount $2\pi Rw$.

Under the modular transformation $\tau \to -1/\tau$, the roles of σ and t are exchanged, so the cut is found in the σ direction, as in figure 6.4b. Such picture represents a 1-loop diagram for a closed string which is closed up to a translation of the coordinate X^{25} by $2\pi Rw$, namely a closed string satisfying the boundary conditions (6.18). This means that to achieve a modular invariant partition function it is absolutely essential to add sectors with non-zero winding; namely, we have additional pieces

$$\sum_{w \in \mathbf{Z}} \mathrm{tr}_{\mathcal{H}_w} \left[e^{-\tau_2 \ell H_w} \, e^{i \tau_1 \ell P} \right] \tag{6.34}$$

where the trace is taken over the Hilbert space of string states in the sector of winding w.

Subsequently, we would have to enforce that in these new sectors the propagating modes are also invariant under translations of X_{25}, by introducing a projector. The total result is the double sum in k, w in (6.29). Sum in w sums over different sectors, whereas the sum in k projects onto states invariant under X^{25} translations.

6.3.2 α' effects I: Enhanced gauge symmetries

At large values of R, one easily recovers that the string spectrum reproduces the spectrum obtained using the field theory approximation. Indeed, winding states are very heavy, so only the $w = 0$ sector has a chance of being light. States with different k are merely KK replicas of the basic fields that exist in the 26d theory.

Forgetting the tachyon and its KK replicas (which can be lighter than $M^2 = 0$ for large enough R), the massless modes are $\alpha_{-1}^M \tilde{\alpha}_{-1}^N |0\rangle$, suitably decomposed according to whether $M, N = 25$, or $M, N = \mu$. Explicitly, we get

$$\alpha_{-1}^{\mu} \tilde{\alpha}_{-1}^{\nu} |0\rangle \tag{6.35}$$

which are the 25d graviton, 2-form, and a scalar (from the trace). We also have

$$\alpha_{-1}^{\mu} \tilde{\alpha}_{-1}^{25} |0\rangle \quad , \quad \alpha_{-1}^{25} \tilde{\alpha}_{-1}^{\mu} |0\rangle \tag{6.36}$$

two 25d gauge bosons. Taking symmetric and antisymmetric combinations, they are easily seen to arise from the 26d metric and 2-form, respectively. Hence the generic gauge symmetry in 25d is $U(1) \times U(1)$.

Finally we also have

$$\alpha_{-1}^{25} \tilde{\alpha}_{-1}^{25} |0\rangle \tag{6.37}$$

which is an additional scalar. This and the trace of (6.35) are the 25d dilaton and geometric moduli.

As in field theory, the charge of states under the gauge boson arising from the 26d graviton is given by their internal momentum, k. It is also easy to argue that the charge of states under the gauge boson arising from the 26d 2-form is given by their winding number w. Namely, starting from the coupling of a string to the 2-form field in 26d

$$\int_{\Sigma} B_M N \, \partial_a X^M \, \partial_b X^N \, \epsilon^{ab} \tag{6.38}$$

It is clear that we obtain a coupling of a string wrapped on $\mathbf{S^1}$ to the mixed component $B_{\mu,25}$,

$$\int dt \int_0^{\ell} d\sigma \, B_{\mu,25} \, \partial_\sigma X^{25} \, \partial_t X^{\mu} \simeq w \int dt \, \hat{A}_{\mu} \partial_t X^{\mu} \tag{6.39}$$

the state behaves as a 25d point particle coupling to $\hat{A}_\mu$ with charge w.

As announced before, as we let R approach the string length scale $L_s = \sqrt{\alpha'}$ new surprising features arise. In fact we can check that at $R = \sqrt{\alpha'}$ there appear new massless states from sectors of non-zero winding. The mass formulae in this point in moduli space are

$$\begin{aligned} \alpha' M_L^2 &= \frac{1}{2}(k+w)^2 + 2(N-1) \\ \alpha' M_R^2 &= \frac{1}{2}(k-w)^2 + 2(\tilde{N}-1) \end{aligned} \tag{6.40}$$

Denoting $|k,w\rangle$ the vacuum in the sector of momentum k and winding w, there are additional massless states, satifying the level matching condition (6.28).

We obtain four additional gauge bosons

$$\begin{aligned} &\alpha_{-1}^\mu|1,-1\rangle \quad , \quad \alpha_{-1}^\mu|-1,1\rangle \\ &\tilde{\alpha}_{-1}^\mu|1,1\rangle \quad , \quad \tilde{\alpha}_{-1}^\mu|-1,-1\rangle \end{aligned}$$

One should recall that they are charged under the generic $U(1)\times U(1)$ gauge symmetry, with charges given precisely by the pairs (k,w). The total gauge group is non-abelian and it is in fact $SU(2)^2$.

We also obtain eight new additional massless scalars

$$\begin{aligned} &\alpha_{-1}^{25}|1,-1\rangle \quad , \quad \alpha_{-1}^{25}|-1,1\rangle \\ &\tilde{\alpha}_{-1}^{25}|1,1\rangle \quad , \quad \tilde{\alpha}_{-1}^{25}|-1,-1\rangle \\ |2,0\rangle \quad , \quad &|-2,0\rangle \quad , \quad |0,2\rangle \quad , \quad |0,-2\rangle \end{aligned} \tag{6.41}$$

Checking the charges under the generic $U(1)^2$ symmetry, it is possible to see that these scalars, along with the radial modulus (6.37) transform in the representation $(\mathbf{3},\mathbf{3})$ of $SU(2)\times SU(2)$. The set of charges for the gauge bosons, and the scalars are shown in figure 6.5, and can be seen to correspon to roots of $SU(2)^2$ and weights of $(3,3)$.

This is a very surprising effect. For a particular value of the compactification radius $R=\sqrt{\alpha'}$, stringy effects (namely the existence of winding) generate an enhanced gauge symmetry in spacetime (enhanced as compared with the symmetry at a generic value of R). Indeed a dramatic effect! This mechanism of generating gauge bosons goes well beyond what was achievable from the field theory KK mechanism.

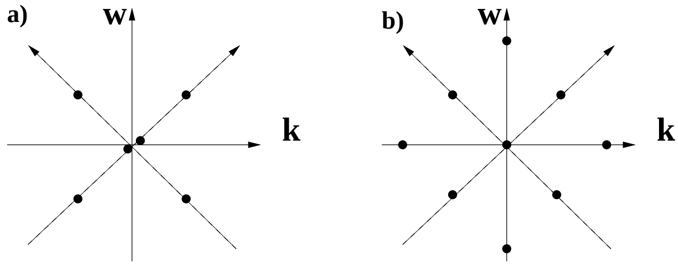


Figure 6.5: Charges of gauge bosons (a) and scalars (b) at the enhanced symmetry point $R = \sqrt{\alpha'}$. The charges tell us that the gauge bosons fill out a $SU(2) \times SU(2)$ group (the roots of each $SU(2)$ factor point along the dashed lines), whereas the scalars fill out a representation $(3,3)$ of $SU(2)^2$.

Of course it is possible to cook up a new 25d effective field theory by including by hand the new massless modes. So this effective field theory would contain gravity and non-abelian $SU(2)^2$ gauge interactions, and a bunch of 9 scalars transforming in the representation $(3,3)$ coupled to these gauge bosons. It is important to understand two facts:

• This effective field theory is *not* derived from the 26d effective field theory by compactification; we know that the latter missed the crucial issue of winding states, and is a good approximation at large R, and not at $R = \sqrt{\alpha'}$

• This effective field theory is a good approximation to the 25d physics for R close to $\sqrt{\alpha'}$. As we will see shortly, going away from $R = \sqrt{\alpha'}$ makes some fields massive, so for R very different from $\sqrt{\alpha'}$ these masses are too large and it is not a good idea to include the corresponding fields in the effective field theory.

It is interesting to understand what happens when we vary slightly the value of R away form the value $\sqrt{\alpha'}$. Since we have solved the string states for all values of R, we simply read off the mass formulae and see that the additional gauge bosons, as well as the additional scalars get masses (proportional to the deviation of R and $\sqrt{\alpha'}$.

This sounds very much like a Higgs mechanism, with gauge bosons becoming massive and some scalars being eaten and becoming the longitudinal components of the massive vector bosons. Indeed this is correct: for small departures from $R = \sqrt{\alpha'}$ the 25d effective field theory language should be appropriate and the breaking of the gauge group is just a Higgs mechanism triggered by the scalars in the $(3, 3)$.

A finer point is that the number of scalars that disappears is larger than the number of gauge bosons becoming massive. This is however consistent. Out of the original 9 massless scalars, 4 of them are eaten by the 4 gauge bosons associated to the broken generators, 1 of the remaining remains massless (and is interpreted as the geometric modulus parametrizing R), and the 4 remaining become massive due to couplings between them and the scalars picking up a vev.

As discussed by Polchinski (around eq (8.3.22), organizing the 9 scalars in a 3×3 matrix M_{ij}, the scalar potential for the theory at $R=\sqrt{\alpha'}$ includes an $SU(2)^2$ invariant term

$$V(M) = \epsilon^{ijk}\,\epsilon^{i'j'k'}\,M_{ii'}\,M_{jj'}\,M_{kk'} \tag{6.42}$$

Giving a vev to one of the scalars, say M_{33}, we generate mass terms

$$\epsilon^{ij}\epsilon^{i'j'}\,M_{ii'}\,M_{jj'} \tag{6.43}$$

for $i, i', j, j' = 1, 2$. Namely four fields become massive due to the scalar potential.

A tantalizing (but more advanced) comment is that the field that has received the vev has flat potential, so it is a modulus, and parametrizes the deviation of R from $\sqrt{\alpha'}$. So it is what we have called the geometric modulus. Increasing the vev for this field would eventually lead us into the large volume regime.

However notice that in principle any of the 9 fields in M_{ij} can be the one in getting the vev. They are in the same $SU(2)^2$ multiplet, so gauge invariance tells us that none of these fields is priviledged. Therefore, starting from the enhanced symmetry point, there seem to exist different regimes which can be interpreted as large volume regimes in suitable variables. This will become clearer after we study next section.

6.3.3 α' effects II: T-duality

The existence of winding states in string theory leads to another amazing surprise. Recall the mass formula (6.44)

$$M_{25d}^2 = \frac{k^2}{R^2} + \frac{R^2}{\alpha'^2}w^2 + \frac{2}{\alpha'}\,(N+\tilde{N}-2) \tag{6.44}$$

It is invariant under the so-called T-duality transformation

$$R \to \frac{\alpha'}{R} \quad ; \quad k \leftrightarrow w \tag{6.45}$$

Namely the complete spectrum of the theory at radius R is the same as the spectrum of the theory at radius α'/R, up to a relabeling of k and w.

This is extremely striking. If we are 25d observers and measure the spectrum of states, we would be unable to distinguish whether it is coming from a string theory compactified on a circle of radius R or α'/R.

Striking again! The theory at large $R \to \infty$ has infinite towers of momentum states becoming massless (the KK step $1/R$ is very small); this is a typical signal of a decompactification limit. On the other hand, in the T-dual theory the radius is going to zero $R' = \alpha'/R \to 0$, and we still recover infinite towers of states becoming massless, but now they are coming from string with winding number w (since the T-dual circle is small, it costs almost no energy to increase the winding number). So the small R limit looks also as a decompactification limit, and it *is* a decompactification limit in T-dual language!

One might thing that this puzzling feature is not a property of full-fledged string theory, but just an accidental property of the spectrum. This is not correct, and one can show that string interactions also respect T-duality. T-duality is the complete physical equivalence of the theories compactified on circles of radius R and α'/R.

In other words, both theories are described by exactly the same worldsheet theory, and differ on how the spacetime coordinates (the spacetime geometry) is recovered from the 2d worldsheet theory.

To be more specific, it is convenient to describe our worldsheet theory as given by two sets of 2d fields $X_L^i(\sigma+t)$ and $X_R^i(\sigma-t)$, which are decoupled. Now there are two ways to construct the true spacetime coordinates $X^i(\sigma,t)$ out of them. One possibility is

$$\begin{aligned} X^i(\sigma,t) &= X_L^i(\sigma+t) + X_R^i(\sigma-t) \quad ; \quad i=2,\ldots,24 \\ X^{25}(\sigma,t) &= X_L^{25}(\sigma+t) + X_R^{25}(\sigma-t) \end{aligned} \tag{6.46}$$

whereas there is another

$$\begin{aligned} X^i(\sigma,t) &= X_L^i(\sigma+t) + X_R^i(\sigma-t) \quad ; \quad i=2,\ldots,24 \\ X^{25}(\sigma,t) &= X_L^{25}(\sigma+t) - X_R^{25}(\sigma-t) \end{aligned} \tag{6.47}$$

The relation between one and the other is

$$p_L^{25} \to p_L^{25} \quad ; \quad p_R^{25} \to -p_R^{25} \quad ; \tag{6.48}$$

which corresponds to the T-duality transformation (B.1).

The implications of this are difficult to overemphasize. It certainly suggests that spacetime is a secondary concept in string theory, and that it is derived from more fundamental concepts like the worldsheet theory. What this means for our understanding of the nature of spacetime in string theory is still unclear.

A final comment we would like to make in this respect is that T-duality is in fact a $\mathbf{Z_2}$ remnant of a gauge symmetry. Indeed, there is a value of R for which the theory is self-dual, this is our old friend $R = \sqrt{\alpha'}$. At this point, the complete spectrum is invariant under $k \leftrightarrow w$.

It is also easy to see that the effect of this transformation is nothing but a gauge transformation within the enhanced gauge group $SU(2)^2$. Finally, it is possible to see that two T-dual deviations from $R = \sqrt{\alpha'}$ are mapped to each other by a relabeling transformation which is a subroup of this group: indeed, regarding $SU(2)$ as $SO(3)$ (the rotation group in 3d) a rotation of π around the axis distinguished by the field getting a vev (the direction 3 if $M_{3,i'}$ gets the vev) in the first $SO(3)$ has the effect of mapping the vev for one of the modulus to its negative. Hence maps a deformation toward $R > \sqrt{\alpha'}$ to a deformation towards $R < \sqrt{\alpha'}$.

This means that two T-dual theories are identified by a gauge transformation, so should not be considered as really different. Hence the moduli space of compactification is not really parametrized by the real line (i.e. possible values of R) but rather by the real line modulo $R \to 1/R$. The moduli space can therefore be described (with no redundancy) by the set of points $R > \sqrt{\alpha'}$.

Again this has amazing implications, since it suggests the existence of a minimum distance in string theory. These issues must be taken with a grain of salt, however, since in the study of D-branes the community has realized that there exist other objets in string theory which are able to probe distances much shorter than L_s [41].

We see that even the simplest compactification is rich enough to illustrate the amazing features of string theory regarding the nature of spacetime.

6.3.4 Additional comments

Let us conclude by pointing out some generalizations of the concepts we have studied in toroidal compactifications

- Toroidal compactification of more than one dimension

  This is studied nicely enough in section 8.4 in [55]. One can proceed in analogy with the circle case. Some of the new features of this situation are the appearance of scalars from the KK reduction of the 26d 2-form. They have flat potential and are new moduli from the viewpoint of the lower-dimensional theory, characterizing the background B-field in the internal space. The complete moduli space (without taking into account dualities) is called Narain moduli space and is described as a coset

$$\frac{O(k,k,\mathbf{R})}{O(k,\mathbf{R})\times\mathbf{O(k,R)}} \tag{6.49}$$

  The set of T-dualities is larger, and is given by the group $O(k,k,\mathbf{Z})$, so the true moduli space is

$$\frac{O(k,k,\mathbf{R})}{O(k,\mathbf{R})\times\mathbf{O(k,R)}\times\mathbf{O(k,k,Z)}} \tag{6.50}$$

  A standard reference on all these issues is [42].

- Buscher's T-duality

  The existence of T-dual configuration does not require spacetime to be a cartesian product with one factor given by a circle. In fact, T-duality can be extended to geometries with one Killing vector with compact orbits (with finite length, at least asymptotically). Buscher's formulae provide the backgroud obtained by applying T-duality along the orbits of this Killing vector. Surprisingly T-duality is even able to relate geometries with different topology.

- Compactification on non-toroidal geometries

  Although this can be considered in bosonic string theory, it has found more applications in the supertring context. We will discuss some of this for heterotic string theories in later lectures.

Chapter 7

Type II Superstrings

We are already familiar with bosonic string theory, and have learned how to solve the issue of reducing it to lower dimensions via compactification. However, we have been unable to construct a theory with fermions in spacetime.

In this and coming lectures we study string theories whose massless spectrum contains spacetime spinor particles. These are the superstring theories, and today we will center on a particular kind of them: type II superstrings (leaving other superstrings, like heterotic strings and type I strings, for later lectures).

Before getting started, let us mention that in order to identify the quantum numbers of states with respect to the spacetime Lorentz group, it is quite crucial to have in mind the representation theory of $SO(2n)$ Lie algebras, which can be found in section 6 of the appendix on group theory.

7.1 Superstrings

7.1.1 Fermions on the worldsheet

To describe a new string theory we have to modify the worldsheet theory. Clearly, if we keep the same field content as in the bosonic string and simply add interactions, the spectrum in spacetime will not be very different from that in the bosonic theory, and in particular it will not contain spacetime fermions. Addint interactions is more similar to just curving the background on which the string is propagating.

Instead, we propose to change the field content of the 2d theory describing the worldsheet. A simple possibility which preserves D-dimensional Poincare invariance is to make the 2d worlsheet theory supersymmetric [1]. Namely, to add 2d fermion fields $\psi^\mu(\sigma, t)$, partners of the 2d bosonic fields $X^\mu(\sigma, t)$, and gravitino partners for the worldsheet metric $g_{ab}(\sigma, t)$ (notice that since supersymmetry commutes with global symmetries, the 2d fermionic fields should transform in the vector representation of the D-dimensional spacetime Lorentz group, just like the 2d bosonic fields). It is important to emphasize that at this stage it is not obvious at all that such theory will lead to spacetime fermions or spacetime supersymmetry; in fact, the 2d fermion fields are bosons with respect to the spacetime Lorentz group!

Two-dimensional theories of this kind are sometimes refered to as 'fermionic strings'. We will *not* write down the 2d action for those fields, etc, but instead use the simple practical rules to give the final result of physical fields and hamiltonian after light-cone quantization. Recall that upon light-cone quantization of the bosonic theory the physical fields where the bosonic fields associated to the transverse coordianates $X^i(\sigma, t)$, $i = 2, \ldots, D-1$, with hamiltonian given by an infinite set of decoupled harmonic oscillators.

The light-cone quantization for the fermionic sector also leaves the transverse fermionic coordinates $\psi^i(\sigma, t)$, $i = 2, \ldots, D-1$ as the only remaining physical fields. Their hamiltonian corresponds to an infinite set of fermionic harmonic oscillators.

In closed string theories it is possible to carry out the quantization etc independently for left- and right-moving degrees of freedom. This is quite convenient for us, so we split our degrees of freedom in $X_L^i(\sigma+t)$, $\psi_L^i(\sigma+t)$, $X_R^i(\sigma-t)$, $\psi_R^i(\sigma-t)$, and work with just the left moving piece. The level matching constraints etc will be discussed at a later stage.

[1]One may wonder if 2d susy is really necessary to achieve spacetime fermions. In our discussion it would seem that we are emphasizing just the need of worldsheet fermions, and that 2d susy appears as an accidental symmetry in the system of decoupled fermionic and bosonic harmonic oscillators; however it is possible to argue as in the first section of chapter 10 in [71] that the equation of motion for spacetime spinors arises from the conserved supercurrent of the 2d theory. From this viewpoint 2d susy is quite crucial. In fact, even in our simplified discussion spatime fermions are seen to arise from *fermionic zero modes* in the R sector, where the zero point energy exactly vanishes due to 2d susy; hence susy turns out to be crucial as well in our description, although not in a very explicit way.

7.1.2 Boundary conditions

We are interested in discussing closed fermionic strings in flat D-dimensional Minkowski space. To have closed string in flat space, the 2d bosocic fields must be periodic in σ

$$X_L^i(\sigma + t + \ell) = X_L^i(\sigma + t) \tag{7.1}$$

and we have the oscillator expansion

$$X_L^i(\sigma + t) = \frac{x^i}{2} + \frac{p_i}{2p^+}(t+\sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n \in \mathbf{Z}-\{\mathbf{0}\}} \frac{\alpha_n^i}{n} e^{-2\pi i n(\sigma+t)/\ell} \tag{7.2}$$

with modes having commutation relations

$$[x^i, p_j] = i\delta_{ij} \quad ; \quad [\alpha_n^i, \alpha_m^j] = m\,\delta_{ij}\,\delta_{m,-n} \tag{7.3}$$

and hamiltonian

$$\begin{aligned} H_B &= \frac{\sum_i p_i p_i}{4p^+} + \frac{1}{\alpha' p^+}\left[\sum_{n>0} \alpha_{-n}^i \alpha_n^i + E_0^B\right] \\ E_0^B &= -\frac{D-2}{24} \end{aligned} \tag{7.4}$$

For fermions, there is a subtlety in discussing boundary conditions. In the two-dimensional worldsheet field theory, as in any quantum field theory, the only observables are expressions that go like products of two fermion fields. That means thae periodicity in σ of observables is consistent with antiperiodicity of the fermion fields . Hence there are two consistent boundary conditions

$$\begin{array}{lll} \text{Neveu} - \text{Schwarz} & \text{NS} & \psi_L^i(\sigma + t + \ell) = -\psi_L^i(\sigma + t) \\ \text{Ramond} & \text{R} & \psi_L^i(\sigma + t + \ell) = \psi_L^i(\sigma + t) \end{array} \tag{7.5}$$

These can be chosen independently for left and right sectors. It is important to notice that consistency, e.g. Lorentz invariance, already requires that in a given sector, fermions fields ψ_L^i for all i are all periodic or all antiperiodic.

Hence it would seem that we can define four different kinds of closed strings, according to whether the left and right sectors have NS or R fermions;

namely we would have NS-NS, NS-R, R-NS and R-R strings. Very surprisingly, we will see that modular invariance requires these different boundary conditions to coexist within the same theory. In a sense, in the same way that a consistent string theory requires us to sum over different worlsheet topologies (topological sectors of the embedding functions X^i), it also requires us to sum over different topological sectors (boundary conditions) for the 2d fermion fields, in a precise way dictated by the requirement to get a modular invariant partition function. This has been formulated very precisely as a sum over spin structures on the worldsheet [94].

7.1.3 Spectrum of states for NS and R fermions

Before going further, it will be useful to compute the oscillator expansion, hamiltonian and spectrum of states for 2d fermions with NS and R boundary conditions. We describe this for the left-moving sector, being analogous (and independent) for the right-moving one.

NS sector

Antiperiodic boundary conditions require the oscillator modding to be half-integer. We have the oscillator expansion

$$\psi_L^i(\sigma+t) \;=\; i\sqrt{\frac{\alpha'}{2}}\,\sum_{r\in\mathbf{Z}}\psi_{r+1/2}^i\, e^{-2\pi i(r+1/2)(\sigma+t)/\ell} \tag{7.6}$$

Notice that there are no zero modes in the expansion. The oscillators have anticommutation relations

$$\{\psi_{n+1/2}^i,\psi_{m+1/2}^j\} \;=\; \delta^{ij}\,\delta_{m+1/2,-(n+1/2)} \tag{7.7}$$

The hamiltonian for the fermionic degrees of freedom is

$$H_{F,NS} \;=\; \frac{1}{\alpha' p^+}\,\left[\sum_{r=0}^{\infty}(r+\frac{1}{2})\,\psi_{-r-1/2}^i\,\psi_{r+1/2}^i \;+\; E_0^{F_{NS}}\right] \tag{7.8}$$

where the zero point energy for NS fermionic oscillators is

$$E_0^{F_{NS}} \;=\; -\frac{1}{2}\sum_{n=0}^{\infty}(n+\frac{1}{2}) \tag{7.9}$$

evaluated with the exponential regularization. It is useful to compute in general (for $\alpha > 0$)

$$Z_\alpha = \frac{1}{2}\sum_{n=0}^{\infty}(n+\alpha) \tag{7.10}$$

as the $\epsilon \to 0$ limit of the finite part of

$$\begin{aligned}
Z_\alpha(\epsilon) &= \frac{1}{2}\sum_{n=0}^{\infty}(n+\alpha)\,e^{-(n+\alpha)\epsilon} = -\frac{1}{2}\frac{\partial}{\partial\epsilon}\sum_{n=0}^{\infty}e^{-(n+\alpha)\epsilon} = -\frac{1}{2}\frac{\partial}{\partial\epsilon}\left(\frac{e^{-\alpha\epsilon}}{1-e^{-\epsilon}}\right) = \\
&= -\frac{1}{2}\frac{\partial}{\partial\epsilon}\left[(1-\alpha\epsilon+\alpha^2/2\epsilon^2+\mathcal{O}(\epsilon^3))\,(\frac{1}{\epsilon}+\frac{1}{2}+\frac{1}{12}\epsilon+\mathcal{O}(\epsilon^2))\right] = \\
&= -\frac{1}{2}\frac{\partial}{\partial\epsilon}\left[\frac{1}{\epsilon}+\frac{1}{2}+\frac{1}{12}\epsilon-\alpha-\frac{1}{2}\alpha\epsilon+\frac{1}{2}\alpha^2\epsilon+\mathcal{O}(\epsilon^2)\right] = \\
&= \frac{1}{2\epsilon^2}-\frac{1}{24}+\frac{1}{4}\alpha(1-\alpha)+\mathcal{O}(\epsilon)
\end{aligned} \tag{7.11}$$

so we get

$$Z_\alpha = -\frac{1}{24}+\frac{1}{4}\alpha(1-\alpha) \tag{7.12}$$

and

$$E_0^{F_{NS}} = -\frac{1}{48}(D-2) \tag{7.13}$$

The total bosonic and fermionic hamiltonian for the 2d theory in the NS sector is

$$H_L = \frac{\sum_i p_i p_i}{4p^+} + \frac{1}{\alpha' p^+}\left[\sum_{n>0}\alpha^i_{-n}\alpha^i_n + \sum_{r=0}^{\infty}(r+\frac{1}{2})\,\psi^i_{-r-1/2}\,\psi^i_{r+1/2} + (D-2)\frac{-1}{16}\right] \tag{7.14}$$

The contribution of the left-moving sector to the spacetime mass is

$$m_L^2 = 2p^+ H_L - \frac{1}{2}\sum_i p_i^2 \tag{7.15}$$

namely

$$\alpha' m_L^2/2 = \left[\sum_{n>0}\alpha^i_{-n}\alpha^i_n + \sum_{r=0}^{\infty}(r+\frac{1}{2})\,\psi^i_{-r-1/2}\,\psi^i_{r+1/2} - \frac{(D-2)}{16}\right] \tag{7.16}$$

The spectrum in the NS sector is obtained by defining a groundstate $|k\rangle_{NS}$ with spacetime momenta k_i, and annihilated by all positive modding oscillators

$$\begin{aligned} \psi^i_{n+1/2}|k\rangle_{NS} = 0 \quad &; \quad \forall n \geq 0 \\ \alpha^i_n|k\rangle_{NS} = 0 \quad &; \quad \forall n > 0 \end{aligned} \tag{7.17}$$

and applying negative modding oscillators in all possible ways.

The lightest left moving states (for zero spacetime momentum) are

$$\begin{array}{cc} \text{State} & \alpha' m_L^2/2 \\ |0\rangle_{NS} & -\frac{(D-2)}{16} \\ \psi^i_{-1/2}|0\rangle_{NS} & \frac{1}{2} - \frac{(D-2)}{16} \end{array} \tag{7.18}$$

Now we realize that the first excited state is a vector with respect to spactime Lorentz transformations, and that it only has $D-2$ components. So it forms a representation of the group $SO(D-2)$, which is the little group of a massless particle in a Lorentz invariant D-dimensional theory. This means that in order to be consistent with Lorentz invariance, the state should be massless, and this requires $(D-2)/16 = 1/2$, namely $D = 10$. Namely we obtain the result that the string theory at hand propagates consistently only in a spacetime of *ten* dimensions.

The states we have transform under the $SO(8)$ group as

State	$\alpha' m_L^2/2$	$SO(8)$
$\|0\rangle_{NS}$	$-1/2$	$\mathbf{1}$
$\psi^i_{-1/2}\|0\rangle_{NS}$	0	$\mathbf{8_V}$

where 8_V is the vector representation of $SO(8)$.

Ramond sector

Periodic boundary conditions require integer modding for fermionic oscillators

$$\psi^i_L(\sigma+t) \, = \, i\sqrt{\frac{\alpha'}{2}} \sum_{r\in\mathbf{Z}} \psi^i_r \, e^{-2\pi i r(\sigma+t)/\ell} \tag{7.19}$$

An important difference with respect to the NS sector is the existence of fermion zero modes ψ^i_0.

The anticommutation relations read

$$\{\psi^i_n, \psi^j_m\} = \delta^{ij}\, \delta_{m,-n} \tag{7.20}$$

The hamiltonian for the fermionic degrees of freedom is

$$H_{F,R} = \frac{1}{\alpha' p^+} \left[\sum_{r=1}^{\infty} r\, \psi^i_{-r}\, \psi^i_r \,+\, E_0^{F_R} \right] \tag{7.21}$$

with $E_0^{F_R} = (D-2)\times(-1/2)\sum_{r=1}^{\infty} r$, which for $D=10$ equals $E_0^{F_R} = 8\times\frac{1}{24}$. The total bosonic plus fermionic zero point energies cancel in the R sector [2]

The total bosonic and fermionic hamiltonian for the 2d theory in the NS sector is

$$H_L = \frac{\sum_i p_i p_i}{4p^+} + \frac{1}{\alpha' p^+} \left[\sum_{n>0} \alpha^i_{-n}\alpha^i_n \,+\, \sum_{r=1}^{\infty} r\, \psi^i_{-r}\, \psi^i_r \right] \tag{7.22}$$

The contribution of the left-moving sector to the spacetime mass is

$$m_L^2 = 2p^+ H_L - \frac{1}{2}\sum_i p_i^2 \tag{7.23}$$

namely

$$\alpha' m_L^2/2 = \left[\sum_{n>0} \alpha^i_{-n}\alpha^i_n \,+\, \sum_{r=0}^{\infty} r\, \psi^i_{-r}\, \psi^i_r \right] \tag{7.24}$$

To compute the spectrum we have to be careful with the definition of the ground state, because of fermion zero modes. Given a groundstate, application of some ψ^i_0 costs no energy and we get another groundstate. The system has a degenerate set of groundstates, and we have to find how the fermionic operators act on them. Clearly we can require that positive modding operators annihilate it; however we cannot require that all fermionic zero modes annihilate it, since this is not consistent with the zero mode anticommutators

$$\{\psi^i_0, \psi^j_0\} = \delta^{ij} \tag{7.25}$$

[2] In the NS sector the local 2d susy is broken by the different boundary conditions betwen bosons and fermions, leaving a finite zero point energy contribution; in the R sector the 2d susy is globally preserved by the boundary conditions, so the zero point energies cancel.

which is a Clifford algebra (see section 6 of the lesson on group theory). In fact, defining the action of the ψ_0^i on the set of groundstates is constructing a representation of the corresponding Cliffor algebra

By now we know that to construct such a representation we should define the operators

$$A_a^{\pm} = \psi_0^{2a} \pm i\psi_0^{2a-1} \quad \text{for} \;\; a = 1, \ldots, 4 \tag{7.26}$$

define a lowest weight state by $A_a^-|0\rangle = 0$, and build the set of states by application of the A_a^+ operators

$$\begin{array}{cc} |0\rangle & A_{a_1}^+|0\rangle \\ A_{a_1}^+ A_{a_2}^+|0\rangle & A_{a_1}^+ A_{a_2}^+ A_{a_3}^+|0\rangle \\ A_1^+ A_2^+ A_3^+ A_4^+|0\rangle & \end{array} \tag{7.27}$$

A representation of the Clifford algebra splits into two spinor representations, of different chiralities, of the $SO(8)$ Lie algebra. These correspond to the two above columns; we denote the corresponding states by $\mathbf{8_S}$ and $\mathbf{8_C}$, or equivalently by the corresponding weights $\frac{1}{2}(\pm,\pm,\pm,\pm)$ with the number of $-$'s even for $\mathbf{8_S}$ and odd for $\mathbf{8_C}$.

The Hilbert space in the R sector is obtained by applying the negative modding operators to these groundstates in all possible ways. At the massless level, the only states are the groundstates, transforming under $SO(8)$ as

$$\mathbf{8_S} + \mathbf{8_C} \tag{7.28}$$

Our results, to summarize, are that the light modes in the NS and R sectors are

	State	$\alpha' m_L^2/2$	$SO(8)$
NS	$\vert 0\rangle_{NS}$	$-1/2$	$\mathbf{1}$
	$\psi_{-1/2}^i \vert 0\rangle_{NS}$	0	$\mathbf{8_V}$
R	$\frac{1}{2}(\pm,\pm,\pm,\pm)$ $\#- =$ even	0	$\mathbf{8_S}$
	$\frac{1}{2}(\pm,\pm,\pm,\pm)$ $\#- =$ odd	0	$\mathbf{8_C}$

We can choose these states independently for left and right movers. We now need to discuss how to glue them together to form physical states. One conditions is the level matching constraint, which amounts to

$$m_L^2 = m_R^2 \tag{7.29}$$

The glueing is also constrained from modular invariance. Namely, a string in one of these sectors, namely NS for left movers and NS for right movers, is *not* modular invariant.

The real, physical, string theories are formed by combining NS and R sectors in a way consisntent with modular invariance. In a sense we need to sum over boundary conditions for the fermions, i.e. combine the spectra of different sectors.

7.1.4 Modular invariance

We would like to discuss the partition function

$$Z(\tau) = \mathrm{tr}_{\mathcal{H}} \left(e^{-\tau_2 \ell H} \, e^{i\tau_1 \ell P} \right) \tag{7.30}$$

In order to keep discussion about left and right movers independently it is useful to recall that the trace over the physical level-matched Hilbert space of a string theory can be extended to a trace over an unconstrained Hilbert space, with independend left and right sectors, with level matching imposed upon integration of the τ_1 piece of the modular parameter (see lesson on modular invariance).

Using that

$$H = \frac{\sum_i p_i^2}{2\alpha' p^+} + H_L + H_R \quad ; \quad P = H_L - H_R \tag{7.31}$$

with $H_L = \frac{1}{\alpha' p^+}(N + E_0)$, $H_R = \frac{1}{\alpha' p^+}(\tilde{N} + \tilde{E}_0)$, the expression for the partition function can be written as

$$\begin{aligned} Z(\tau) &= \mathrm{tr}_{\mathcal{H}} \, e^{-\pi\alpha'\tau_2 \sum p_i^2} \, q^{N+E_0} \, \overline{q}^{\tilde{N}+\tilde{E}_0} = \mathrm{tr}_{\mathcal{H}_{c.m.}} \, e^{-\pi\alpha'\tau_2 \sum p_i^2} \, \mathrm{tr}_{\mathcal{H}_L} \, q^{N+E_0} \, \mathrm{tr}_{\mathcal{H}_R} \, \overline{q}^{\tilde{N}+\tilde{E}_0} = \\ &= (4\pi^2 \alpha' \tau_2)^{-4} \, \mathrm{tr}_{\mathcal{H}_L} \, q^{N+E_0} \, \mathrm{tr}_{\mathcal{H}_R} \, \overline{q}^{\tilde{N}+\tilde{E}_0} \end{aligned} \tag{7.32}$$

where factorization follows from considering the left and right movers independently.

Within each sector we have such factorization. We would now like to compute the left movers partition functions for NS and R boundary conditions. At this point, it will be useful to recall some useful modular functions, (see appendix of the lesson on modular invariance), which we gather in the appendix.

The partition function in the left sector contains a trace over the bosonic oscillators, which is computed just like in bosonic string theory

$$\mathrm{tr}_{\mathcal{H}_{bos}} q^{N_B+E_0^B} = \eta(\tau)^{-8} \quad (7.33)$$

To obtain the partition function over the infinite set of fermionic oscillators, consider first the simplified situation of the partition function of a single fermionic harmonic oscillator. It has just two states, the vacuum $|0\rangle$ and $\psi_{-\nu}|0\rangle$, where ν denotes the oscillator moding. For this system we have

$$\mathrm{tr}_{\mathcal{H}} q^{N_F+E_0^F} = q^{E)^F} (1+q^{\nu}) \quad (7.34)$$

For several decoupled fermionic harmonic oscillators, we simply get the product of partition functions for the individual ones.

NS fermions

Using this, the partition function for 8 NS fermionic coordinates is the product of partiton functions for eight infinite sets of fermionic harmonic oscillators with half-integer moddings $n+1/2$, namely

$$\mathrm{tr}_{\mathcal{H}_{NS}} q^{N_F+E_0^F} = \left[q^{-1/48} \prod_{n=1}^{\infty} (1+q^{n-1/2}) \right]^8 = \frac{\vartheta \begin{bmatrix} 0 \\ 0 \end{bmatrix}^4}{\eta^4} \quad (7.35)$$

R fermions

This is the product over the partition function of eight infinite sets of fermionic harmonic oscillator with integer modding, times the multiplicity of 16 due to the degenerate ground state, namely

$$\mathrm{tr}_{\mathcal{H}_R} q^{N_F+E_0^R} = 16 \left[q^{1/24} \prod_{n=1}^{\infty} (1+q^{n}) \right]^8 = \frac{\vartheta \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}^4}{\eta^4} \quad (7.36)$$

Now we easily observe that modular transformations may mix different boundary conditions, and even require the introduction of new pieces in the partition function. For instance

$$\frac{\vartheta \begin{bmatrix} 0 \\ 0 \end{bmatrix}^4}{\eta^4} \xrightarrow{\tau\to\tau+1} \frac{\vartheta \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}^4}{\eta^4} \xrightarrow{\tau\to -1/\tau} \frac{\vartheta \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}^4}{\eta^4} \xrightarrow{\tau\to -1/\tau} \quad (7.37)$$

Clearly a modular invariant partition function must be a sum over sectors with different boundary conditions.

7.1.5 Type II superstring partition function

Instead of working by trial and error, let us simply give the final result of a possible modular invariant partition function, and then interpret it in terms of the physical spectrum of the theory.

Consider the two partition functions for left movers

$$Z_\pm = \frac{1}{2}\,(4\pi^2\alpha'\tau_2)^{-4}\,\eta^{-8}\,\eta^{-4}\left(\vartheta\begin{bmatrix}0\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^4 - \vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^4 \pm \vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^4\right) \quad (7.38)$$

The first piece is half of the contribution from spacetime momenta, then follows the piece from left bosonic oscillators, then the piece from left fermionic oscillators. Either of the two choices is invariant under $\tau \to -1/\tau$, and they transform as $Z_\pm \to -Z_\pm$ under $\tau - \to \tau + 1$. Therefore, it is possible to cook up several modular invariant partition functions for the complete left times right theory. Namely we consider the partition functions

$$Z_+\overline{Z}_+ \quad ; \quad Z_-\overline{Z}_- \quad ; \quad Z_+\overline{Z}_- \quad ; \quad Z_-\overline{Z}_+ \qquad (7.39)$$

This means that there are four consistent string theories! (in fact, we will see later on that there are only two inequivalent ones).

7.1.6 GSO projection

It is now time to address the question of what is the meaning of pieces like $\vartheta\begin{bmatrix}0\\1/2\end{bmatrix}$ or $\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}$ in the partition function. For NS fermions it is easy to realize that

$$\begin{aligned}
\eta^{-1}\,\vartheta\begin{bmatrix}0\\0\end{bmatrix} &= q^{-1/48}\prod_{n=1}^{\infty}(1+q^{n-1/2})^2 = \mathrm{tr}_{\mathcal{H}_{NS}}\, q^{N+E_0^F} \\
\eta^{-1}\,\vartheta\begin{bmatrix}0\\1/2\end{bmatrix} &= q^{-1/48}\prod_{n=1}^{\infty}(1-q^{n-1/2})^2 = \mathrm{tr}_{\mathcal{H}_{NS}}\, q^{N+E_0^F}\,(-1)^F
\end{aligned} \qquad (7.40)$$

On the second line we sum over NS fermions, weighting each fermionic oscillator mode by a minus sign; this can be implemented in the trace as the insertion of an operator $(-1)^F$ which anticommutes with all fermionic oscillator operators.

Using this, we are now ready to interpret the meaning of one of the pieces of the left partition functions $Z_{\pm}$. Namely

$$\eta^{-4}\left(\vartheta\begin{bmatrix}0\\0\end{bmatrix}^4-\vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^4\right) = \tfrac{1}{2}\,\mathrm{tr}_{\mathcal{H}_{NS}}\,q^{N+E_0^F}\;-\;\mathrm{tr}_{\mathcal{H}_{NS}}\,q^{N+E_0^F}\,(-)^F = \\ = \mathrm{tr}_{\mathcal{H}_{NS}}\,q^{N+E_0^F}\,\tfrac{1}{2}(1-(-)^F) \qquad (7.41)$$

The operator $\frac{1}{2}(1-(-)^F)$ is a projector that allows to propagate only modes with an odd number of fermionic oscillators. This piece of the partition function traces over 8 fermions with NS boundary conditions, projecting out modes with an even number of fermionic oscillators. This is the GSO projection in the NS sector.

The effect on the light NS states is to remove the tachyonic groundstate $|0\rangle_{NS}$ from the physical spectrum, and leave the states $\psi^i_{-1/2}|0\rangle_{NS}$.

Similarly, the remaining pieces of the partition function correspond to

$$\eta^{-4}\left(\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^4\pm\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^4\right) = \tfrac{1}{2}\,\mathrm{tr}_{\mathcal{H}_{R}}\,q^{N+E_0^F}\;\pm\;\mathrm{tr}_{\mathcal{H}_{R}}\,q^{N+E_0^F}\,(-)^F = \\ = \mathrm{tr}_{\mathcal{H}_{R}}\,q^{N+E_0^F}\,\tfrac{1}{2}(1\pm(-)^F) \qquad (7.42)$$

which implements a GSO projection on the R sector. Namely, for the partition function Z_+ the GSO projection leaves states with even number of excitations over the groundstate 8_C and states with odd number of excitations over the groundstate 8_S (and projects out other possibilities); while Z_- leaves states with odd number of excitations over the groundstate 8_C and states with even number of excitations over the groundstate 8_S (and projects out other possibilities).

7.1.7 Light spectrum

The product form of the left times right partition function implies that left NS and R sectors can combine with right NS and R sectors. More explicitly, the fermionic piece of the partition function has the structure

$$\begin{aligned} Z_\psi(\tau) &= (\,\mathrm{tr}_{\mathcal{H}_{NS,GSO_-}} - \mathrm{tr}_{\mathcal{H}_{R,GSO_-}}\,)\times(\,\mathrm{tr}_{\mathcal{H}_{NS,GSO_-}} - \mathrm{tr}_{\mathcal{H}_{R,GSO_\pm}}\,)^* = \\ &= \mathrm{tr}_{\mathcal{H}_{NS,GSO_-}}\mathrm{tr}^*_{\mathcal{H}_{NS,GSO_-}} - \mathrm{tr}_{\mathcal{H}_{NS,GSO_-}}\mathrm{tr}^*_{\mathcal{H}_{R,GSO_\pm}} - \\ &\quad - \mathrm{tr}_{\mathcal{H}_{R,GSO_-}}\mathrm{tr}^*_{\mathcal{H}_{NS,GSO_-}} + \mathrm{tr}_{\mathcal{H}_{R,GSO_-}}\mathrm{tr}^*_{\mathcal{H}_{R,GSO_\pm}} \end{aligned} \qquad (7.43)$$

where the subindex $GSO_\pm$ implies we trace only over the states surviving the GSO projection $\frac{1}{2}(1 \pm (-)^F)$. Notice the minus sign in the contributions from the NS-R and R-NS partition function, which implies that loops of the corresponding spacetime fields are weighted with a minus sign, namely they are fermions. We will see that these states have half-integer spin, so these string theories automatically implement the spin-statistics relation.

We discuss the light (in fact massless) spectrum of the theories in what follows.

Type IIB superstring

Consider the theory described $Z_+\overline{Z}_+$. Using the above projections, it is easy to realize that (both for left and right sectors) the massless NS states are simply the $\psi^i_{-1/2}|0\rangle$, transforming in the 8_V, while in the R sector the states surviving the GSO projection transform as 8_C. These states can be glued together satisfying the level matching condition.

The $SO(8)$ representation of the complete states is obtained by tensoring the representations of the left and right pieces. Hence we have

NS-NS	$8_V \otimes 8_V$	$1 + 28_V + 35_V$
NS-R	$8_V \otimes 8_C$	$8_S + 56_S$
R-NS	$8_C \otimes 8_V$	$8_S + 56_S$
R-R	$8_C \otimes 8_C$	$1 + 28_C + 35_C$

The NS-NS sector contains an scalar (dilaton), a 2-index antisymmetric tensor (2-form $B_{\mu\nu}$), and a 2-index symmetric tensor (graviton $G_{\mu\nu}$).

The R-NS and NS-R sectors contain fermions, in fact the 56_S arising from a vector and a spinor under $SO(8)$ is a gravitino (a spin 3/2 particle).

The RR sector contains a bunch of p-forms, namely p-index completely antisymmetric tensors. In particular, a 0-form (scalar) a, a 2-form $\tilde{B}_2$, and a 4-form (of self-dual field strength) A_4^+. It is sometimes convenient to introduce the Hodge duals of these, which are a 6-form B_6, an 8-form C_8. Finally, it is also useful to introduce a 10-form C_{10}, which does not have any propagating degrees of freedom, since it has no spacetime kinetic term (since its field strength would be a 11-form in 10d spacetime).

The theory is invariant under spacetime coordinate reparametrization, and gauge transformations of the p-forms. It is also invariant under local supersymmetry. It is easy to verify from the tables in [46] that the massless spectrum is that of 10d $N = 2$ chiral supergravity. String theory is providing a finite ultraviolet completion of this supergravity theory, remarkable indeed!

Finally, this theory is chiral in 10d, and has potential gravitational anomalies. It was checked in [45] that the chiral sector of the theory is precisely such that all anomalies automatically cancel (in a very non-trivial, almost miraculous, way).

This is the TYPE IIB superstring.

Consider now the theory described by $Z_-\overline{Z}_-$. It is similar to the above by simply exchanging $C \leftrightarrow S$ in the $SO(8)$ representations. Hence, clearly the two theories are the same up to a redefinition of what we mean by left and right chirality in 10d (namely, up to a parity transformation). So we do not obtain a new theory from $Z_-\overline{Z}_-$. Similarly $Z_-\overline{Z}_+$ and $Z_+\overline{Z}_-$ are related, and is enough to study just one of them.

Type IIA superstring

Consider the theory described $Z_+\overline{Z}_-$. Using the above projections, the massless sector is

NS-NS	$8_V \otimes 8_V$	$1 + 28_V + 35_V$
NS-R	$8_V \otimes 8_S$	$8_C + 56_C$
R-NS	$8_C \otimes 8_V$	$8_S + 56_S$
R-R	$8_C \otimes 8_S$	$8_V + 56_V$

The NS-NS sector contains an scalar (dilaton), a 2-index antisymmetric tensor (2-form $B_{\mu\nu}$), and a 2-index symmetric tensor (graviton $G_{\mu\nu}$).

The R-NS and NS-R sectors contain fermions, in fact the 56_S, 56_C arising from a vector and a spinor under $SO(8)$ are gravitinos (a spin 3/2 particle).

The RR sector contains a bunch of p-forms, namely p-index completely antisymmetric tensors. In particular, a 1-form (scalar) A_1, and a 3-form C_3. It is sometimes convenient to introduce the Hodge duals of these, which are a 5-form C_5, a 7-form A_7. Finally, it is also useful to introduce a 9-form C_9, which does not contain much dynamics (and is related to Romans massive IIA supergravities [76]).

The theory is invariant under spacetime coordinate reparametrization, and gauge transformations of the p-forms. It is also invariant under local supersymmetry. It is easy to verify from the tables in [46] that the massless spectrum is that of 10d $N = 2$ non-chiral supergravity. String theory is providing a finite ultraviolet completion of this supergravity theory, remarkable indeed!

Finally, this theory is non-chiral in 10d, hence is automatically anomaly free.

This is the TYPE IIA superstring.

Some comments

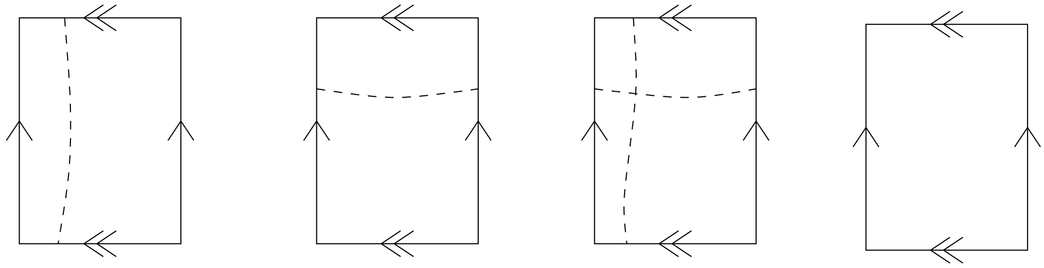


Figure 7.1: The four theta function contributions to the partition function can be understood as for possible boundary conditions in σ and t for fermions in a 2-torus. (Anti)periodicity in t is correlated with the presence of $(-)^F$ insertions in the trace, while (anti)periodicity in σ is correlated with the choice of NS or R fermions. Clearly modular transformations relate different contributions, so that a modular invariant theory needs to combine all of them.

• The construction we have described seems a bit intricate. However, it follows naturaly from the underlying worldsheet geometry of the string, namely from modular invariance, i.e. invariance under (large) coordinate transformations on the worldsheet. The reason why modular transformations mix different boundary conditions can be understood intuitively from figure 7.1: Starting with a GSO projected trace over NS states, the piece involving the $(-1)^F$ insertion implies that 2d fermions pick up a minus sign as they evolve in t; upon the modular transformation $\tau \to -1/\tau$, we obtain that fermions pick up an additional sign as σ varies, namely the boundary condition is not NS any longer, but is flipped to R in this sector. All contributions in the partition function may be understood in this language.

• We re-emphasize that the appearance of spacetime fermions is subtle, and is not automatically obtained from the existence of 2d fermions. Indeed, in the NS sector we have 2d fermions but no spacetime fermions. Similarly, the existence of spacetime supersymmetry does not automatically follow from 2d susy, rather it is implemented due to the GSO projection. This is one of the remarkable features of string theory, the deep relation between physics of the worldsheet (modular invariance, etc) and spacetime physics (spacetime susy).

• Spacetime supersymmetry is not manifest in the formalism we have described. It would be nice to find a formalism which describes type II superstring, and which makes spacetime supersymmetry manifest. Intuitively, we would like to describe the worldsheet theory by describing string configurations by an embedding of the worldsheet into 10d superspace, namely a set of embedding superfunctions $(X^\mu(\sigma,t),\Theta^\alpha(\sigma,t))$, where Θ^α transform in the spinor representation of the spacetime Lorentz group and parametrize the fermionic dimensions of superspace. Such a formulation exists and is known as the Green-Schwarz superstrings. For type II theories it is equivalent to the formulation we used (called the NSR formulation), but it is more difficult in some respects. Some useful comments on it may be found in section 12.6 in [71].

• Recall that the partition function is the vacuum energy of the spacetime theory. Spacetime supersymmetry implies that the spectrum is fermion/boson degenerate, and that this vacuum energy vanishes. Indeed, the theta functions satisfy the 'abstruse identities'

$$\vartheta\begin{bmatrix}0\\0\end{bmatrix}^4-\vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^4-\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^4=0\quad;\quad\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}=0 \qquad (7.44)$$

So the 1-loop cosmological constant vanishes in these theories.

• If the partition function is exactly zero, why should we bother about whether it is modular invariant or not?? The key observation is that modular invariance of the vacuum amplitude (without use of abstruse identities) guarantees that other more complicated amplitudes (with external legs) are also invariant under large coordinate reparametrizations on the worldsheet.

• Recall that the contribution $Z(\tau)$ must be integrated over the fundamental domain in τ to get the complete contribution. As discussed in the bosonic theory, the ultraviolet region is related, namely is equivalent geometrically, to the infrared region. A difference with the bosonic theory is that the type II superstrings do not contain tachyons, so there are no infrared divergences.

7.2 Type 0 superstrings

We would like to discuss (the only) other possible modular invariant partition functions that one can construct with the basic building blocks we have, namely the 2d fields of the (2d supersymmetric) strings. Interestingly enough, the theories we are about to construct, called type 0 theories, are *not* spacetime supersymmetric, and moreover do *not* contain spacetime fermions.

So they clearly illustrate the fact that 2d fermions/susy do not guarantee spacetime fermions/susy.

The complete left times right partition function is given by

$$Z_{\pm} = \frac{1}{2}(4\pi^2\alpha'\tau_2)^{-4}\,|\eta|^{-16}|\eta|^{-8}\left(\left|\vartheta\begin{bmatrix}0\\0\end{bmatrix}\right|^8 + \left|\vartheta\begin{bmatrix}0\\1/2\end{bmatrix}\right|^8 + \left|\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}\right|^8 \pm \left|\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}\right|^8\right) \quad (7.45)$$

We obtain two new inequivalent theories, whose structure in the fermionic partition function is

$$\begin{aligned} &\mathrm{tr}\,_{\mathcal{H}_{NS,GSO_+}}\mathrm{tr}\,^*_{\mathcal{H}_{NS,GSO_+}} + \mathrm{tr}\,_{\mathcal{H}_{NS,GSO_+}}\mathrm{tr}\,^*_{\mathcal{H}_{NS,GSO_-}} + \\ &+\mathrm{tr}\,_{\mathcal{H}_{R,GSO_+}}\mathrm{tr}\,^*_{\mathcal{H}_{NS,GSO_\pm}} + \mathrm{tr}\,_{\mathcal{H}_{R,GSO_-}}\mathrm{tr}\,^*_{\mathcal{H}_{R,GSO_\mp}} \end{aligned} \quad (7.46)$$

The lightest modes of the two theories are

Type 0A

Sector	States	$SO(8)$	$\alpha' m^2$	Fields
NS-NS	$1\otimes 1$	1	-2	T
	$8_V\otimes 8_V$	$1+28_V+35_V$	0	$\phi,\, B_2,\, G_{\mu\nu}$
R-R	$8_C\otimes 8_S$	8_V+56_V	0	$A_1,\, C_3$
	$8_S\otimes 8_C$	8_V+56_V	0	$A_1',\, C_3'$

Type 0A

Sector	States	$SO(8)$	$\alpha' m^2$	Fields
NS-NS	$1\otimes 1$	1	-2	T
	$8_V\otimes 8_V$	$1+28_V+35_V$	0	$\phi,\, B_2,\, G_{\mu\nu}$
R-R	$8_C\otimes 8_C$	$1+28_C+35_C$	0	$a,\, \tilde{B}_2,\, A_4^+$
	$8_S\otimes 8_S$	$1+28_S+35_S$	0	$a',\, \tilde{B}_2',\, A_4^-$

The theories contain a tachyon in the NS-NS sector. As usual, one interprets the tachyon as an instability of the theory, which is sitting at the top of some potential for the corresponding field. There are many speculations on what is the stable vacuum of type 0 theories, and even whether it exists or not. The issue remains for the moment as an open question.

Due to this feature, and to lack of fermions, most research is centered on type II strings, rather than type O.

7.3 Bosonization*

We would like to finish with some comments on bosonization. Bosonization/fermionization is a phenomenon relating certain two-dimensional field theories; it is the complete physical equivalence of a 2d quantum field theory with bosonic degrees of freedom and one with fermionic degrees of freedom. This can happen two dimensions since all representations of the $SO(2)$ Lorentz group are one-dimensional, there is no real concept of spin.

For our simplified discussion, we will be interested in discussing simply the equivalences of partition functions of the corresponding 2d theories. But let us emphasize that bosonization/fermionization is complete equivalence of all physical quantities in both theories). Notice however that equivalence of partition functions implies a one-to-one map between states in the two Hilbert spaces, and agreement in their energies.

A simple example of bosonization/fermionization is that the 2d theory of two left-moving free fermions (with NS boundary conditions on the circle) is equivalent to the 2d theory of one left-moving boson compactified on a circle of radius $R = \sqrt{\alpha'}$. Indeed, let us compute the partition function of the theory with two fermions

$$Z_{2\psi} = \left[q^{-1/48} \prod_{n=1}^{\infty} (1 - q^{n-1/2}) \right]^2 = \frac{\vartheta \begin{bmatrix} 0 \\ 0 \end{bmatrix} (\tau)}{\eta(\tau)} \tag{7.47}$$

This final expression can be rewritten using (A.5) as

$$\frac{1}{\eta \tau} \sum_{n \in \mathbf{Z}} q^{n^2/2} \tag{7.48}$$

which corresponds to the partition function of one left-moving boson parametrizing a compact direction of radius $\sqrt{\alpha'}$. The η corresponds to the trace over the oscillator degrees of freedom, while the sum over n corresponds to the sum over left-moving momentum p_L. Finally, purely left-moving bosons with no right-moving partner have no center of mass degrees of freedom, so there is no trace over center of mass momentum. Some of these issues will appear back in the study of the heterotic.

Using this kind of computations, it is possible to bosonize the complete left-moving sector of a type II superstring. Indeed it is possible to recast the left-moving fermion partition function in terms of a bosonic interpretation. In fact, starting with the GSO projected fermionic partition function

$$Z_{\pm} = \frac{1}{2} \eta^{-4} \left(\vartheta \begin{bmatrix} 0 \\ 0 \end{bmatrix}^4 - \vartheta \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}^4 - \vartheta \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}^4 \pm \vartheta \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}^4 \right) \tag{7.49}$$

and writing the ϑ functions as infinite sums, we obtain

$$\begin{aligned} Z_{\pm} &= \frac{1}{2}\,\eta^{-4}\left(\sum_{n_1,n_2,n_3,n_4} q^{\sum_i n_i^2} - \sum_{n_1,n_2,n_3,n_4} q^{\sum_i n_i^2}\, e^{\pi i \sum_i n_i} - \right. \\ &\left. - \sum_{n_1,n_2,n_3,n_4} q^{\sum_i (n_i+1/2)^2} \pm \sum_{n_1,n_2,n_3,n_4} q^{\sum_i (n_i+1/2)^2}\, e^{\pi i \sum_i (n_i+1/2)}\right) \end{aligned}$$

By gathering terms we may write

$$Z_{\pm} = \eta^{-4}\left(\sum_{\vec{r}=(n_1,n_2,n_3,n_4)} q^{\vec{r}^2}\,\frac{1}{2}(1-(-1)^{\sum_i n_i}) - \sum_{r=(n_1+1/2,\ldots,n_4+1/2)} q^{\vec{r}^2}\,\frac{1}{2}(1\pm(-1)^{\sum_i n_i})\right)$$

Defining lattices $\Lambda^{\pm}$ of vectors of the form

$$\begin{array}{llll} (n_1,n_2,n_3,n_4) & ; & n_i\in\mathbf{Z} & ; \quad \sum_{\mathbf{i}} \mathbf{n_i} = \text{odd} \\ (n_1+\frac{1}{2},n_2+\frac{1}{2},n_3+\frac{1}{2},n_4+\frac{1}{2}) & ; & n_i\in\mathbf{Z} & ; \quad \sum_{\mathbf{i}} \mathbf{n_i} = \text{odd, even for}\ \mathbf{\Lambda}^+,\ \mathbf{\Lambda}^- \end{array} \tag{7.50}$$

we can write

$$Z_{\pm} = \eta^{-4}\sum_{r\in\Lambda^{\pm}} q^{r^2} \tag{7.51}$$

Which corresponds to the partition function of four left-moving bosons parametrizing a four-torus defined by the lattice $\Lambda^{\pm}$. Recall that this is not a fake trick, but a complete physical equivalence of 2d theories.

We will not use much this bosonic description. However, it is sometimes used in discussing more complicated models, like orbifolds, since it provides an easy bookeeping of the GSO projections in terms of a lattice.

Chapter 8

Heterotic superstrings

8.1 Heterotic superstrings in bosonic formulation

8.1.1 Heteroticity

We have discussed that in closed string theories the left and right moving sectors have independent hamiltonian evolution. The only relation between both is in the construction of physical states, the level matching conditions.

We have also discussed two consistent (say, left moving) sectors. That of the bosonic string, given (in the light-cone gauge) by 24 2d bosons $X_L^i(\sigma+t)$, $i = 2, \ldots, 25$ and that of the superstring, given by 8 bosons $X_L^i(\sigma + t)$ and 8 fermions $\psi_L^i(\sigma + t)$, $i = 2, \ldots, 9$.

The basic idea in the construction of the heterotic string theories is to consider using the bosonic 2d content for the left moving sector and the superstring 2d content for the right moving sector [1]. Let us denote our right movers by $X_R^i(\sigma - t)$ $\psi_R^i(\sigma - t)$, and our left movers by $X_L^i(\sigma + t)$, $X_L^I(\sigma + t)$, with $i = 2, \ldots, 9$, $I = 1, \ldots, 16$.

The theory is rather peculiar at first sight. The left moving bosons $X_L^i(\sigma + t)$ can combine with the right moving ones $X_R^i(\sigma - t)$ to make out

[1]That this can be done is already very non-trivial. In a Polyakov description we are coupling a 2d *chiral* field theory (since it is not invariant under 2d parity, i.e. exchange of left and right) to a 2d metric. In order for the path integral over 2d metrics to be well defined the 2d field theory must be free of 2d gravitational anomalies; this is true precisely for the matter content of left and right moving degrees of freedom that we have proposed.

the coordinates of physical spacetime (which therefore has ten dimensions). On the other hand, it is not clear what meaning the remaining left moving bosons $X_L^I(\sigma+t)$ have. We will see that, in a precise sense to be explained below, they do not correspond to physical spacetime dimensions, but rather should be though of as parametrizing a 16d compact torus, with very small and fixed radius $R=\sqrt{\alpha'}$. Since this distance is of order the string scale, it is not very meaningful to assing a geometric interpretation to the corresponding dimensions.

8.1.2 Hamiltonian quantization

The worldsheet action is the expected one, namely the Polyakov action for left and right movers independently, with the right moving sector coupling also to a 2d gravitino. Since we will be interested in the light cone quantization, we simply say that it proceeds as usual, and that the only physical fields left over are those mentioned above. We now review the main features

Right movers

In the right moving sector, bosons parametrize non-compact directions, so they must be periodic in σ

$$X_R^i(\sigma-t+\ell) \,=\, X_R^i(\sigma-t) \tag{8.1}$$

They have the usual integer mode expansion

$$X_R^i(\sigma-t) \;=\; \frac{x^i}{2} \,+\, \frac{p_i}{2p^+}\,(t-\sigma) \,+\, i\sqrt{\frac{\alpha'}{2}}\sum_{n\in\mathbf{Z}-\{\mathbf{0}\}}\frac{\tilde{\alpha}_n^i}{n}\,e^{2\pi i\,n\,(\sigma-t)/\ell} \tag{8.2}$$

Right moving fermions can be either periodic (R) or antiperiodic (NS)

$$\begin{array}{ll} \text{NS} & \psi_L^i(\sigma+t+\ell) = -\psi_L^i(\sigma+t) \\ \text{R} & \psi_L^i(\sigma+t+\ell) = \psi_L^i(\sigma+t) \end{array} \tag{8.3}$$

so we have the mode expansion

$$\psi_R^i(\sigma-t) \,=\, i\sqrt{\frac{\alpha'}{2}}\sum_{r\in\mathbf{Z}}\tilde{\psi}_{r+\nu}^i\,e^{2\pi i(r+\nu)(\sigma-t)/\ell} \tag{8.4}$$

with $\nu=0,1/2$ for R, NS fermions.

The complete right moving hamiltonian is

$$H_R = \frac{\sum_i p_i p_i}{4p^+} + \frac{1}{\alpha' p^+}\,(\tilde{N}_B + \tilde{N}_F + \tilde{E}_0)$$
$$\tilde{N}_B = \sum_{n>0} \tilde{\alpha}^i_{-n}\tilde{\alpha}^i_n \quad ; \quad \tilde{N}_F = \sum_{r=0}^{\infty} (r+\nu)\,\psi^i_{-r-\nu}\,\psi^i_{r+\nu} \quad ; \quad \tilde{E}_0 = -2\nu(1-\nu) \tag{8.5}$$

Left movers

For the left sector, the bosons $X_L^i(\sigma+t)$ are paired with the right moving bosons, so they are pediodic

$$X_L^i(\sigma+t+\ell) = X_L^i(\sigma+t) \tag{8.6}$$

and have a mode expansion

$$X_L^i(\sigma+t) = \frac{x^i}{2} + \frac{p_i}{2p^+}\,(t+\sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbf{Z}-\{\mathbf{0}\}} \frac{\alpha_n^i}{n}\, e^{-2\pi i\, n\,(\sigma+t)/\ell} \tag{8.7}$$

We now need to propose mode expansions for the remaining left moving bosons $X^I(\sigma+t)$. To put it in a heuristic way, we propose a mode expansion that corresponds to the left moving sector of a bosonic theory compactified on a 16d torus, consistently with making the corresponding right moving degrees of freedom identically vanish.

Namely, recall the mode expansion for left and right moving bosons in a circle compactification of the bosonic theory (see lesson on toroidal compactification), in the sector of momentum k and winding w (k,

$$X_L(\sigma+t) = \tfrac{x}{2} + \tfrac{p_L}{2p^+}\,(t+\sigma) + i\sqrt{\tfrac{\alpha'}{2}} \textstyle\sum_{n\in\mathbf{Z}-\{\mathbf{0}\}} \tfrac{\alpha_n^i}{n}\, e^{-2\pi i\, n\,(\sigma+t)/\ell}$$
$$X_R(\sigma-t) = \tfrac{x}{2} + \tfrac{p_R}{2p^+}\,(t-\sigma) + i\sqrt{\tfrac{\alpha'}{2}} \textstyle\sum_{n\in\mathbf{Z}-\{\mathbf{0}\}} \tfrac{\tilde{\alpha}_n^i}{n}\, e^{2\pi i\, n\,(\sigma-t)/\ell} \tag{8.8}$$

with

$$p_L = \frac{k}{R} + \frac{wR}{\alpha'} \quad ; \quad p_R = \frac{k}{R} - \frac{wR}{\alpha'} \tag{8.9}$$

In order to be compatible with making all right handed dynamics trivial, namely $X_R \equiv 0$, we need

$$x = 0 \quad ; \quad \tilde{\alpha}_n = 0 \quad ; \quad k = w \quad ; \quad R = \sqrt{\alpha'} \tag{8.10}$$

So the center of mass position degree of freedom is removed, momentum is related to winding, and the internal torus is frozen at fixed radius $\sqrt{\alpha'}$.

Generalizing to 16 dimensions, we propose the following expansion for the left moving fields $X^I(\sigma+t)$

$$X_L^I(\sigma+t) \;=\; \tfrac{P^I}{2p^+}\,(t+\sigma) \,+\, i\sqrt{\tfrac{\alpha'}{2}}\,\textstyle\sum_{n\in\mathbf{Z}-\{\mathbf{0}\}} \tfrac{\alpha_n^i}{n}\, e^{-2\pi i\, n\,(\sigma+t)/\ell} \qquad (8.11)$$

where P^I is a 16d vector in a lattice Λ of internal quantized momenta. The whole right moving sector can be though of as consistenty truncated from the theory (to check complete consistency would require to verify that right handed dynamical modes are not excited in interactions, either; we skip this more involved issue).

The total left moving hamiltionian is

$$\begin{aligned} H_L &= \frac{\sum_i p_i^2}{4p^+} + \frac{\sum_I P^I P^I}{4p^+} + \frac{1}{\alpha' p^+}\,(N-1) \\ N &= \sum_i \sum_N \alpha_{-n}^i \alpha_n^i \sum_I \sum_N \alpha_{-n}^I \alpha_n^I \end{aligned} \qquad (8.12)$$

We have the spacetime mass formulae

$$\begin{aligned} \alpha' m_R^2/2 &= \tilde{N}_B + \tilde{N}_F - 2\nu(1-\nu) \\ \alpha' m_L^2/2 &= N_B + \frac{P^2}{2} - 1 \end{aligned} \qquad (8.13)$$

and the level matching conditions are given by

$$m_L^2 = m_R^2 \qquad (8.14)$$

8.1.3 Modular invariance and lattices

Let us describe a modular invariant partition function and then discuss what kind of physical spectrum it is describing. We can assume the simple ansatz that the complete partition function factorizes as a product of a left and a right moving piece, namely

$$Z(\tau) \;=\; (4\pi\alpha'\tau_2)^{-4}\,|\eta(\tau)|^{-16}\,\overline{Z}_\psi(\tau)\,Z_{\alpha^I}(\tau) \qquad (8.15)$$

The first factor corresponds to tracing over the 10d spacetime momentum degrees of freedom, the second to the trace over the oscillators of the X_R^i,

X_L^i. The factor $\overline{Z}_\psi(\tau)$ is the trace over the right moving fermionic oscillators. From our experience with type II superstrings, an almost modular invariant partition function for this sector is

$$\overline{Z}_\psi \,=\, (\eta^{-4})^* \left(\vartheta \begin{bmatrix} 0 \\ 0 \end{bmatrix}^4 - \vartheta \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}^4 - \vartheta \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}^4 \pm \vartheta \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}^4 \right)^* \tag{8.16}$$

The two choices for the sign eventually lead to the same theory (up to a 10d parity transformation), so for concreteness we pick the $-$ sign.

For $Z_{\alpha^I}(\tau)$ we have the trace over the oscillators and the 16d momentum degrees of freedom

$$Z_{\alpha^I}(\tau) \,=\, \eta(\tau)^{-16} \sum_{P\in\Lambda} q^{P^2/2} \tag{8.17}$$

Now we need to require modular invariance, and this will impose some restrictions on the possible choices of Λ.

i) As $\tau \to \tau+1$, the momentum and bosonic oscillator part is invariant, while we have

$$\vartheta \begin{bmatrix} 0 \\ 0 \end{bmatrix}(\tau+1) = \vartheta \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}(\tau) \quad ; \quad \vartheta \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}(\tau+1) = \vartheta \begin{bmatrix} 0 \\ 0 \end{bmatrix}(\tau)$$
$$\vartheta \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}(\tau+1) = e^{-\pi i/4}\,\vartheta \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}(\tau) \quad ; \quad \vartheta \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}(\tau+1) = e^{-\pi i/4}\,\vartheta \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}(\tau)$$
$$\eta(\tau+1) = e^{\pi i/12}\,\eta(\tau)$$

and hence

$$\overline{Z}_\psi(\tau+1) \,=\, e^{4\pi i/3}\,\overline{Z}_\psi \tag{8.18}$$

Hence we need

$$Z_{\alpha^I}(\tau+1) \,=\, e^{2\pi i/3}\,Z_{\alpha^I} \tag{8.19}$$

This is so, provided

$$\sum_{P\in\Lambda} e^{2\pi i(\tau+1)\,P^2/2} \,=\, \sum_{P\in\Lambda} e^{2\pi i\tau\,P^2/2} \tag{8.20}$$

Namely, we need $P^2 \in 2\mathbf{Z}$ for any $P \in \Lambda$. Lattices with this property are called *even*.

For future use (see next footnote), let us point out that even lattice are always *integer* lattices. An integer lattice is such that for any $v, w \in \Lambda$, we have $v \cdot w \in \Lambda$. To show this, notice that in an even lattice, for any v, w we have $(v+w)^2$ is even, but $(v+w)^2 = v^2 + w^2 + 2v \cdot w$. Since v^2, w^2 are enen, it follows that $v \cdot w \in \mathbf{Z}$ and Λ is integer.

ii) As $\tau \to -1/\tau$, the spacetime momentum times spactime bosonic oscillator piece is invariant. For the fermionic piece we have

$$
\begin{aligned}
&\vartheta\left[\begin{array}{c}0\\0\end{array}\right](-1/\tau) = (-i\tau)^{1/2}\,\vartheta\left[\begin{array}{c}0\\0\end{array}\right](\tau) \qquad \vartheta\left[\begin{array}{c}0\\1/2\end{array}\right](-1/\tau) = (-i\tau)^{1/2}\,\vartheta\left[\begin{array}{c}1/2\\0\end{array}\right](\tau)\\
&\vartheta\left[\begin{array}{c}1/2\\0\end{array}\right](-1/\tau) = (-i\tau)^{1/2}\,\vartheta\left[\begin{array}{c}0\\1/2\end{array}\right](\tau) \quad \vartheta\left[\begin{array}{c}1/2\\1/2\end{array}\right](-1/\tau) = i(-i\tau)^{1/2}\,\vartheta\left[\begin{array}{c}1/2\\1/2\end{array}\right](\tau)\\
&\eta(-1/\tau) = (-i\tau)^{1/2}\,\eta(\tau)
\end{aligned}
$$

and hence

$$\overline{Z}_\psi(-1/\tau) \;=\; \overline{Z}_\psi(\tau) \tag{8.21}$$

So we need

$$Z_{\alpha^I}(-1/\tau) \;=\; Z_{\alpha^I}(\tau) \tag{8.22}$$

The left hand side reads

$$Z_{\alpha^I}(-1/\tau) \;=\; (-i\tau)^{-8}\,\eta(\tau)^{-16}\sum_{P\in\Lambda} e^{2\pi i(-1/\tau)P^2/2} \tag{8.23}$$

Using the Poisson resummation formula [2]

$$
\begin{aligned}
&\textstyle\sum_{v\in\Lambda}\exp\left[\,-\pi(v+\theta)\cdot A\cdot(v+\theta) + 2\pi i\,(v+\theta)\cdot\phi\,\right] =\\
&= \tfrac{1}{|\Lambda^*/\Lambda|\sqrt{\det A}}\textstyle\sum_{k\in\Lambda^*}\exp\left[\,-\pi(k+\phi)\cdot A^{-1}\cdot(k+\phi) - 2\pi i k\theta\,\right]
\end{aligned} \tag{8.24}
$$

we have

$$Z_{\alpha^I}(-1/\tau) \;=\; (-i\tau)^{-8}\,\eta(\tau)^{16}\,\frac{1}{|\Lambda^*/\Lambda|}(-i\tau)^8\sum_{K\in\Lambda^*} e^{-2\pi i\tau\,K^2/2} \tag{8.25}$$

[2] Here Λ^* is the lattice dual to Λ, which is formed by the vectors k such that $k \cdot v \in \mathbf{Z}$ for any $v \in \Lambda$. For integer lattices, Λ is a sublattice of Λ^*, and the quotient Λ^*/Λ is a finite set. Its cardinal $|\Lambda^*/\Lambda|$ is called the index of Λ in Λ^*.

So we have invariance if $\Lambda* = \Lambda$. Such lattices are called *self-dual*.

The compactification lattice Λ must be even and self-dual to obtain a consistent modular invariant theory. Even self-dual lattices (with euclidean signature scalar product) have been proved by mathematicians to be extremely constrained. They only exist in dimensions multiple of eight; happily we need 16d lattices, so the dimension is in the allowed set of values.

Moreover there are only *two* inequivalent 16d even and self-dual lattices. These are the following

i) The $E_8 \times E_8$ lattice

It is spanned by vectors of the form

$$\begin{array}{ccc} (n_1, \ldots, n_8; n'_1, \ldots, n'_8) & ; & (n_1 + \frac{1}{2}, \ldots, n_8 + \frac{1}{2}; n'_1, \ldots, n'_8) \\ (n_1, \ldots, n_8; n'_1 + \frac{1}{2}, \ldots, n'_8 + \frac{1}{2}) & ; & (n_1 + \frac{1}{2}, \ldots, n_8 + \frac{1}{2}; n'_1 + \frac{1}{2}, \ldots, n'_8 + \frac{1}{2}) \end{array} \quad (8.26)$$

with $n_I, n'_I \in \mathbf{Z}$, and $\sum_I n_I = \text{even}$, $\sum_{I'} n'_I = \text{even}$

ii) The Spin(32)/$\mathbf{Z_2}$ lattice

Spanned by vectors of the form

$$\begin{array}{c} (n_1, \ldots, n_{16}) \\ (n_1 + \frac{1}{2}, \ldots, n_{16} + \frac{1}{2}) \end{array} \quad (8.27)$$

So these define two consistent heterotic superstring theories.

8.1.4 Spectrum

The spectrum of these theories is found by constructing left and right moving states in the usual way (constructing ladder operators and Hilbert spaces, and applying the GSO projections dictated by the partition function), and glueing them together satisfying level-matching.

We will simply discuss massless states, although the rules to build the whole tower of string states should be clear.

The right moving sector is exactly the same as one of the sides of the type II superstrings. The two choices of Z_ψ give two final theories which differ by a 10d parity operation, so are equivalent; hence we choose one of them. The massless states suriviving the GSO projection are

Sector	State	$SO(8)$
NS	$\tilde{\psi}^i_{-1/2}\|0\rangle$	8_V
R	$\tilde{A}^+_a\|0\rangle$	8_C
	$\tilde{A}^+_{a_1}\tilde{A}^+_{a_2}\tilde{A}^+_{a_3}\|0\rangle$	

We will denote the states in the R sector by $\frac{1}{2}(\pm,\pm,\pm,\pm)$ (with odd number of $-$'s), i.e. by the $SO(8)$ weights.

For the left movers, the mass formula is given by

$$\alpha' m_L^2/2 \,=\, N_B \,+\, \frac{P^2}{2} \,-\, 1 \tag{8.28}$$

Lightest states are

	State	$\alpha' m_L^2/2$	$SO(8)$
$N_B = 0, P = 0$	$\|0\rangle$	-1	1
$N_B = 1, P = 0$	$\alpha^i_{-1}\|0\rangle$	0	8_V
$N_B = 1, P = 0$	$\alpha^I_{-1}\|0\rangle$	0	1
$N_B = 0, P^2 = 2$	$\|P\rangle$	0	1

Notice that there is a tachyon, but it will not lead to any physical state in spacetime since it has no tachyonic right-moving state to be level-matched with.

The latter states with $P^2 = 2$ are different for the two choices of lattice. For the $E_8 \times E_8$ lattice, these states have internal momentum P of the form

$$\begin{array}{ll} (\underline{\pm,\pm,0,0,0,0,0,0};0,0,0,0,0,0,0,0) & \\ \frac{1}{2}(\pm,\pm,\pm,\pm,\pm,\pm,\pm,\pm;0,0,0,0,0,0,0,0) & \#- = \text{even} \\ (0,0,0,0,0,0,0,0;\underline{\pm,\pm,0,0,0,0,0,0}) & \\ \frac{1}{2}(0,0,0,0,0,0,0,0;\pm,\pm,\pm,\pm,\pm,\pm,\pm,\pm) & \#- = \text{even} \end{array} \tag{8.29}$$

We note that these are the non-zero root vectors of $E_8 \times E_8$ (hence the name of the lattice).

States with $P^2 = 2$ in the $Spin(32)/\mathbf{Z_2}$ lattice have P of the form

$$(\underline{\pm,\pm,0,0,0,0,0,0,0,0,0,0,0,0,0,0}) \tag{8.30}$$

We note that these are the non-zero root vectors of $SO(32)$ (hence the name of the lattice). Notice that momenta of the form $P = 1/2(\pm,\ldots,\pm)$ have $P^2 = 4$ and give rise to massive states.

We should now glue together left and right states. The schematic structure of massless states is

$$(\tilde{8}_V + \tilde{8}_C) \times (8_V + \alpha^I + |P\rangle) \tag{8.31}$$

Namely, we have the states

$$\begin{array}{cc} \tilde{\psi}^i_{-1/2}|0\rangle \times \alpha^j_{-1}|0\rangle & 8_V \times 8_V = 1 + 28_V + 35_V \\ \frac{1}{2}(\pm,\pm,\pm,\pm) \times \alpha^j_{-1}|0\rangle & 8_C \times 8_V = 8_S + 56_S \end{array}$$

The massless fields are a scalar dilaton ϕ, a graviton $G_{\mu\nu}$, a 2-form B_2, and fermion superpartners, including a 10d chiral gravitino (56_S). This is the $N = 1$ 10d supergravity multiplet, so the theory turns out to have $N = 1$ spacetime susy. Notice that this is half the susy of type II theories, since we have GSO projection only on one of the sides, and this produces half as many gravitinos.

We also obtain the states

$$\begin{array}{cc} \psi^i_{-1/2}|0\rangle \times \alpha^I_{-1}|0\rangle & 8_v \\ \frac{1}{2}(\pm,\pm,\pm,\pm) \times \alpha^I_{-1}|0\rangle & 8_C \end{array}$$

they correspond to 16 gauge bosons and superpartner gauginos. The gauge group is $U(1)^{16}$.

Finally we have the states

$$\begin{array}{cc} \tilde{\psi}^i_{-1/2}|0\rangle \times |P\rangle & 8_v \\ \frac{1}{2}(\pm,\pm,\pm,\pm) \times |P\rangle & 8_C \end{array}$$

These are also gauge bosons and gauginos. It is possible to see that they are charged under the $U(1)^{16}$ gauge symmetries (this is analogous to how winding and momentum states are charged with respect to the gauge symmetries obtained in toroidal compactifications), so the gauge group will be enhanced to a non-abelian symmetry. We would like to identify what is the final gauge group, for each of the two choices of internal lattice. The $U(1)^{16}$ gives the Cartan subalgebra of the group, which hence has rank 16. The charge of a state $|P\rangle$ under the I^{th} $U(1)$ factor is given by P^I, hence the vectors P must correspond to the non-zero roots of the gauge group. As we have mentioned before, the $P^2 = 2$ states of the compactification lattices precisely correspond to the non-zero roots of the groups $E_8 \times E_8$ and $SO(32)$, respectively for each of the lattices. Hence states from the α^I oscillators and from momentum P give altogether 10d $N = 1$ vector multiplets of $E_8 \times E_8$ or $SO(32)$.

The complete massless spectrum for the two consistent (spacetime supersymmetric) heterotic theories is 10d $N = 1$ supergravity coupled to $E_8 \times E_8$ or $SO(32)$ vector multiplets. These theories are chiral, so there is a very stringent consistency issue arising from 10d anomalies. This will be reviewed later on in this lecture.

Notice that the spectrum of these theories is very exciting. It contains non-abelian gauge symmetries and charged chiral fermions. In later lectures we will see that this structure allows to obtain interesting theories with charged chiral 4d fermions upon compactification. In particular this is possible due to the existence of fundamental vector multiplets in the higher dimensional theory, therefore avoiding diverse no-go theorems about getting charged chiral fermions in Kaluza-Klein theories with pure (super)gravity in the higher dimensional theory.

8.2 Heterotic strings in the fermionic formulation

In this section we discuss a different construction of the same heterotic string theories as before. Readers comfortable with the above bosonic formulation may therefore skip this section.

We refer the reader to the last section in the lesson about type II superstring to the discussion of bosonization/fermionization. There we discussed that a theory of k left-moving boson paramerizing compactified directions is equivalent to a theory $2k$ fermions with a sum over boundary conditions determined by the compactification lattice.

This motivates introducing a different description of the heterotic strings we have constructed. Indeed, we construct a string theory whose worldsheet degrees of freedom (already in the light-cone gauge) are right moving fields $X_R^i(\sigma - t)$, $\psi_R^i(\sigma - t)$, $i = 2, \ldots, 9$ and left-moving fields $X_L^i(\sigma + t)$, $\lambda_L^A(\sigma + t)$, with $i = 2, \ldots, 9$ and $A = 1, \ldots, 32$.

The quantization of these is standard: Bosons $X_{L,R}^i$ are periodic in σ and give rise to integer-modded oscillators, fermions ψ^i can be NS or R and have consequently half-integer or integer modded oscillators. Finally fermions λ^A

can also be NS or R, but in contrast with the preivous λ^A's with different boundary condition can coexist in the same sector (recall the ψ's must be all NS or all R in order not to violate spacetime Lorentz invariance).

With these ingredients, we can construct two possible modular invariant partition functions, which have the familiar GSO projection on the right-moving piece. They define two consistent heterotic string theories, which will turn out to be the two heterotic strings constructed above, but described in 2d fermionic language.

The two partition functions have the structure

$$Z(\tau) \,=\, (4\pi\alpha'\tau_2)^{-4}\,|\eta(\tau)|^{-16}\,\overline{Z}_\psi(\tau)\,Z_\lambda(\tau) \tag{8.32}$$

with two possible opctions for Z_λ

$$\begin{aligned}
\textbf{i)}\quad Z_\lambda(\tau) \,&=\, \frac{\vartheta\begin{bmatrix}0\\0\end{bmatrix}^{16}+\vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^{16}+\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^{16}+\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^{16}}{\eta(\tau)^{16}}\\
\textbf{ii)}\quad Z_\lambda(\tau) \,&=\, \left(\frac{\vartheta\begin{bmatrix}0\\0\end{bmatrix}^{8}+\vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^{8}+\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^{8}+\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^{8}}{\eta(\tau)^{8}}\right)^2
\end{aligned} \tag{8.33}$$

They differ in the way the 32 fermions λ^A are grouped. It is possible to use the expressions of the ϑ functions as infinite sums and write the above partition functions as sums over momenta in the $Spin(32)/\mathbf{Z_2}$ and $E_8\times E_8$ lattices, thus showing the equivalence with the bosonic formulations above. We have recovered exactly the same heterotic string theories starting from a different world-sheet formulation (related to the previous by bosonization/fermionization).

It is however interesting to construct the spectrum directly in the fermionic formulation. We review it now, with special emphasis on the massless sector.

The right-moving sector is very familiar, and works exactly as one of the sides of the type II superstring. At the massless level, we obtain NS states $\tilde{\psi}^i_{-1/2}|0\rangle$ in the 8_V of $SO(8)$ and R states in the 8_C.

For the left-moving sector, we treat the two possible cases separately.

The $SO(32)$ heterotic in fermionic language

We start with **i)**, the partition function Z_λ has the structure

$$\mathrm{tr}_{\,\mathcal{H}_{NS}}(1+(-)^F)+\mathrm{tr}_{\,\mathcal{H}_R}(1+(-)^F) \tag{8.34}$$

Hence the 32 fermions are all with NS or all with R boundary conditions. In each sector there is an overall GSO projection.

NS sector

The mass formula is given by

$$\alpha' m_L^2/2 \;=\; N_B + N_F - 1 \tag{8.35}$$

There are no fermion zero modes, so the vacuum is non-degenerate; the Hilbert space is obtained by applying negative modding oscillators on it. The GSO projection requires the number of fermion oscillators to be even for physical states. The lightest states are

State	$\alpha' m_L^2/2$
$\|0\rangle$	-1
$\alpha_{-1}^i\|0\rangle$	0
$\lambda_{-1/2}^A\lambda_{-1/2}^B\|0\rangle$	0

The latter states correspond to antisymmetric combinations of the indices A and B. Therefore and for future convenience we associate them to the generators of an $SO(32)$ Lie algebra (whose generators in the vector representation are given by antisymmetric matrices).

As before, the left-moving tachyon cannot be level-matched with any right-moving state and does not lead to spacetime tachyon states.

R sector The mass formula is given by

$$\alpha' m_L^2/2 \;=\; N_B + N_F + 1 \tag{8.36}$$

There are 32 fermion zero modes, so the vacuum is 2^{16}-fold degenerate, split in two chiral spinor irreps of the underlying $SO(32)$ symmetry (acting on the Λ^A). The GSO projection selects states with even number of fermion oscillators on one of them, and states with odd number of fermion oscillators on the other. All states in the R sector are however massive, hence we will not be too interested in them.

The total spectrum is found by glueing left and right moving states in a level-matched way. The states

$$8_V \times \alpha_{-1}^i|0\rangle \quad ; \quad 8_C \times \alpha_{-1}^i|0\rangle \quad ; \tag{8.37}$$

reproduce the 10d $N=1$ supergravity multiplet $1+28_V+35_V+8_S+56_S$. The states

$$8_V \times \lambda_{-1/2}^A\lambda_{-1/2}^B|0\rangle \quad ; \quad 8_C \times \lambda_{-1/2}^A\lambda_{-1/2}^B|0\rangle \quad ; \tag{8.38}$$

reproduce 10d $N = 1$ vector multiplets with gauge group $SO(32)$ (as can be guessed by noticing that we have $32 \times 32/2$ states associated with antisymmetric combinations of indices in the vector of $SO(32)$.

Hence we have reproduced the (massless) spectrum of the $SO(32)$ heterotic superstring.

The $E_8 \times E_8$ heterotic in fermionic language

We now study **ii)**, the partition function Z_λ has the structure

$$\left[\operatorname{tr}_{\mathcal{H}_{NS}} + \operatorname{tr}_{\mathcal{H}_R}(1 + (-)^F)\right]^2 \tag{8.39}$$

Hence the 32 fermions are split in two sets of 16, which we denote λ^A, $\lambda^{A'}$. They have equal boundary conditions within each set, but with independent boundary conditions. For each set of 16 fermions: the NS boundary conditions imply the groundstate is unique, and GSO requires an even number of fermion oscillators to be applies; the R boundary conditions imply a 2^8-fold degenerate groundstate, split as two chiral spinor irreps of the underlying $SO(16)$, denoted 128 and 128', with GSO requiring even number of fermion oscillators acting on 128 and odd number on 128'.

With this information we can construct the complete left-moving spectrum. The lightest states which will finally level-match with right-moving ones are the massless ones, so we look only at these

$\mathbf{NS_{16}NS_{16}}$

The mass formula is

$$\alpha' m_L^2/2 \;=\; N_B + N_F - 1 \tag{8.40}$$

The massless states are

State	Remark
$\alpha^i_{-1}\|0\rangle$	8_V of $SO(8)$
$\lambda^A_{-1/2}\lambda^B_{-1/2}\|0\rangle$	Adj. of $SO(16)$
$\lambda^{A'}_{-1/2}\lambda^{B'}_{-1/2}\|0\rangle$	Adj. of $SO(16)'$

$\mathbf{R_{16}NS_{16}}$

The vacuum is 2^7-fold degenerate due to the 16 R fermion zero modes. The mass formula is

$$\alpha' m_L^2/2 \;=\; N_B + N_F \tag{8.41}$$

The massless states are the groundstates, which transform as 128 of the $SO(16)$

$\mathbf{NS}_{16}\mathbf{R}_{16}$

Similarly to the above, the massless states are the groundstates, which transform as 128 of the $SO(16)'$

$\mathbf{R}_{16}\mathbf{R}_{16}$

In this sector even the groundstate is massive.

The total massless spectrum is obtained by tensoring the right-moving $8_V + 8_C$ with the above left handed states. It is easyy to recover the 10d $N = 1$ supergravity multiplet by tensoring the right-moving $8_V + 8_C$ with the left-moving 8_V. On the other hand, by tensoring the right-moving $8_V + 8_C$ with the left-moving $SO(8)$ singlets we obtain 10d $N = 1$ vector multiplets with gauge group $E_8 \times E_8'$. The gauge group can be guessed by rememebering that the adjoint of E_8 decomposes as an adjoint plus a 128 of $SO(16)$. Hence we recover the complete massless spectrum of the $E_8 \times E_8$ heterotic.

8.3 Spacetime Non-susy heterotic string theories

There are other ways to construct modular invariant partition functions, beyond the factorized proposal used above. These are more easily constructed using the fermionic formulation of superstrings (a bosonized formulation is also possible, but more involved since it would require lattices mixing the internal bosons and spacetime fermionic degrees of freedom).

Without aiming at a general classification, let us simply give one example of such a modular invariant partition function

$$\frac{1}{\overline{\eta}^4\,\eta^{16}} \quad \left(\overline{\vartheta}\begin{bmatrix} 0 \\ 0 \end{bmatrix}^4 \vartheta\begin{bmatrix} 0 \\ 0 \end{bmatrix}^{16} - \overline{\vartheta}\begin{bmatrix} 1/2 \\ 0 \end{bmatrix}^4 \vartheta\begin{bmatrix} 1/2 \\ 0 \end{bmatrix}^{16} - \right.$$
$$\left. \overline{\vartheta}\begin{bmatrix} 0 \\ 1/2 \end{bmatrix}^4 \vartheta\begin{bmatrix} 0 \\ 1/2 \end{bmatrix}^{16} - \overline{\vartheta}\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}^4 \vartheta\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}^{16} \right) \tag{8.42}$$

The interpretation in terms of the GSO projection is that we correlate the $(-1)^F$ quantum number of the right moving fermions with the $(-1)^F$ quantum number of the internal left-moving fermions.

Schematically the spectrum at the massless level is

Sector	State	$\alpha' m^2$	$SO(8)$	internal
NS-NS	$\|0\rangle \otimes \psi^I_{-1/2}\|0\rangle$	-2	1	32
	$\tilde{\psi}^i_{-1/2}\|0\rangle \otimes \alpha^i_{-1}\|0\rangle$	0	$1 + 28_V + 35_V$	1
	$\tilde{\psi}^i_{-1/2}\|0\rangle \otimes \psi^I_{-1/2}\psi^J_{-1/2}\|0\rangle$	0	8_V	$SO(32)$

Notice that the left moving R states has only massive modes, so by level matching the NS-R, R-NS and R-R sector have only massive modes. The theory contains the graviton, 2-form and dilaton field, as well as $SO(32)$ gauge bosons. The theory is spacetime non-supersymmetric, and contains tachyons, transforming in the 32 of $SO(32)$. As in other cases of tachyons in closed string theories, the fate of this instability is not known. Finally, the theory contains fermions, but all of them are massive. Overall, the theory is not too interesting, and is given just as an example of non-supersymmetric heterotic strings.

This heterotic string can also be constructed in the bosonic formulation, by reading off the required lattice from the above partition function. Note as we said that the lattice would involve the internal bosons as well as the bosonization of the right moving fermions.

We conclude by pointing out that all 10d non-supersymmetric heterotic theories contains tachyons, except for the so-called $SO(16) \times SO(16)$ heterotic. Details on this can be found in [71] (although discussed in a language perhaps not too transparent).

8.4 A few words on anomalies

Anomaly cancellation in theories with chiral 10d spectrum is an astonishing example of self-consistency of string theory. Therefore it is an interesting topic to be covered. We leave its discussion for the evaluation project.

8.4.1 What is an anomaly?

Let us start giving a set of basic facts about anomalies, directed towards understanding in what situations they may appear. A complete but formal introduction may be found in [50].

When a classical theory has a symmetry which is not present in the quantum theory, we say that the symmetry has an anomaly or that the theory is anomalous. Namely, what happens is that quantum corrections generate terms in the effective action which are not invariant under the symmetry. Since the classical lagrangian was invariant, such terms cannot be removed with local counterterms, and the quantum theory is not invariant.

In the path integral formalism of quantum field theory, the lack of invariance of the quantum theory (the anomaly) arises from the non-invariance of the measure of the functional integration (this is Fujikawa's method of computing anomalies).

Notice that if there exists some regularization which preserves a classical symmetry of the classical theory, then the symmetry is not anomalous. Namely, the regularized theory is still invariant under the symmetry, so regularized quantum corrections preserve the symmetry, and when the cutoff is taken to infinity the symmmetry is still preserved. Hence the only symmetries which can be anomalous are those for which no symmetry-preserving regularization exists.

This has the important consequence that only chiral fields can contribute to anomalies. The contribution from non-chiral fields can always be regularized by using the Pauli-Villars regularization, which preserves all the symmetries of the system.

This implies that anomalies can arise only in even dimensions [3] $D = 2n$ because only then there exist chiral representations of the Lorentz group. Anomalies arise from very precise diagrams, they appear only from contributions at one loop (and not at higher order, this is Adler's theorem), in a diagram of one loop of chiral fields (usually fermions) with $n+1$ external legs of the fields associated to the symmetry (gauge bosons for gauge symmetries, gravitons for diffeomorphism invariance (gravitaional anomalies), and external currents for global symmetries). For instance, in 10d theories, anomalies arise from hexagon diagrams (see fig 8.1 with external legs corresponding to gravitons and/or gauge bosons, if they are present in the theory.

We will center on gauge anomalies, which are lethal for the theories. Namely, in preserving unitarity of the theory it is essential that unphysical polarization modes decouple, and this happens as a consequence of gauge invariance. If an anomaly spoils the gauge invariance in the quantum theory, the latter is inconsistent (non-unitary, etc). Namely, by scattering physical polarization modes we can create unphysical ones by processes mediated by

[3] There exists different class of anomalies, called global anomalies (what I mean here is different from anomalies for global symmetries), which are different from the ones we study here and may also exist in odd dimensions; for instance the parity anomaly in odd dimensions.

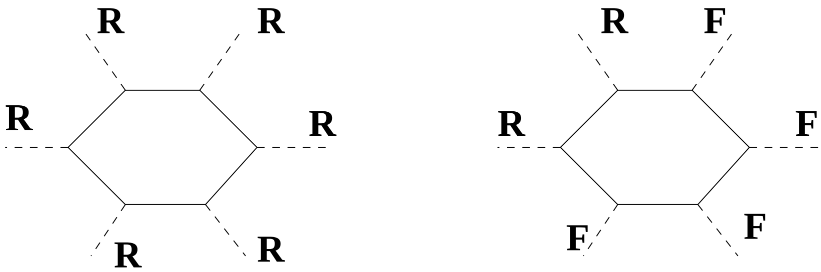


Figure 8.1: Different hexagon diagrams contributing to gravitaional, gauge and mixed anomalies.

the anomaly diagram. Hence, the latter must vanish in order to have a consistent unitary theory.

Anomalous gauge variations of the effective action can be obtained from the so-called anomaly polynomial I which is a formal $(2n+2)$-form constructed as a polylnomial in the gravitational and gauge curvature 2-forms, R and F, resp. It is therefore closed and gauge invariant. For instance, in a 10d theory with gravitons and gauge bosons, the anomaly polynomial is a linear combination of things like $\mathrm{tr}\, R^6$, $\mathrm{tr}\, R^4 \mathrm{tr}\, R^2$, $\mathrm{tr}\, F^6$, $\mathrm{tr}\, F^4 \mathrm{tr}\, F^2$, $(\mathrm{tr}\, F^2)^3$, etc, with wedge products implied. Coefficients of the anomaly polynomial are determined by the spectrum of chiral fields of the theory. The anomalous variantion of the 1-loop effective action under a symmetry transformation with gauge parameter λ is of the form

$$\delta_\lambda S_{\mathrm{eff}} = \int I^{(1)} \tag{8.43}$$

where $I^{(1)}$ is an n-form, obtained by the so-called Wess-Zumino descent procedure, as follows. Since the anomaly polynomial I is closed, it is locally exact and can be written as $I = dI^{(0)}$, with $I^{(0)}$ a $(2n+1)$-form. It can be shown that the gauge variation of $I^{(0)}$ under any symmetry transformation is closed, hence it is also locally exact and we can write $\delta_\lambda I^{(0)} = dI^{(1)}$, where λ is the gauge parameter and $I^{(1)}$ is the above n-form. Hence we have

$$I = dI^{(0)} \quad ; \quad \delta_\lambda I^{(0)} = dI^{(1)} \tag{8.44}$$

To give one simple example, consider a 4d $U(1)$ gauge theory with n chiral fermions carrying charge $+1$. The anomaly polynomial is given by $I = n\, F^3$.

We then have $I^{(0)} = n\, AF^2$ and $\delta_\lambda I^{(0)} = n\, d\lambda\, F^2$, hence $I^{(1)} = n\lambda F^2$, leading to the familiar form of the 4d anomaly.

Notice that the fact that the anomaly is a topological quantity is related to the fact that it is determined by the spectrum of chiral fermions. The latter is unchanged by continuous changes of the parameters of the theory, like coupling constants, etc, hence so is the anomaly, i.e. it is a topological quantity.

The fact that all anomalies in a theory can be derived from a unique anomaly polynomial implies that the anomalies for diverse symmetries (and for diagrams involving different kinds of gauge fields) obey the so-called Wess-Zumino consistency conditions. Roughly speaking, they imply that if a gauge variation wrt a symmetry 'a' generates a term involving the gauge curvature of a symmetry 'b', then a gauge variation of 'b' should generate terms involving the curvature of 'a'. This is clear from the fact that the diagram mediating the anomalies contains external legs of both 'a' and 'b'.

8.4.2 Anomalies in string theory and Green-Schwarz mechanism

In string theory, the spacetime theory is often chiral, for instance type IIB or heterotic superstrings in 10d (also type I, see next lectures).

From the string theory viewpoint, the theory is however finite and gauge invariant. This implies that the underlying string theory is providing a regularization of the corresponding effective field theory containing the chiral fields. From this viewpoint it is clear that string theory should lead to theories free of gauge and gravitational anomalies (In fact, the relation between modular invariance (ultimately responsible for finiteness of string theory) and absence of anomalies has been explored in the literature [51]).

In type IIB theory, the fields contributing to the gravitational anomalies are the 8_S, 56_S and 35_C, i.e. the fermions and the self-dual 4-form. With this matter content there is a miraculous cancellation of all terms in the anomaly polynomial, which then automatically vanishes. The theory is therefore non-anomalous [52].

In heterotic theories, the field content also leads to some miraculous cancellations of terms in the anomaly polynomial. For instance, the fact that the gauge group has 496 generators leads to the absence of $\mathrm{tr}\, F^6$ terms. This is called cancellation of the irreducible anomaly. However, even after these

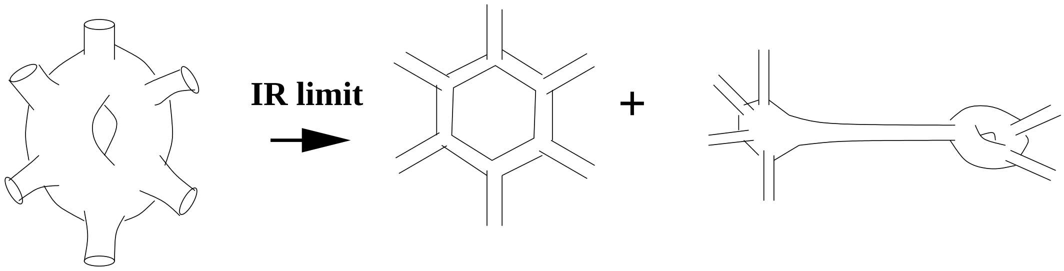


Figure 8.2: The low-energy limit of the six-point function for gravitons and gauge bosons contains two contributions, the familiar field theory hexagon, and a diagram of exchange of closed string modes at tree level with tree and one-loop level couplings to external legs.

miracles, the anomaly polynomaly still is non-vanishing, but has a special structure, it is of the form

$$I \simeq \mathrm{tr}\, F^4\, (\mathrm{tr}\, F^2 - \mathrm{tr}\, R^2) \tag{8.45}$$

This residual anomaly, known as reducible anomaly, is cancelled by a special contribution to the six-point function of gauge bosons and gravitons, which does not have the standard field theory hexagon interpretation. As is shown in figure 8.2, the contribution to the 1-loop amplitude with six external legs lead to two kinds of low-energy contributions. One of them is the familiar field thery hexagon diagram, of massless particles running in a loop. The second is however of the form of an exchange of massless modes along a tree level diagram, and a subsequent 1-loop coupling to some gauge fields.

The existence of the second contribution was noticed by Green and Schwarz [4], who provided the right field theory interpretation for it. The massless mode propagating along the tube is the 2-form B_2 (or its dual B_6) which has couplings to the curvatures as follows

$$\int_{10d} B_2 \wedge \mathrm{tr}\, F^4 \quad ; \quad \int_{10d} B_6 \wedge (\mathrm{tr}\, F^2 - \mathrm{tr}\, R^2) \quad ; \tag{8.46}$$

which arise at tree level and 1-loop respectively. The last coupling is often expressed by saying that B_2 obeys the modified Bianchi identity

$$dH_3 = \mathrm{tr}\, F^2 - \mathrm{tr}\, R^2 \tag{8.47}$$

[4] In fact, they noticed it in type I, which is similar to the $SO(32)$ heterotic at the field theory level.

Using these couplings, the gauge variation of the effective action is

$$\delta \int_{10d} H_3 \wedge (\mathrm{tr}\, F^4)^{(0)} \;=\; \int_{10d} H_3 \wedge \delta(\mathrm{tr}\, F^4)^{(0)} \;=\; \int_{10d} H_3 \wedge d(\mathrm{tr}\, F^4)^{(1)} \;= \tag{8.48}$$
$$\int_{10d} dH_3 \wedge (\mathrm{tr}\, F^4)^{(1)} \;=\; \int_{10d} (\mathrm{tr}\, F^2 - \mathrm{tr}\, R^2)(\mathrm{tr}\, F^4)^{(1)} \simeq \int_{10d} [\,(\mathrm{tr}\, F^2 - \mathrm{tr}\, R^2)(\mathrm{tr}\, F^4)\,]^{(1)}$$

The total anomalous variation therefore vanishes. This is the so-called Green-Schwarz mechanism. This is very remarkable, indeed so remarkable that triggered a lot of interest in string theory since the mid 80's.

Chapter 9

Open strings

In this lecture we discuss open strings. The motivation is clear: they are strings of a kind very different from the ones we have studied up to now, so it is interesting to analyze their main features. Moreover, it is essential to have some familiarity with open strings to construct the type I superstring (see next lecture) since it contains sectors of open strings.

9.1 Generalities

Open strings are string with endpoints; they are described by worldsheets with boundaries, see figure 9.1

The basic interaction between open strings is that two endpoints glue together; the basic interaction vertex corresponds to two open strings joining into a single one, figure **??**a). Notice that the endpoints that glue together may belong to the same open string, so that this basic interaction also implies the existence of a vertex of two open strings joining into a closed one, figure **??**b). This has the remarkable consequence that **theories with open strings necessarily contain closed strings** (notice that we know that there exist theories of closed strings with no open strings; i.e. closed strings may be consistent by themselves, but open string theories necessarily must be coupled to closed string theories).

The worldsheet geometry forces us to include two sectors (open strings and closed strings) in the theory. The total spectrum of spacetime particles is given by the spectrum of oscillation modes of the closed string plus the spectrum of oscillation modes of the open string.

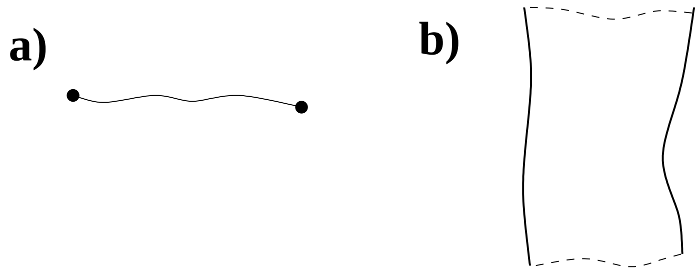


Figure 9.1: Open strings have endpoints. As open strings move in time they sweep out a worldsheet with boundaries.

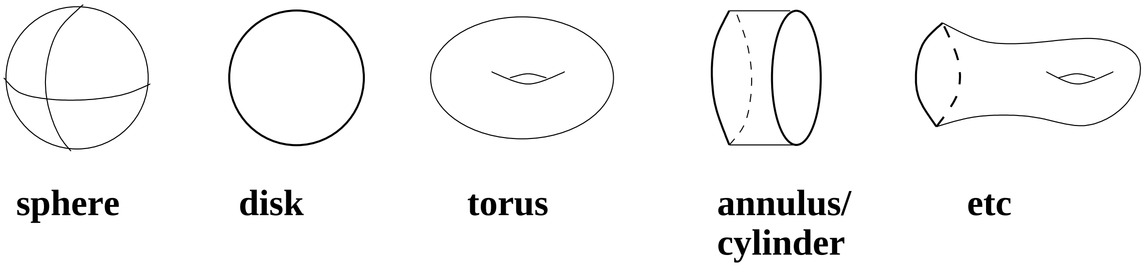


Figure 9.2: .

Any amplitude is obtained by summing over geometries of 2d surfaces interpolating between in and out states. This genus expansion contains contributions from surfaces with handles and boundaries, which is weighted by a factor of $g_s^{-\chi}$ where $\chi = 2 - 2g - n_b$, with g, n_b the number of handles and boundaries. Some examples are given in fig 9.2.

Finally, we would like to make the following important remark. The fact that open strings couple to closed strings implies that the local structure of the worldsheet of open strings is the same as that of closed strings. This implies that the local 2d dynamics for open and closed strings must be the same (with the only differences arising, as we will see, from boundary conditions on the 2d fields).

A related issue is that there exist diagrams which admit two different interpretations, regarded as open string diagrama or closed string diagramas. Namely, the annulus can be regarded as vacuum diagram of open string states running in a loop, or as a tree level diagram of closed string appearing from and disappearing into the vacuum. Both interpretations are possible because the local structure of the worldsheet is the same for open and closed strings. Both interpretations are related by a relabeling of the worldsheet coordinates σ, t. The requirement that a single geometry can receive both interpretations is a strong consisntency condition known as open/closed duality.

9.2 Open bosonic string

For this analysis we follow section 1.3 of [55]. This is an open string whose local worldsheet dynamics is described by 26 2d bosons $X^\mu(\sigma, t)$ and a 2d metric $g_{ab}(\sigma, t)$, with the Polyakov action

$$S_{\rm P} = -\frac{1}{4\pi\alpha'}\int_\Sigma d^2\xi\,(-g)^{1/2}\,g^{ab}(\sigma,t)\,\partial_a X^\mu\,\partial_b X^\nu \eta_{\mu\nu} \tag{9.1}$$

The corresponding closed string sector is therefore the closed bosonic string. Here we center on the quantization of the open string sector, that is quantization of the above 2d field theory living on the interval (with boundary conditions to be specified below).

9.2.1 Light-cone gauge

The gauge freedom of the 2d theory is fixed in the same way as we did for the closed bosonic string. Again we have several steps

1. Reparametrization of t

Fix the t reparametrization freedom by setting the so-called light-cone condition

$$X^+(\sigma, t) = t \tag{9.2}$$

2. Reparametrization of σ

For slices of constant t, define a new spatial coordinate σ' for each point of the slice, as the (diffeomorphism and Weyl) invariant distance to one of the endpoints

$$\sigma' = c(t)\int_{\sigma_0}^{\sigma} f(\sigma, t)\,d\sigma \tag{9.3}$$

where

$$f(\sigma) = (-g)^{-1/2}\,g_{\sigma\sigma}(\sigma, t) \tag{9.4}$$

and $c(t)$ is a σ independent coefficient used to impose that the total length of the string is fixed, a constant in t which we call ℓ. Notice that, in contrast with closed string, there is a preferred reference line (so we do not impose level matching constraints to get physical states). In what follows σ' will be denoted simply σ.

3. Weyl invariance

Now we use Weyl invariance to impose that

$$g = -1 \qquad \forall \sigma, t \tag{9.5}$$

The gauge fixing conditions imply, just like for the closed bosonic string, that

$$\partial_\sigma g_{\sigma\sigma} = 0 \tag{9.6}$$

The quantization is very similar to quantization of the closed bosonic string, and the result is exactly the same local dynamics (e.g. hamiltonian). The reader satisfied with this explanation is welcome to jump to eq. (9.17).

9.2.2 Boundary conditions

It is now convenient to obtain what kind of boundary conditions we need to impose at $\sigma = 0, \ell$. To obtain them let us vary the action (9.1)

$$\begin{aligned}
\delta S_P &= -\frac{1}{2\pi\alpha'} \int_\Sigma d^2\xi \, g^{ab} \, \partial_a X^\mu \, \partial_b X_\mu = \\
&= -\frac{1}{2\pi\alpha'} \int_{-\infty}^{\infty} dt \int_0^\ell d\sigma \, \partial_a (\, g^{ab} \, \delta X^\mu \, \partial_b X_\mu \,) + \frac{1}{2\pi\alpha'} \int_\Sigma d^2\xi \, \delta X^\mu \, \partial_a (\, g^{ab} \partial_b X_\mu \,) \\
&= -\frac{1}{2\pi\alpha'} \int_{-\infty}^{\infty} dt \, (\, g^{\sigma b} \, \delta X^\mu \, \partial_b X_\mu \,) \, |_{\sigma=0}^{\sigma=\ell} + \frac{1}{2\pi\alpha'} \int_\Sigma d^2\xi \, \delta X^\mu \, \partial_a (\, g^{ab} \partial_b X_\mu \,)
\end{aligned} \tag{9.7}$$

The second term is the variation that leads to the equations of motion for the 2d fields just like in the closed string. To recover them, we need the first term to vanish. If δX^μ is unconstrained [1] , we then need

$$g^{\sigma b} \partial_b X^\mu(\sigma, t) \, |_{\sigma=0}^{\sigma=\ell} = 0 \tag{9.8}$$

Using this for $X^+ = t$, we get

$$g_{\sigma t} = 0 \ \text{ at } \ \sigma = 0, \ell. \tag{9.9}$$

[1] This is not the case for open string sectors describing lower-dimensional D-branes (to be studied in later lectures).

For the transverse coordinates X^i we get

$$g^{\sigma\sigma}\partial_\sigma X^\mu(\sigma,t)\,|_{\sigma=0}^{\sigma=\ell} = 0 \tag{9.10}$$

We cannot satisfy this equation by requiring $g_{\sigma\sigma} = 0$ at $\sigma = 0, \ell$, since (9.6) would then imply $g_{\sigma\sigma} \equiv 0$ is non-dynamical, in contrast with the situation in closed bosonic strings. Therefore we have to impose

$$\partial_\sigma X^i\,|_{\sigma=0,\ell} \;=\; 0 \tag{9.11}$$

These are Neumann boundary conditions on both open string endpoints, so this kind of open strings are also called Neumann-Neumann or NN.

9.2.3 Hamiltonian

The lagrangian in light-cone gauge is

$$\begin{aligned} L \;&= -\tfrac{1}{4\pi\alpha'}\textstyle\int_0^\ell d\sigma \quad [-2\,g^{tt}\partial_t X^+\,\partial_t X^- \,+ g^{tt}\partial_t X^i\,\partial_t X^i \,-\, 2\,g^{\sigma t}\,\partial_t X^+\,\partial_\sigma X^- \,+ \\ &\qquad +2\,g^{\sigma t}\,\partial_\sigma X^i\,\partial_t X^i \,+ g^{\sigma\sigma}\,\partial_\sigma X^i\,\partial_\sigma X^i\,] \;= \\ &= -\tfrac{1}{4\pi\alpha'}\textstyle\int_0^\ell d\sigma \quad [\,g_{\sigma\sigma}\,(2\,\partial_t X^- \,-\partial_t X^i\,\partial_t X^i) \,-\, 2\,g_{\sigma t}\,(\partial_\sigma X^- - \partial_\sigma X^i\,\partial_t X^i) \,+ \\ &\qquad g_{\sigma\sigma}^{-1}\,(1-g_{\sigma t}^2)\,\partial_\sigma X^i\,\partial_\sigma X^i\,] \end{aligned} \tag{9.12}$$

Defining the center of mass and relative coordinates $x^-(t)$, $Y^-(\sigma,t)$

$$\begin{aligned} x^-(t) \;&=\; \frac{1}{\ell}\int_0^\ell d\sigma\,X^-(\sigma,t) \\ X^-(\sigma,t) \;&=\; x^-(t) \,+\, Y^-(\sigma,t) \end{aligned} \tag{9.13}$$

we obtain

$$\begin{aligned} L \;\;=\;\; &-\frac{\ell}{2\pi\alpha'}\,g_{\sigma\sigma}\,\partial_t x^-(t) \,-\, \frac{1}{4\pi\alpha'}\int_0^\ell d\sigma\,[\,-\,g_{\sigma\sigma}\,\partial_t X^i\,\partial_t X^i \,+ \\ &-2\,g^{\sigma t}\,(\partial_\sigma Y^- - \partial_\sigma X^i\,\partial_t X^i) \,+\, g_{\sigma\sigma}^{-1}\,(1-g_{\sigma t}^2)\,\partial_\sigma X^i\,\partial_\sigma X^i\,] \end{aligned} \tag{9.14}$$

The $Y^-(\sigma,t)$ acts as a Lagrange multiplier imposing

$$\partial_\sigma g_{\sigma,t}(\sigma,t) \;=\; 0 \quad \forall \sigma, t \tag{9.15}$$

From (9.9) we get

$$g_{\sigma,t}(\sigma,t) = 0 \quad \forall \sigma, t \tag{9.16}$$

The lagrangian becomes

$$L \;=\; -\frac{\ell}{2\pi\alpha'}\, g_{\sigma\sigma}\, \partial_t x^-(t) \;+\; \frac{1}{4\pi\alpha'} \int_0^\ell d\sigma \,[\, g_{\sigma\sigma}\, \partial_t X^i\, \partial_t X^i \;-\; g_{\sigma\sigma}^{-1}\, \partial_\sigma X^i\, \partial_\sigma X^i \,]$$

exactly as for closed strings. Following the computations there, the hamiltonian then reads

$$H \;=\; \frac{\ell}{4\pi\alpha' p^+} \int_0^\ell d\sigma \,[\, 2\pi\alpha'\, \Pi_i\, \Pi_i \;+\; \frac{1}{2\pi\alpha'}\, \partial_\sigma X^i\, \partial_\sigma X^i \,] \tag{9.17}$$

9.2.4 Oscillator expansions

From the above hamiltonian, the equations of motion for the 2d fields $X^i(\sigma, t)$ read

$$\partial_t^2 X^i \;=\; \partial_\sigma^2 X^i \tag{9.18}$$

where we have again set $\ell = 2\pi\alpha' p^+$. Again, the general solution will be a superposition of left- and right-moving waves $X_L^i(\sigma + t)$, $X_R^i(\sigma - t)$. These have the general oscillator expansion

$$\begin{aligned} X_L^i(\sigma+t) \;&=\; \frac{x^i}{2} \;+\; \frac{p_i}{2p^+}\,(t+\sigma) \;+\; i\sqrt{\frac{\alpha'}{2}}\, \sum_\nu \frac{\alpha_n^i}{n}\, e^{-\pi i\,\nu\,(\sigma+t)/\ell} \\ X_R^i(\sigma-t) \;&=\; \frac{x^i}{2} \;+\; \frac{p_i}{2p^+}\,(t-\sigma) \;+\; i\sqrt{\frac{\alpha'}{2}}\, \sum_\nu \frac{\tilde{\alpha}_n^i}{n}\, e^{\pi i\,\nu\,(\sigma-t)/\ell} \end{aligned} \tag{9.19}$$

with ν a modding to be fixed by the boundary conditions. Notice that for convenience the exponents we use differ from those in closed strings in a factor of two.

Now we have to impose the boundary conditions

$$\partial_\sigma X_L^i \;+\; \partial_\sigma X_R^i \;=\; 0 \;\text{ at }\; \sigma = 0, \ell \tag{9.20}$$

We compute

$$\partial_\sigma X_L^i \;+\; \partial_\sigma X_R^i \;=\; i\sqrt{\frac{\alpha'}{2}}\, \frac{i\pi}{\ell}\, \sum_\nu \left[-\alpha_\nu^i e^{-\pi i\nu(\sigma+t)/\ell} \;+\; \tilde{\alpha}_\nu^i e^{\pi i\nu(\sigma-t)/\ell} \right] \tag{9.21}$$

Imposing the boundary condition at $\sigma = 0$ we obtain

$$\alpha_\nu^i = \tilde{\alpha}_\nu^i \tag{9.22}$$

The boundary conditions for open strings relate the left and right movers, which are no longer independent. This alos means that the Hilbert space of an open string will be exactly like one of the sides (say the left-moving sector) of a closed string (the right-moving one not being an independent one). Notice that this also means that open strings can couple only left-right symmetric closed string sectors; for instance, there are no heterotic open strings.

Imposing the boundary condition at $\sigma = \ell$ we obtain

$$\alpha_\nu^i \sin \pi\nu \;=\; 0 \tag{9.23}$$

Which implies $\nu \in \mathbf{Z}$

The hamiltonian in terms of the oscillator modes reads

$$H \;=\; \frac{p_i p_i}{2p^+} + \frac{1}{2\alpha' p^+} \left[\sum_{n>0} [\, \alpha_{-n}^i \alpha_n^i \right] + E_0 \tag{9.24}$$

with $E_0 = 24 \times (-1/24) = -1$. This is exactly the hamiltonian for the left-moving sector of the closed bosonic string, except for a factor of two arising from that in the oscillator expansion.

9.2.5 Spectrum

The spectrum is obtained just like the left-moving sector of the closed string theory. The spacetime mass formula is

$$\alpha' m^2 \;=\; N_B \;-\; 1 \;\; \text{with} \;\; N_B \;=\; \sum_{n>0} \alpha_{-n}^i \alpha_n^i \tag{9.25}$$

We define the vacuum by $\alpha_n^i |0\rangle_o = 0$ for $n > 0$, and construct the Hilbert space by applying creation oscillators to it. The lightest modes are

State	$\alpha' m^2$	$SO(24)$
$\|0\rangle_o$	-1	1
$\alpha_{-1}^i \|0\rangle_o$	0	24

(Notice that we get the right Lorentz little group for the massless particles). We obtain a 26d $U(1)$ massless gauge boson and a neutral tachyonic 26d scalar.

To the open string states we have to add the closed string states. Recall thay are given by

State	$\alpha' m^2$	$SO(24)$
$\|0\rangle_c$	-4	1
$\alpha_{-1}^i \tilde{\alpha}_{-1}^j \|0\rangle_c$	0	24×24

where $|0\rangle_c$ is the closed string vacuum. This leads to the 26d closed string tachyon and the massless 26d graviton, 2-form and dilaton.

We would like to briefly mention that, in contrast with the closed string tachyon, there is a general consensus on the meaning of the open string tachyon. It signals an instability because we are expanding the theory around a maximum of the potential for this field. In order to correct this, we should look for a minimum of the tachyon potential and expand the theory around it. The potential indeed has a minimum, and very surprisingly the proposal is that the theory sitting at this minimum is just the closed bosonic string theory, with no open string sector.

The intuition underlying this proposal by A.Sen (and which is a bit advanced for this lecture) is that the open string sector is associated to an underlying objetc which is filling spacetime (a D25-brane). The open string tachyon signals an instability of this object, which decays and disappears. The theory left over is just closed string theory with no open string sector.

Although open string sectors of the bosonic theory are 'unstable' in this sense, it is useful to study them to learn more about string theory, and as background material for other open string sectors without this kind of tachyons.

9.2.6 Open-closed duality

In this section we would like to study how theories with open strings deal with ultraviolet regimes. Consider the simplest 1-loop open string diagram, namely the vacuum to vacuum amplitude given by the annulus. This corresponds to an open string evolving for some time $2T\ell$ and glueing back to itself, see figure B.1a.

This can be computed easily as a trace over the open string Hilbert space. An important difference with respect to the torus in the closed bosonic string is that now we have a fixed reference line, we cannot glue back the open string with a shift in σ; hence we do not have the analog of τ_1. One could imagine to glue back the strin up to an exchange of the roles of the two string endpoints, but this would lead to a worldsheet with the topology of the Moebius string, rather than an annulus. Such worldsheets exist for unoriented open strings,

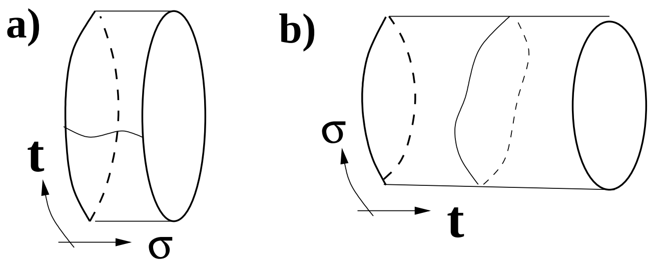


Figure 9.3: The annulus diagram regarded in the open and in the closed string channel.

which couple to unoriented closed string. Since the closed string theories we have studied are oriented, so are our open strings, and we will not consider Moebius strips. In next lecture, type I superstring is an unoriented string theory and will contain such diagrams.

Let us evaluate the annulus amplitude. It is given by a sum over all possible annulus geometries, namely integrating over the parameter T we have

$$Z = \int_0^\infty \frac{dT}{2T} Z(T) \tag{9.26}$$

with

$$Z(T) = \mathrm{tr}_{\mathcal{H}_{\mathrm{op.}}} e^{-2T\ell H_{\mathrm{op.}}} \tag{9.27}$$

Recalling

$$H_{\mathrm{op.}} = \frac{\sum_i p_i^2}{2p^+} + \frac{1}{2\alpha' p^+}(N_B - 1) \tag{9.28}$$

we have

$$Z(T) = \mathrm{tr}_{\mathrm{mom.}} e^{-2\pi\alpha' T \sum_i p_i^2} \mathrm{tr}_{\mathrm{osc.}} e^{-2\pi T(N-1)} \tag{9.29}$$

Defining $q = e^{2\pi i(iT)}$, the traces will organize in modular functions with parameter $\tau = iT$. Computing the traces in a by now familiar way we have

$$Z = \int_0^\infty \frac{dT}{2T} (8\pi^2\alpha' T)^{-12} \eta(iT)^{-24} \tag{9.30}$$

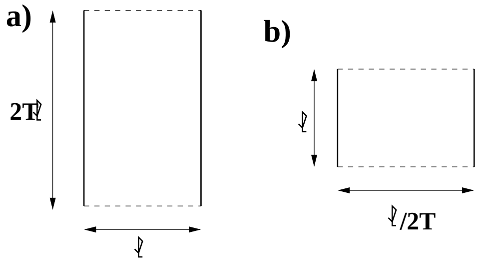


Figure 9.4: Open-closed duality. An open string propagating a time $2T\ell$ is geometrically the same as a closed string propagating a time $T'\ell$ with $T' = 1/(2T)$.

Open-closed duality is the fact that the annulus diagram can be regarded, in a dual channel, as a diagram where closed strings appear from and disappear into the vacuum, at tree level, see figure B.1b). Notice that the ultraviolet regime in the open string channel corresponds to the infrared in the closed string channel, see figure 10.7. Hence the ultraviolet regime is mapped to an infrared regime due to the appearance of a dual channel once stringy energies are reached.

In order to see more manifestly how the amplitude (9.30) can be regarded as a closed string one, notice that in exchanging the roles of σ and t in the annulus there is a redefinition of the new σ to bring it back to the light-cone convention (total length equal to ℓ for closed strings) and hence the closed string propagates for a time $T'\ell$ with $T' = 1/(2T)$. Using the modular transformation properties

$$\eta(i/(2T')) \;=\; (2T)^{1/2}\,\eta(2iT') \tag{9.31}$$

we can write

$$Z \;=\; \int_0^\infty \frac{dT'}{2T'}\,(8\pi^2\alpha')^{-12}\,\eta(2iT')^{-24} \tag{9.32}$$

The same amplitude now has the structure of a sum over closed string states with some peculiarities: there is not power-like dependence on T', meaning that the closed states are created out of the vacuum with zero momentum (due to momentum conservation); also, there is no analog of τ_1 since the closed string does not come back to itself; finally, due to the absence of integration over τ_1 (because there is no τ_1) the level matching on closed states has to be imposed explicitly, this leads to the argument of the oscillator η functions to be doubled.

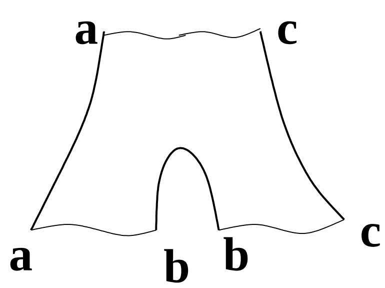


Figure 9.5: Open string interaction vertex with Chan-Paton factors.

9.3 Chan-Paton factors

We now turn to the discussion of an essentially new feature of open strings. It is consistent to have more than one kind of open string sector in a string theory. The most general possibility is to introduce a discrete degree of freedom, in one out of N possible states, at each string endpoint. Hence, each open string is characterized by two indices, a, b, with $a, b = 1, \ldots, N$, denoted Chan-Paton indices, specifying in which states the endpoints are. Notice that the labels are ordered for oriented open strings.

These degrees of freedom are non-dynamical, so the label of an endpoint simply propagates unchanged along the endpoint worldline. The rules for interactions are clear, there is one label per boundary, and one should sum over all possible labels in internal boundaries. The basic interaction vertex is shown in figure 9.5.

The quantization of open strings with Chan-Paton factors is straightforward. Since Chan-Paton indices are non-dynamical, they do not enter in the hamiltonian, and the quantization of each ab sector proceeds as for a single open string without Chan-Paton factors. The existence of the indices only implies that there are N^2 states of each kind. The lightest states are as follows

State	$\alpha' m^2$	$SO(24)$
$\|0\rangle_{ab}$	-1	1
$\alpha^i_{-1}\|0\rangle_{ab}$	0	24

where $|0\rangle_{ab}$ denotes the groundstate of the ab open string. Hence we obtain N^2 gauge bosons and N^2 scalar tachyons. The N^2 gauge bosons A_{ab} can be seen to correspond to a gauge group $U(N)$. This can be seen by analyzing their interactions as follows, see fig 9.6.

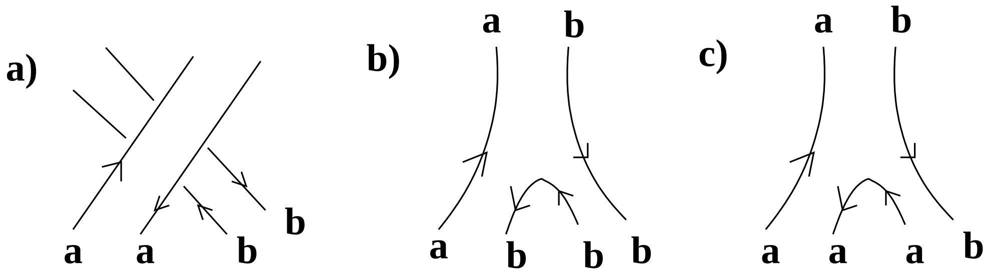


Figure 9.6: Interactions between open string with Chan-Paton factors.

• The gauge bosons A_{aa}, A_{bb} for $a \neq b$ do not interact among themselves, since they do not have common indices, fig 9.6a. This means that the corresponding generators of the gauge group commute. In fact, they generate a $U(1)^N$ Cartan subalgebra.

• The gauge boson A_{ab} interacts with, i.e. is charged under A_{aa}, A_{bb}, as shown in figures 9.6b,c. The orientations of the boundary are inherited from the orientation on whe worldsheet. The orientations imply that 9.6b, c differ by a sign. Fixing a convention, we say that A_{ab} carries charge $+1$ and -1 under A_{aa}, A_{bb}.

Since charge under Cartan generators correspond to weights, and since weights in the adjoint representation (in which gauge bosons must transform) are roots, we obtain that the gauge group has $N^2 - N$ non-zero roots of the form

$$(\underline{+, -, 0, \ldots, 0}) \tag{9.33}$$

Going back to the lecture on group theory, we see that these are the non-zero roots of $U(N)$.

Performing a similar discussion it is easy to see that all states in the open string tower transform in the adjoint representation of $U(N)$.

An alternative way to understand the appearance of $U(N)$ is to consider general states, linear combinations of the basic states $|\ \rangle_{ab}$

$$|\ \rangle = \sum_{ab} \lambda_{ab} |\ \rangle_{ab} \tag{9.34}$$

where the matrix of coefficients λ is hermitian. These hermitian matrices are providing an N-dimensional representation of the $U(N)$ generators. Notice that a single Chan-Paton index a can be thought of as transforming in the fundamental or antifundamental representation of the $U(N)$ group, depending on whether it sits at the endpoint where the string starts from or ends at.

It is very remarkable that the simple non-dynamical Chan-Paton degrees of freedom lead to the rich dynamics of non-abelian gauge symmetry from the viewpoint of spacetime. Also very remarkably, we have uncovered a brand new way to obtain non-abelian gauge symmetries in string theory.

As a final comment, it is easy to see that open-closed duality is satisfied for any choice of the Chan-Paton rank N. The annulus amplitude is exactly as the above up to a multiplicity factor of N^2. Upon going to the closed channel, this implies there is an additional factor of N on the disk diagrams creating or annihilating the closed string from or into the vacuum.

Notice finally that the number of open string tachyons increases with N. Hence the more open string sectors the theory has, the more unstable it is in this sense. As with the single open string case, condensation of these tachyons leads to the disappearance of the open string sectors, leaving behind just the closed bosonic string theory.

9.4 Open superstrings

Let us try to consider describing open superstrings. We know that they will couple to some closed superstring, which must be of the kind studied in previous lectures. Since the local 2d worldsheet must be left-right symmetric, the natural possibility to be considered is open string theories coupling to type IIB closed string sectors.

At the end of this section we will see that in superstrings there is an additional consistency condition, called RR tadpole cancellation condition, which is not satisfied by the models we are about to construct. Nevertheless, the material we cover will turn out to be useful for the construction of type I theory, which is consistent, in next lecture.

9.4.1 Hamiltonian quantization

In the light-cone gauge the dynamical 2d fields are $X_L^i(\sigma+t)$, $\psi_L^i(\sigma+t)$, $X_R^i(\sigma-t)$, $\psi_R^i(\sigma-t)$, with $i=2,\ldots,9$. The quantization of the bosonic piece works exactly like in the open bosonic string, and will not be reviewed here.

Centering on the 2d fermions, let us simply state, without entering into details, that there are two possible boundary conditions which lead to the correct equations of motion locally on the worldsheet. The possibilities are

$$\begin{aligned} \psi^i_L &= e^{2\pi i\rho}\,\psi^i_R \quad \text{at } \sigma=0 \\ \psi^i_L &= e^{2\pi i\rho'}\,\psi^i_R \quad \text{at } \sigma=\ell \end{aligned} \tag{9.35}$$

with $\rho,\rho'=0,1/2$. Redefining $\psi^i_R(\sigma-t)\to e^{-2\pi i\rho'}\psi^i_R(\sigma-t)$ we can trivialize the condition at $\sigma=\ell$, hence we are left with two possible sectors, which we call NS and R

$$\begin{array}{llllll} \text{NS} & \psi^i_L=-\psi^i_R & \text{at } \sigma=0 & \text{R} & \psi^i_L=\psi^i_R & \text{at } \sigma=0 \\ & \psi^i_L=\psi^i_R & \text{at } \sigma=\ell & & \psi^i_L=\psi^i_R & \text{at } \sigma=\ell \end{array}$$

The mode expansion in both cases reads

$$\begin{aligned} \psi^i_L(\sigma+t) &= i\sqrt{\frac{\alpha'}{2}}\sum_\nu \psi^i_\nu\, e^{-\pi i\nu(\sigma+t)/\ell} \\ \psi^i_R(\sigma-t) &= i\sqrt{\frac{\alpha'}{2}}\sum_\nu \tilde\psi^i_\nu\, e^{\pi i\nu(\sigma-t)/\ell} \end{aligned} \tag{9.36}$$

For NS boundary conditions, we have

$$\begin{array}{lll} \sigma=0 & \sum_\nu(\psi^i_\nu+\tilde\psi^i_\nu)\,e^{-\pi i\nu t/\ell}=0 & \to\quad \psi^i_\nu=-\tilde\psi^i_\nu \\ \sigma=\ell & \sum_\nu \psi^i_\nu\,\cos\pi\nu\,e^{-\pi i\nu t/\ell}=0 & \to\quad \nu\in\mathbf{Z}+\mathbf{\frac{1}{2}} \end{array} \tag{9.37}$$

For R boundary conditions, we have

$$\begin{array}{lll} \sigma=0 & \sum_\nu(\psi^i_\nu-\tilde\psi^i_\nu)\,e^{-\pi i\nu t/\ell}=0 & \to\quad \psi^i_\nu=\tilde\psi^i_\nu \\ \sigma=\ell & \sum_\nu \psi^i_\nu\,\sin\pi\nu\,e^{-\pi i\nu t/\ell}=0 & \to\quad \nu\in\mathbf{Z} \end{array} \tag{9.38}$$

So left and right movers are linked together. NS fermions are half-integer modded and R fermions have integer moddings. Everything behaves as with the left moving sector of a superstring.

9.4.2 Spectrum for NS and R sectors

Being careful with the factor of 2 from the different exponent in the oscillator expansions, the hamiltonian and mass formula are similar to the left moving ones in a superstring. They are given by

$$\begin{aligned} H &= \frac{\sum_i p_i^2}{2p^+} + \frac{1}{2\alpha' p^+} \left(N_B + N_F + E_0 \right) \\ \alpha' m^2 &= N_B + N_F + E_0 \end{aligned} \tag{9.39}$$

with $E_0 = -1/2, 0$ for NS, R sectors.

In the NS sector, we take the groundstate annihilated by positive modding operators

$$\alpha_n|0\rangle = 0 \quad , \quad \psi_{n-1/2}|0\rangle = 0 \quad , \quad \text{for } n > 0 \tag{9.40}$$

and build the Hilbert space by applying negative modding oscillators to it. The lightest states are

State	$\alpha' m_L^2/2$	$SO(8)$
$\|0\rangle$	$-1/2$	$\mathbf{1}$
$\psi^i_{-1/2}\|0\rangle$	0	$\mathbf{8_V}$

In the R sector, we define the groundstates as annihilated by positive modding operators

$$\alpha_n|0\rangle = 0 \quad , \quad \psi_n|0\rangle = 0 \quad , \quad \text{for } n > 0 \tag{9.41}$$

The groundstate is degenerate due to fermion zero modes, and hence forms a representation of the Cliffor algebra generated by them. Introducing the operators $A_a^\pm = \psi_0^{2a} \pm i\psi_0^{2a+1}$, and the state $|0\rangle$ annihilated by the raising operator, the groundstates are

$$\begin{array}{cc} |0\rangle & A_{a_1}^+|0\rangle \\ A_{a_1}^+ A_{a_2}^+|0\rangle & A_{a_1}^+ A_{a_2}^+ A_{a_3}^+|0\rangle \\ A_1^+ A_2^+ A_3^+ A_4^+|0\rangle & \end{array} \tag{9.42}$$

The two columns correspond to the two chiral irreps of $SO(8)$, 8_S and 8_C respectively. Finally the spectrum is obtained by applying negative modding oscillators to these groundatates. The lightest modes are the groundstates themselves

	State		$\alpha' m_L^2/2$	$SO(8)$
R	$\frac{1}{2}(\pm,\pm,\pm,\pm)$	$\#- =$ even	0	$\mathbf{8_S}$
	$\frac{1}{2}(\pm,\pm,\pm,\pm)$	$\#- =$ odd	0	$\mathbf{8_C}$

9.4.3 GSO projection

A natural question is now how (or whether) to combine NS and R sectors in constructing the open string spectrum (as was required by modular invariance in closed superstrings). Clearly, the fact that the open strings we want to construct couple to type IIB closed string imposes a constraint on the physical spectrum of the open string. Indeed, the physical spectrum of the closed sector had a GSO projection; if no constraint is imposed on the open string spectrum, it would be possible to create unphysical closed string states (with the wrong GSO behaviour) by scattering open string states.

In other words, open/closed duality (the fact that the open 1-loop annulus diagram can be regarded as a closed string amplitude (with only GSO projected states propagating) requires the open string sector to have a specific mixture of NS and R boundary condition, i.e. a GSO projection.

Indeed, it turns out that the GSO projection in the open string sector is exactly that on one of the sides in a type II superstring. Namely, it eliminates the NS groundstate, and the 8_S R groundstate. Hence the open string tachyon disappears, and the only massless states are a 10d $U(1)$ gauge boson and a 10d chiral fermion. They fill out a vector multiplet of 10d $\mathcal{N} = 1$ supersymmetry.

The complete spectrum is given by this open string spectrum, plus the closed type IIB string spectrum, which at the massless level is 10d $\mathcal{N} = 2$ supersymmetry. This supersymmetry in the closed sector is not a symmetry of the full theory, and it would be broken to $\mathcal{N} = 1$ by interactions with open strings.

Let us finish by mentioning that addition of Chan-Paton indices is straightforward and leads to the same result as for bosonic open strings, namely the gauge group becomes non-abelian $U(N)$ and all states transform in the adjoint representation. This leads to a new situation, very different from heterotic, with non-abelian gauge symmetries and charged fermions. So it in principle provides an interesting starting point for model building of theories similar to the Standard Model (see future lectures on D-branes worlds).

9.4.4 Open-closed duality

Let us verify that the annulus constructed in the open string channel indeed reproduces a GSO projected closed string amplitude in the dual channel.

The annulus amplitude is

$$Z = \int_0^\infty \frac{dT}{2T}\, Z(T) \tag{9.43}$$

with

$$\begin{aligned} Z(T) &= \mathrm{tr}_{\mathcal{H}_{\rm op.}}\, e^{-2T\ell H_{\rm op.}} = \mathrm{tr}_{\rm mom.} e^{-2\pi\alpha' T\sum_i p_i^2}\, \mathrm{tr}_{\rm bos.}\, e^{-2\pi T(N_B - E_0^B)} \times \\ &\times \left(\mathrm{tr}_{NS,GSO}\, e^{-2\pi T(N_F - E_0^F)} - \mathrm{tr}_{R,GSO}\, e^{-2\pi T(N_F - E_0^F)}\right) \end{aligned} \tag{9.44}$$

We have

$$\begin{aligned} \mathrm{tr}_{\rm mom.} e^{-2\pi\alpha' T\sum_i p_i^2} &= (8\pi^2\alpha' T)^{-4} \\ \mathrm{tr}_{\rm bos.}\, e^{-2\pi T(N_B - E_0^B)} &= \eta(iT)^{-8} \\ \mathrm{tr}_{NS,GSO}\, e^{-2\pi T(N_F - E_0^F)} &= \frac{1}{2}\left(\mathrm{tr}_{NS} q^{N_F + E_0^F} + \mathrm{tr}_{NS}(q^{N_F+E_0^F}\,(-)^F\right) = \\ & \frac{1}{2}\,\eta^{-4}\left(\vartheta\begin{bmatrix}0\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^4\right) \\ \mathrm{tr}_{R,GSO}\, e^{-2\pi T(N_F - E_0^F)} &= \frac{1}{2}\left(\mathrm{tr}_{R} q^{N_F + E_0^F} + \mathrm{tr}_{R}(q^{N_F+E_0^F}\,(-)^F\right) = \\ & \frac{1}{2}\,\eta^{-4}\left(\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^4\right) \end{aligned} \tag{9.45}$$

In total

$$Z(T) = \frac{1}{2}\,(8\pi^2\alpha' T)^{-4}\,\eta^{-8}\,\eta^{-4}\left(\vartheta\begin{bmatrix}0\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^4 - \vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^4 \pm \vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^4\right)(2iT) \tag{9.46}$$

It is clear that replacing $T = 1/(2T')$ and using the modular properties of the eta and theta functions we recover a correctly GSO projected closed string amplitude.

9.4.5 RR tadpole cancellation condition

Although everything looks fine, clearly there must be something wrong in the above construction. In previous lectures we mentioned that the field content of type IIB theory is free of gravitational anomalies in a very intricate and miraculous manner. Here we are seemingly constructing a bunch of

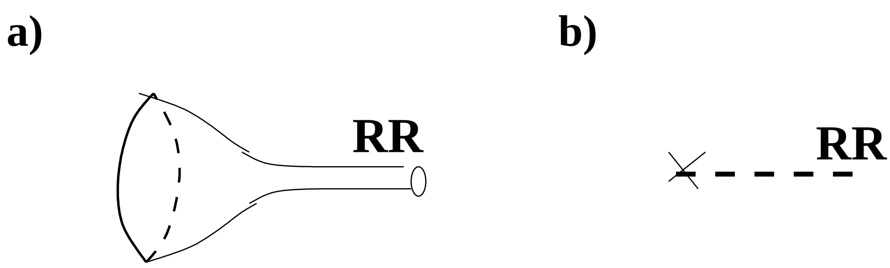


Figure 9.7: Disk diagram leading to a tadpole term for some closed string mode.

theories which include the anomaly free type IIB field content, plus a bunch of additional chiral fields arising from the open string sectors.

The additional sets of fields in these theories are anomalous, so it is *not* possible that these theories with open string sectors are consistent.

Indeed we are going to learn that in theories with open superstrings there is a consistency condition which we had not satisfied, and which renders inconsistent all the above theories unless $N = 0$, namely unless open string sectors are absent.

Let us discuss the physical idea, since the computations will be done in some more detail in the lecture on type I superstrings. The key idea is that the theory contains tadpole interactions due to disk diagrams of the kind shown in figure 9.7. From the spacetime viewpoint, these are terms in the effective action, which are linear in the closed sector field, schematically

$$Q \int d^{10}x \, \varphi(x) \tag{9.47}$$

with Q the coefficient of the disk tadpole, and φ the corresponding closed string field.

It is possible to compute explicitly in string theory which closed string fields get this kind of tadpoles, but much can learnt from simple considerations. First, the terms should be Poincare invariant in order to appear in the effective action. In the RR sector, massless fields are p-forms in spacetime, for all possible even p degrees. The only p-form for which the tadpole term is Poincare invariant is the 10-form C_{10}. This field is very peculiar, since its field strength would be an 11-form which is identically zero in a 10d spacetime. Hence, and although it has a vertex operator in string theory, it has not kinetic term. The only place where it appears in the spacetime action is in fact the tadpole term. Hence we have

$$S_{C_{10}} = Q \int_{M_{10}} C_{10} \tag{9.48}$$

The equation of motion for this field is therefore

$$Q = 0 \tag{9.49}$$

Namely, rather than a condition on the field, it is a consistency condition on the theory. It requires that the RR tadpole is absent from the theory. This is the RR tadpole cancellation condition.

It is possible to check that the coefficient of the tadpole diagram is non-zero if there are open string sectors. Indeed, the standard way to compute the disk (see lecture on type I) is to compute the annulus and take the infinite T' limit in the closed string channel, where the amplitude factorizes as the square of the disk. Recalling that with N Chan-Paton factors, the annulus goes like N^2, the disk and hence the tadpole is proportional to N. Consequently (9.49) requires $N = 0$, namely no open string sectors.

This is our result. The derivation was a bit crude, in particular since it involved spacetime considerations. Nevertheless the result is robust and has been derived (in a very technical way) purely from worldsheet consideration [54].

We would like to conclude with two comments. In addition to the RR tadpole, there is also a tadpole for NSNS fields. This tadpole is not a dangerous one, since all fields in the NSNS sector have kinetic terms, hence their equations of motion impose conditions on the fields and not consistency conditions on the theory [2]. This is analogous to open bosonic strings, where disk tadpoles exist for fields with kinetic terms, hence do not signal inconsistencies.

Finally, let us mention what theories are affected by the inconsistency. The precise statement is that it is not possible to couple open strings to type IIB closed string in a 10d Poincare invariant way. In further lectures we will encounter consistnte situations with open superstrings, which avoid the above problem: either because the open strings are unoriented and couple to an unoriented version of type IIB theory (but not to just type IIB theory);

[2] In any event, supersymmetry relates NSNS and RR tadpoles, so that often in imposing RR tadpole cancellation conditions one obtains NSNS tadpole cancellation, although the latter is not required for consistency.

or beacuse the open string sectors do not preserve 10d Poincare invariance (see lecture on D-branes).

Chapter 10

Type I superstring

10.1 Unoriented closed strings

10.1.1 Generalities

Consider a closed oriented string theory which is left-right symmetric, e.g. closed bosonic string theory or type IIB theory. Consider modding it out, quotienting, by the operation Ω, worldsheet parity, that exchanges left and right movers. Namely, construct the quotient theory, where states related by left-right exchange are considered equivalent

$$|a\rangle_L \otimes |b\rangle_R \quad |a\rangle_R \otimes |b\rangle_R \tag{10.1}$$

This operation is called orientifolding the theory by Ω (this is also called gauging the global symmetry Ω).

The genus expansion in the quotient theory is drastically different from the original one. Consider for instance 1-loop vacuum diagrams. As usual we have the torus, which corresponds to closed string states $A_L \times B_R$ which evolve and are glued back to the original state. In theories where states related by Ω are considered equivalent, there is a new diagram. It corresponds to starting with a closed string state $A_L \times B_R$ letting it evolve and glueing it back to the original up to the action of Ω. This is shown graphically in figure 10.1 where we can see the worldsheet is a non-orientable surface, a Klein bottle. The result generalizes to other amplitudes as the statement that the genus expansion of unoriented theories contains non-orientable worldsheets.

A general worldsheet (including oriented and unoriented ones) can be described as a sphere with an arbitrary number of handles and crosscaps. A

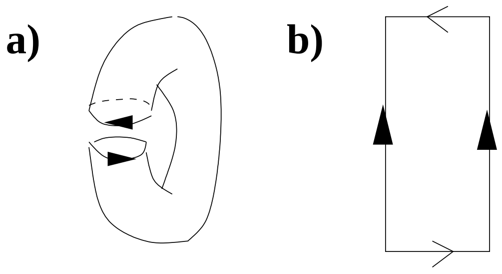


Figure 10.1: Two pictures representing the Klein bottle. In b) we construct it as a rectangle with vertical sides identified with the same orientation and horizontal sides glued with the reversed orientation, as suggested by the arrow.

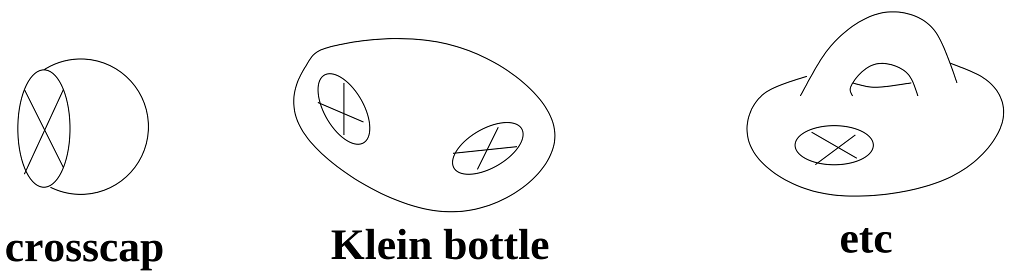


Figure 10.2: Several examples of non-orientable surfaces constructed by glueing crosscaps to a sphere.

crosscap can be described as cutting a small disk in the surface and identifying antipodal points in the resulting boundary to close back the surface. Several non-orientable surfaces are shown in figure 10.2. In theories with open string sectors (see later) the genus expansion contains worldsheets with boundaries. Recalling the discussion in the review lectures, an amplitude mediated by a worldsheet with g handles, n_c crosscaps and n_b boundaries is weighted by a factor of $e^{-\xi\phi}$, where ϕ is the dilaton vev and $\xi = 2 - 2g - n_c - n_b$ is the Euler characteristic of the worldsheet.

The spectrum of the unoriented theory is obtained from the spectrum of the 'parent' oriented theory very simply. Namely, one takes the original spectrum and keeps only the Ω invariant states (or linear combinations of states). The same result is obtained from our description of the genus expansion: One way to obtain the spectrum of a theory is to see what states contribute in the one-loop vacuum amplitude. The sum over the two contributions, the torus and the Klein bottle, can be written in terms of the original Hilbert space of states as

$$\begin{aligned} \mathrm{tr}_{\mathcal{H}_{\rm oriented}} (\ldots) \,+\, \mathrm{tr}_{\mathcal{H}_{\rm oriented}} (\ldots\,\Omega), = \\ = \mathrm{tr}_{\mathcal{H}_{oriented}} [\ldots\, \tfrac{1}{2}(1+\Omega)] \end{aligned} \tag{10.2}$$

the piece $\frac{1}{2}(1+\Omega)$ is a projector that only keeps Ω invariant states, so that only the later contribute to the amplitude. For non-invariant states, the contributions from the torus and Moebius strip cancel each other; the sum over topologies projects out those states.

10.1.2 Unoriented closed bosonic string

Let us obtain the precise action of Ω on closed string states in a systematic way (to be used in other cases as well). The action of Ω on the 2d bosonic field $X(\sigma,t)$ is to transform it into a field $X^{i'} = \Omega X^i \Omega^{-1}$ such that

$$X^{i'}(\sigma,t) \,=\, X^i(\ell-\sigma,t) \tag{10.3}$$

Introducing the oscillator expansion

$$X^i(\sigma,t) \;=\; x^i \,+\, \frac{p^i}{p^+}\,t \,+ i\sqrt{\frac{\alpha'}{2}}\,\sum_{n\neq 0}\left[\frac{\alpha_n^i}{n}\,e^{-2\pi i\,n(\sigma+t)\ell} \,+\, \frac{\tilde{\alpha}_n^i}{n}\,e^{2\pi i\,n(\sigma-t)\ell}\right] \tag{10.4}$$

we obtain

$$x^{i\prime} = x^i \quad ; \quad p^{i\prime} = p^i \quad ; \quad \alpha_n^{i\,\prime} = \tilde{\alpha}_n^i \quad ; \quad \tilde{\alpha}_n^{i\,\prime} = \alpha_n^i \tag{10.5}$$

which corresponds to an exchange of the left and right movers, as expected.

The quotient theory is obtained by taking the vacuum of the original theory

$$\alpha_n^i|0\rangle = 0 \quad ; \quad \tilde{\alpha}_n^i|0\rangle = 0 \quad \forall n > 0 \tag{10.6}$$

and applying left and right oscillators forming Ω invariant states. The spacetime mass of these states is given by the original formula

$$\alpha' m^2/2 \;=\; N_B \,+\, \tilde{N}_B \,-\, 2 \tag{10.7}$$

The lightest modes are

State	$\alpha' m^2/2$	Lorentz rep
$\|0\rangle$	-2	scalar
$\alpha_{-1}^{(i}\tilde{\alpha}_{-1}^{j)}\|0\rangle$	0	graviton
$\sum_i \alpha_{-1}^i\tilde{\alpha}_{-1}^i\|0\rangle$	0	dilaton

We see that the 2-form of the original theory is odd under Ω and is projected out. The complete spectrum is easily obtained.

This concludes the construction of our theory, which can be checked to be completely consistent. In the following sections we try to construct an unoriented version of the (type IIB) superstring.

10.1.3 Unoriented closed superstring theory IIB/Ω

The worldsheet theory is in this case described by the 2d bosonic and fermionic fields $X^i(\sigma,t)$, $\psi^i(\sigma,t)$. The bosonic fields are discussed exactly as above. On the fermionic fields, the action of Ω is such that

$$\psi^{i\prime}(\sigma,t) \;=\; \psi^i(\ell-\sigma,t) \tag{10.8}$$

Using the oscillator expansion

$$\psi^i(\sigma,t) \;=\; i\sqrt{\frac{\alpha}{2}}\sum_{r\in\mathbf{Z}}\left[\,\psi_{r+\nu}^i\, e^{-2\pi i\,(r+\nu)(\sigma+t)/\ell} \;+\; \tilde{\psi}_{r+\nu}^i\, e^{2\pi i\,(r+\nu)(\sigma-t)/\ell}\,\right] \tag{10.9}$$

where $\nu = 1/2, 0$ for NS and R fermions, we obtain

$$\psi^i_{r+\nu}{}' = \tilde{\psi}^i_{r+\nu} \quad ; \quad \tilde{\psi}^i_{r+\nu}{}' = \psi^i_{r+\nu} \quad ; \tag{10.10}$$

We can now obtain the spectrum of the unoriented theory, which is simply obtained by taking the Ω invariant states of the original theory. There is an interesting subtlety in the action of Ω on RR states; since the left and right moving pieces in this sector are spacetime spinors, they anticommute, so that a state $A_L \times B_R$ is mapped by Ω to $A_R \times B_L = -B_L \times A_R$. The Ω invariant states are therefore of the form $A_L \times B_R - B_L \times A_R$. Notice also that states in the NS-R sector must combine with states in the R-NS sector to form invariant combinations.

The light spectrum is given by

Sector	State	$SO(8)$	Field
NS-NS	$\psi^{(i}_{-1/2}\vert 0\rangle \otimes \psi^{j)}_{-1/2}\vert 0\rangle$	$1 + 35_v$	dilaton, graviton
NS-R+R-NS	$\psi^i_{-1/2}\vert 0\rangle \otimes \tilde{8}_C + 8_C \otimes \tilde{\psi}^i_{-1/2}\vert 0\rangle$	$56_S + 8_S$	gravitinos
R-R	$[8_C \otimes \tilde{8}_C]$	28_C	2-form

This spectrum corresponds to the gravity multiplet of 10d $\mathcal{N} = 1$ supergravity. Notice in particular that the orientifold projection kills one linear combination of the two gravitinos of the original $\mathcal{N} = 2$ supersymmetric type IIB theory.

This theory as it stands is clearly not consistent. A theory whose spectrum is just the gravity multiplet of 10d $\mathcal{N} = 1$ theory has 10d gravitational chiral anomalies. Clearly we have missed an important consistency condition in the construction of the theory.

The consistency condition is RR tadpole cancellation. Unoriented theories contain a diagram, given by a crosscap with an infinite tube attached to it (see fig. 10.3), which leads to a tadpole for certain massless fields. In particular there is a tadpole for a RR field, which due to 10d Poincare invariance, must be the non-propagating 10-form C_{10} (which can be seen to survive the orientifold projection). This RR tadpole renders the theory inconsistent.

The fact that the problem in constructing a theory of just unoriented closed strings is very similar to the problem of constructing a theory of open strings coupled to type IIB theory leads to the following suggestion. One can attempt to construct a theory free of RR tadpoles by considering the Ω

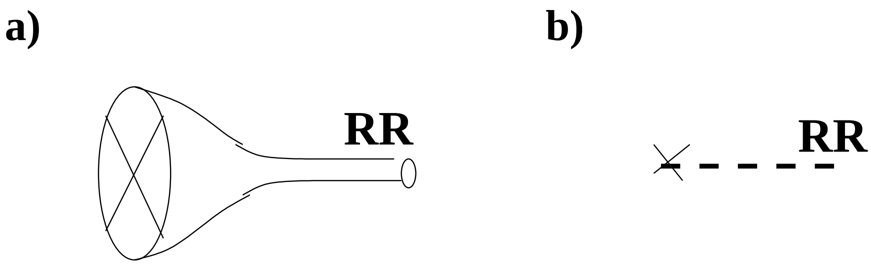


Figure 10.3: Crosscap diagram leading to a tadpole term for some closed string mode.

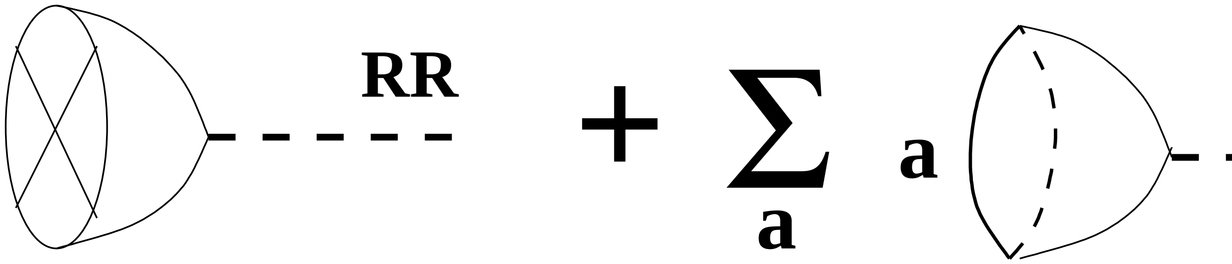


Figure 10.4: Cancellation of RR tadpoles from crosscap and disk diagrams.

orientifold of type IIB coupled to a sector of open strings. Namely, we can attempt to construct a theory where the RR tadpoles for C_{10} arising from open string sectors (disk diagrams) and unorientability (crosscap diagrams) cancel each other. This is the so-called type I superstring theory.

In other words, the equation of motion from the action for the 10-form

$$S_{C_{10}} = (Q_{\rm crosscap} + Q_{rmdisks}) \int C_{10} \qquad (10.11)$$

would be satisfied

$$Q_{\rm crosscap} + Q_{\rm disk} = 0 \qquad (10.12)$$

This is pictorially shown in figure 10.4. Open string sectors coupling to unoriented closed string must be unoriented as well. Hence if one is able to construct such a theory, it would be a theory of unoriented open and closed strings. Hence we need to know a bit about unoriented open strings before the final construction.

10.2 Unoriented open strings

10.2.1 Action of Ω on open string sectors

As mentioned in previous lectures, the local structure on the 2d worldsheet for open strings should be the same as for the corresponding closed sector. Hence, the action on the bosonic coordinates is such that

$$X^{i\prime}(\sigma,t) = X^i(\ell-\sigma,t) \tag{10.13}$$

Using the oscillator expansion for open strings,

$$X^i(\sigma,t) \;=\; x^i \,+\, \frac{p^i}{p^+}t \,+\, i\sqrt{\frac{\alpha'}{2}}\sum_n \frac{\alpha_n^i}{n}\cos\frac{\pi n\sigma}{\ell}\, e^{-\pi i n t/\ell} \tag{10.14}$$

we obtain

$$x^{i\prime} = x^i \quad ; \quad p^{i\prime} = p^i \quad ; \quad \alpha_n^{i\,\prime} = (-1)^n\alpha_n^i \tag{10.15}$$

The action on fermions is such that

$$\psi^{i\prime}(\sigma,t) = \psi^{'}(\ell-\sigma,t) \tag{10.16}$$

Using the expansion

$$\psi^i(\sigma,t) \;=\; i\sqrt{\frac{\alpha'}{2}}\sum_{r\in\mathbf{Z}}\Big[\,\psi_{r+\nu}e^{-\pi i(r+\nu)(\sigma+t)/\ell} \,+\, (-1)^{2\nu}\,\psi_{r+\nu}e^{\pi i(r+\nu)(\sigma-t)/\ell}\Big] \tag{10.17}$$

with $\nu = 1/2, 0$ for NS and R fermions, resp, we obtain

$$\psi^{i\prime}{}_{r+\nu} = (-1)^{r+\nu}\psi_{r+\nu} \tag{10.18}$$

It should be pointed out at this stage that there is a non-trivial action of Ω on the open string NS groundstate, namely

$$\Omega|0\rangle_{NS} = e^{-i\pi/2}|0\rangle_{NS} \tag{10.19}$$

Finally, we also need to specify the action of Ω on the Chan-Paton indices in cases where they are present. Clearly Ω exchanges the order of the labels ab, since it reverses the orientation of the open string.

A general state with fixed operator structure may be written as a linear combination of the corresponding state in the different open string sectors, of the form $\lambda_{ab}|ab\rangle$. The $N \times N$ matrix λ_{ab} is known as the Chan-Paton wavefunction of the state. The action of Ω on Chan-Paton labels can be encoded into an action on λ

$$\lambda \xrightarrow{\Omega} \gamma_\Omega \lambda^T \gamma_\Omega^{-1} \tag{10.20}$$

where γ_Ω is an $N \times N$ unitary matrix or order two. There are two canonical choices, distinguished by the symmetry of γ_Ω

$$\begin{aligned} \text{i)} \qquad \gamma_\Omega &= \mathbf{1}_N \\ \text{ii)} \qquad \gamma_\Omega &= \begin{pmatrix} 0 & i\mathbf{1}_{N/2} \\ -i\mathbf{1}_{N/2} & 0 \end{pmatrix} \end{aligned} \tag{10.21}$$

The first option i) is also often described as

$$\gamma_\Omega = \begin{pmatrix} 0 & \mathbf{1}_{(N/2} \\ \mathbf{1}_{N/2} & 0 \end{pmatrix} \text{ for } N = \text{even} \ ; \ \gamma_\Omega = \begin{pmatrix} 1 & & \\ & 0 & \mathbf{1}_{(N-1)/2} \\ & \mathbf{1}_{(N-1)/2} & 0 \end{pmatrix} \text{ for } N = \text{odd} \tag{10.22}$$

A more transparent interpretation of these actions on Chan-Paton labels is as follows (we take N even for simplicity). Consider splitting the set of labels into two sets, running from 0 to $N/2$ and from $N/2+1$ to N, and label them by indices a, and a'. Denoting the Chan Paton index part of a state by e.g. $|ab\rangle$, the actions above are

$$\begin{aligned} |ab\rangle \to |b'a'\rangle \quad &; \qquad |a'b'\rangle \to |ba\rangle \\ |ab'\rangle \to \pm|ba'\rangle \quad &; \qquad |a'b\rangle \to \pm|ba'\rangle \end{aligned} \tag{10.23}$$

with $+,-$ signs for symmetric or antisymmetric γ_Ω.

10.2.2 Spectrum

It is now easy to obtain the spectrum of the unoriented open string sector, by simply keeping the states of the original theory invariant under the combined action of Ω on the oscillator operators, the vacuum and the Chan Paton labels. We center on the massless sector.

In the NS sector, the states $\lambda\psi^i_{-1/2}|0\rangle$ transform as

$$\lambda\, \psi^i_{-1/2}\,|0\rangle \ \xrightarrow{\Omega}\ -\gamma_\Omega \lambda^T \gamma_\Omega^{-1}\, \psi^i_{-1/2}\,|0\rangle \tag{10.24}$$

Invariant states correspond to components of the matrix λ surviving the projection

$$\lambda = -\gamma_\Omega \lambda^T \gamma_\Omega^{-1} \tag{10.25}$$

In case i), we obtain $\lambda = -\lambda^T$, so there are $N(N-1)/2$ surviving gauge bosons. This number, and the relation with antisymmetric matrices as generators, suggest that the gauge bosons fill out a gauge group $SO(N)$.

In case ii), writing $\lambda = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, the projection imposes $A = -D^T$, $B = B^T$, $C = C^T$. There are $N(N+1)/2$ gauge bosons, and this suggests that the gauge group is $USp(N)$.

In the R sector, the GSO projection selects the groundstate transforming as 8_C. The action of Ω turns out to introduce a minus sign on it, so the projection condition on λ is again

$$\lambda = -\gamma_\Omega \lambda^T \gamma_\Omega^{-1} \tag{10.26}$$

Hence in cases i) and ii) we get 10d fermions in the adjoint representation of $SO(N)$ and $USp(N)$ respectively. The NS and R sectors altogether give an $SO(N)$ or $USp(N)$ vector multiplet of 10d $\mathcal{N} = 1$ supersymmetry. So the open string sector preserves the same amount of supersymmetry as the unoriented closed string sector.

10.3 Type I superstring

As discussed above, the idea in the construction of type I superstring is to add (unoriented) open string sectors to the unoriented closed string theory in section 1, in such a way that the contribution of disks and crosscaps to the 10-form RR tadpole cancels.

10.3.1 Computation of RR tadpoles

The idea

Instead of computing directly the disk and crosscap diagrams with insertions of the massless RR field, there is an indirect but standard way of computing them. In particular it is useful in making sure the disk and crosscaps come out with the same normalization (which is clearly crucial to have correct cancellation).

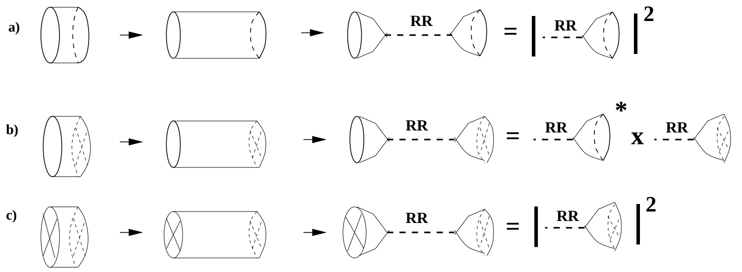


Figure 10.5: Disk and crosscap tadpoles can be recovered in the factorization limit of certain one-loop amplitudes, namely the annulus (a), the Moebius strip (b) and the Klein bottle (c).

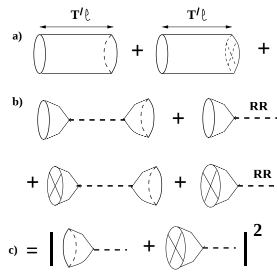


Figure 10.6: The sum of four amplitudes factorizes as the square of the total disk plus crosscap tadpole.

The idea is that since we are interested in computing e.g. the disk with insertion of a massless field, this can be recovered from an annulus amplitude with no insertions, in the limit in which it factorizes in the closed string channel. this is shown in figure 10.5a). Similarly, the amplitude for a crosscap with insertion of massless fields can be recovered from the factorization limit of diagrams in figure 10.5b,c. These diagrams, as we discuss later on, correspond to a Moebius strip and a Klein bottle.

Indeed computing a sum of these diagrams of closed strings propagating for some time $T'\ell$ between disks and crosscaps, as shown in figure 10.6a), and taking the factorization limit $T' \to \infty$ one recovers the expression for the square of the total RR tadpole. This is pictorially shown in fig 10.6, and holds very precisely in the explicit computation to be discussed later on.

These diagrams are most easily computed in the dual channel, where they reduce to traces over Hilbert spaces. The channel in figure 10.6 is recovered by performing a modular transformation, after which we may take the factorization limit. Let us consider the different surfaces

The annulus

The diagram with two disks is our old friend the annulus. It can be easily computed as an ampiltude for an open string to travel for some time $2T\ell$ and glue back to itself. Taking into account the trace over Chan-Paton indices, it reads

$$Z_A = N^2 \int_0^\infty \frac{dT}{2T} \mathrm{tr}_{\mathcal{H}_{\rm open}} e^{-2T\ell H_{\rm open}} \tag{10.27}$$

The trace is over open oriented string states (since it is the sum over worlsheets that implements the orientifold projection, we do not have to impose it explicitly). We have

$$\begin{aligned}
\mathrm{tr}_{\,\mathrm{mom.}} e^{-2\pi\alpha' T \sum_i p_i^2} &= (8\pi^2\alpha' T)^{-4} \\
\mathrm{tr}_{\,\mathrm{bos.}}\, e^{-2\pi T(N_B - E_0^B)} &= \eta(iT)^{-8} \\
\mathrm{tr}_{\,NS,GSO}\, e^{-2\pi T(N_F - E_0^F)} &= \frac{1}{2}\left(\mathrm{tr}_{\,NS} q^{N_F + E_0^F} + \mathrm{tr}_{\,NS}(q^{N_F+E_0^F}(-)^F\right) = \\
& \frac{1}{2}\eta^{-4}\left(\vartheta\begin{bmatrix}0\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^4\right) \\
\mathrm{tr}_{\,R,GSO}\, e^{-2\pi T(N_F - E_0^F)} &= \frac{1}{2}\left(\mathrm{tr}_{\,R} q^{N_F + E_0^F} + \mathrm{tr}_{\,R}(q^{N_F+E_0^F}(-)^F\right) = \\
& \frac{1}{2}\eta^{-4}\left(\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^4\right)
\end{aligned} \tag{10.28}$$

In total

$$Z(T) = \frac{1}{2}(8\pi^2\alpha'\tau_2)^{-4}\eta^{-8}\eta^{-4}\left(\vartheta\begin{bmatrix}0\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^4 - \vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^4 + \vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^4\right) \tag{10.29}$$

As shown in figure 10.7, in going to the dual channel we find a closed string propagating between two disks during a time $T'\ell$ with $T' = \frac{1}{2T}$. We should then replace $T = \frac{1}{2T'}$ in the above expression. To make the formula look like an amplitude in the dual channel we should perform a modular transformation. Leaving the details for a second version of these notes, the amplitude

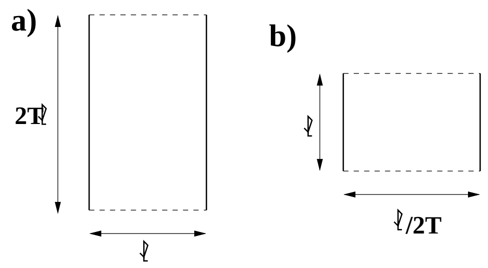


Figure 10.7: An open string propagating a time $2T\ell$ is geometrically the same as a closed string propagating a time $T'\ell$ with $T' = 1/(2T)$.

will read

$$Z_A = \int_0^\infty \frac{dT'}{2T'} \tilde{Z}_A(2T') \tag{10.30}$$

In this amplitude it is easy to identify the propagation of RR modes (upper characteristic of the theta function is 1/2). Taking the limit $T' \to \infty$ in this piece leads to

$$Z_A \to N^2 \tag{10.31}$$

This is proportional to the square of the RR disk tadpole.

Klein bottle

The Klein bottle amplitude corresponds to a closed string that propagates for a time $T\ell$ and is glued back to itself up to the action of Ω, see figure 10.1. The measure is obtained from that of the torus noticing that Ω does not allow for the τ_1 parameter. The amplitude hence reads

$$Z_K = \int_0^\infty \frac{dT}{4T} \operatorname{tr}_{\mathcal{H}_{\text{closed}}} e^{-T\ell H_{\text{closed}}} \tag{10.32}$$

The sum is over the Hilbert space of closed oriented strings. However, states non-invariant under Ω can be written as a sum over an Ω-even and an *Omega*-odd state

$$|A\rangle = \frac{1}{2}(|A\rangle + \Omega|A\rangle) + \frac{1}{2}(|A\rangle - \Omega|A\rangle) + \tag{10.33}$$

which have the same energy and different Ω eigenvalue. Hence their contributions cancel in the trace. Consequently, only states direcly mapped to themselves by Ω can contribute. Since this states are exactly left-right symmetric, we can simply sum over left-moving states and double the energy of each state. We obtain

$$\begin{aligned} Z_K(T) &= \mathrm{tr}_{\,\mathrm{mom.}} e^{-\pi\alpha' T \sum_i p_i^2}\, \mathrm{tr}_{\,\mathrm{bos.}}\, e^{-2\pi T(N_B - E_0^B)} \times \\ &\times \left(\mathrm{tr}_{\,NS,GSO}\, e^{-2\pi T(N_F - E_0^F)} - \mathrm{tr}_{\,R,GSO}\, e^{-2\pi T(N_F - E_0^F)} \right) \end{aligned} \tag{10.34}$$

The result is

$$Z(T) = \frac{1}{2}\,(4\pi^2\alpha' T)^{-4}\,\eta^{-8}(2iT)\,\eta^{-4}\left(\vartheta\begin{bmatrix}0\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^4 - \vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^4 + \vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^4\right)(2iT) \tag{10.35}$$

The Klein bottle is topologically the same surface as a closed string propagating between two crosscaps. This is shown in fig 10.8. In this dual closed channel the closed string propagates for a time $T'\ell$ with $T' = \frac{1}{4T}$. Replacing T in the amplitude and perform a modular transformation (for details, see a forthcoming second version of these notes), the amplitude will read

$$Z_K = \int_0^\infty \frac{dT'}{2T'}\,\tilde{Z}_K(2T') \tag{10.36}$$

Extracting the contribution from RR modes and taking $T' \to \infty$ in leads to

$$Z_K \to (32)^2 \tag{10.37}$$

This is proportional to the square of the RR crosscap tadpole, with same proportionality as in (10.31).

Moebius strip

The Moebius strip corresponds to an aa open string propagating from a time $2T\ell$ and glueing back to itself up to the action of Ω. This kind of diagram does not exist for ab states with $a \neq b$. The amplitude reads

$$Z_M = \pm N \int_0^\infty \frac{dT}{2T}\,\mathrm{tr}_{\,\mathcal{H}_{\mathrm{open}}}\left(e^{-2T\ell H_{\mathrm{open}}}\,\Omega\right) \tag{10.38}$$

The sign is given by the action of Ω on aa states, it can also be written $\mathrm{tr}\,(\gamma_\Omega^{-1}\gamma_\Omega^T)$ and is $+, -$ for cases i), ii) above.

The trace is over open oriented string states. However, in analogy with the Klein bottle, only states directly invariant under Ω contribute to the trace.

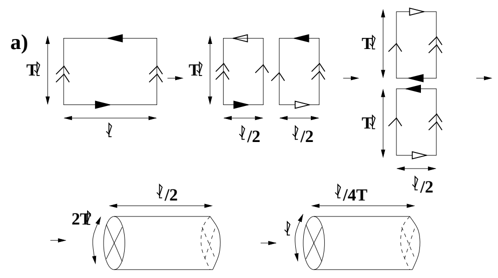


Figure 10.8: Take a Klein bottle as a rectangle with sides identified; cut it in two pieces keeping track of how they were glued; then glue explicitly some of the original identified sides. The result is the same surface now displayed as a surface with two crosscaps.

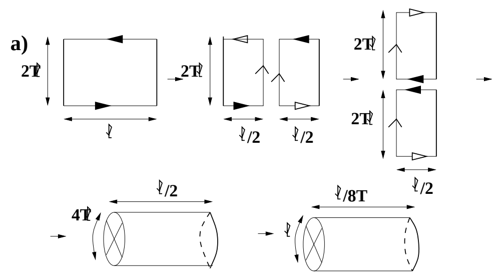


Figure 10.9: Take a Moebius strip as a rectangle with sides identified; cut it in two pieces keeping track of how they were glued; then glue explicitly some of the original identified sides. The result is the same surface now displayed as a surface with one boundary and one crosscap.

The explicit evaluation of this amplitude is easy, but involves slightly more complicated combinations of theta functions than the previous ones. We leave the details for a second version of these notes, and proceed the discussion in a qualitative way.

As shown in figure 10.9, the Moebius strip is topologically the same surface as a closed string propagating between a disk and a crosscaps. In this dual closed channel the closed string propagates for a time $T'\ell$ with $T' = \frac{1}{8T}$. Replacing T in the amplitude and perform a modular transformation, the amplitude will read

$$Z_M = \int_0^\infty \frac{dT'}{2T'}\, \tilde{Z}_M(2T') \tag{10.39}$$

Extracting the contribution from RR modes and taking $T' \to \infty$ in leads to

$$Z_K \to \mp 32\, N \tag{10.40}$$

with $-, +$ corresponding to the cases i), ii) above. This is proportional to the product of the RR disk and crosscap tadpoles, with same proportionality as in (10.31).

RR tadpole cancellation

The sum of the four amplitudes in fig 10.6a in the factorization limit is hence proportional to $(N-32)^2$. This implies that to obtain a consistent theory of unoriented open and closed strings, we need the Chan-Paton indices to run over 32 possible values

$$N = 32 \tag{10.41}$$

and the Ω action on them, γ_Ω, to be a symmetric matrix. This is type I superstring theory.

The spectrum of this theory is obtained straightforwardly. At the massless level the closed string sector corresponds to the 10d $\mathcal{N}=1$ supergravity multiplet, and the open string sector corresponds to 10d $\mathcal{N}=1$ vector multiplets with gauge group $SO(32)$.

Sector	Sector	$SO(8)$	Field
Closed	NS-NS	$1+35_V$	dilaton, graviton
	NS-R+R-NS	8_S+56_S	gravitino
	R-R	28_C	2-form
Open	NS	8_V	$SO(32)$ gauge boson
	R	8_C	gauginos

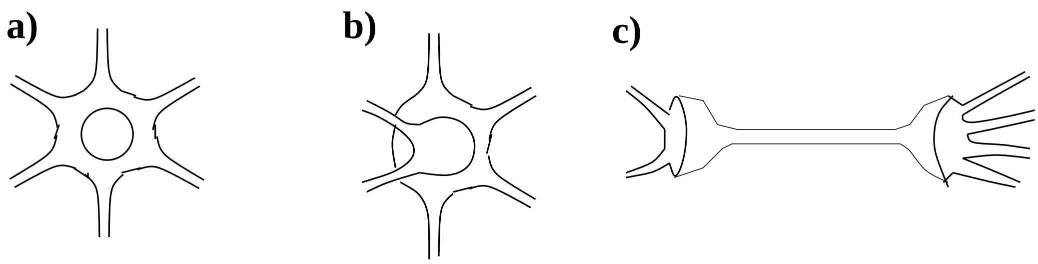


Figure 10.10: Limits of the annulus leading to anomalies in type I theory; a) corresponds to the familiar planar hexagon contribution to irreducible anomalies in field theory, while b) corresponds to a non-planar hexagon field theory contribution anomalies. c) corresponds to a Green-Schwarz diagram exchanging the closed string 2-form field, and which contributes to reducible anomalies.

Notice that this spectrum if free of gravitational and gauge anomalies. For this to be true, it is crucial that the gauge group is $SO(32)$, as we already saw in the discussion of anomalies in the heterotic theories. (interestingly enough, the massless spectrum of the $SO(32)$ and the type I string theories are the same).

In the cancellation of mixed gauge - gravitational anomalies, it is crucial the existence of a Green-Schwarz mechanism. Although at the level of the effective action the description for type I is similar to the one for heterotic (with the difference that the 2-form mediating the interaction is the RR one in type I theory), the string theory origin of the relevant couplings is different. In particular, both the BF^2 and BF^4 terms in type I string theory arise from disk diagrams with open string state insertions (powers of F) and a closed string B-field insertion, see figure 10.10.

10.4 Final comments

Just as with the other superstrings, there exist non-supersymmetric versions of type I superstring. One posibility is to construct orientifold quotients of the type 0 superstrings. We will not discuss these theories in our lectures. Another possibility [?] is to perform a modified Ω projection of type IIB theory which breaks the supersymmetries. We may discuss this theory later

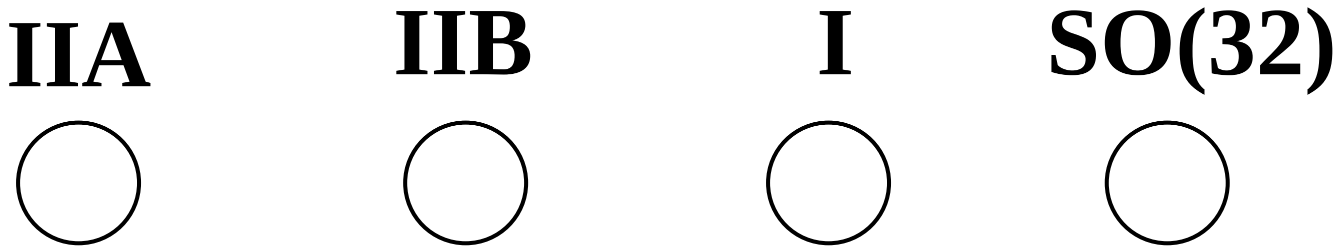


Figure 10.11: .

on in these lectures, since it will be easier to describe it once we learn about D-branes, orientifold planes and antibranes.

This concludes our discussion of the 10d superstring theories. At the moment the picture of string theory that we have is shown in fig 10.11. Five different (spacetime supersymmetric) superstring theories, constructed in different ways and with different features. All of them provide theories which describe gravitational (plus other) interactions in a quantum mechanically consistent way. However this multiplicity is unappealing: we would like to have a more unified description of how to construct consistent theories of gravitational interactions.

In the following lectures we will see that this picture will be drastically modified once we learn about compactification, T-duality and non-perturbative dualities. It turns out that the seemingly different string theories are intimately related, and seem to be just different limits of a unique underlying theory.

It would be very nice if the non supersymmetric strings would also fit into this unified picture. Although there are some ideas in the market, it is much more difficult to find evidence for this proposal.

Chapter 11

Toroidal compactification of superstrings

11.1 Motivation

In this lecture we study toroidal compactification of the (spacetime supersymmetric) superstring theories. The main motivation is to obtain theories which reduce to 4d at low energies. Although the models obtained in this lecture are not interesting to describe the real world (they are non-chiral), they will be useful starting points for further constructions, like orbifolds. Also, toroidal compactification illustrates, just as in bosonic theory, the very striking features of stringy physics. For instance, the phenomenon of T-duality in will reveal that the seemingly different superstring theories are related upon toroidal compactification.

11.2 Type II superstrings

In this discussion we follow section 13.1 of [55].

11.2.1 Circle compactification

Let us consider the type IIA, IIB theories compactified to 9d on a circle $\mathbf{S^1}$ of radius R. The 2d fermion sector is completely unchanged by the compactification; the only effects of the compactification are

i) the possibility of boundary conditions with non-zero winding w for the 2d bosonic fields, namely

$$X^9(\sigma+\ell,t) \,=\, X^9(\sigma,t) \,+\, 2\pi R\, w \tag{11.1}$$

ii) the fact that momentum along x^9 is quantized, $p_9 = k/R$.

In a sector of momentum k and winding w, we have the mode expansion

$$\begin{aligned} X_L(\sigma+t) &= \frac{x_0^9}{2} + \frac{p_{L,9}}{2p^+}(t+\sigma) + \frac{1}{\alpha' p^+}\, N_B \\ X_R(\sigma-t) &= \frac{x_0^9}{2} + \frac{p_{R,9}}{2p^+}(t-\sigma) + \frac{1}{\alpha' p^+}\, \tilde{N}_B \end{aligned} \tag{11.2}$$

with

$$p_L = \frac{k}{R} + \frac{wR}{\alpha'} \quad ; \quad p_R = \frac{k}{R} - \frac{wR}{\alpha'} \tag{11.3}$$

We have the spacetime mass formulae

$$\begin{aligned} M_L^2 &= \frac{p_L^2}{2} + \frac{2}{\alpha'}\,(N_B+N_F+E_0) \\ M_R^2 &= \frac{p_R^2}{2} + \frac{2}{\alpha'}\,(\tilde{N}_B+\tilde{N}_F+\tilde{E}_0) \end{aligned} \tag{11.4}$$

From these expressions we can obtain the spectrum of 9d states at any radius R. For a generic R, the only massless states are in the sector of $k=0$, $w=0$. These states correspond to the zero modes (zero internal momentum) of the KK reduction of the effective field theory of 10d massless modes. Note that these states are present in field theory because they have zero winding.

The proccess of KK reduction to 9d and keeping just the zero mode amounts to simply decomposing the representations with respect to the 10d $SO(8)$ group into representations of the 9d $SO(7)$ group. Working first with e.g. the purely left moving sector, at the massless level we have

Sector	State	$SO(8)$	$SO(7)$
NS	$\psi^i_{-1/2}\vert 0\rangle$	8_V	$7+1$
R	$(\pm\frac{1}{2},\pm\frac{1}{2},\pm\frac{1}{2},\pm\frac{1}{2})$	8_S	8
		8_C	8

Notice that the chiral 10d spinors of different chirality reduce to the same spinor representation of $SO(7)$, which does not have chiral representations (there is no chirality in odd dimensions).

In order to glue left and right movers, we may tensor the $SO(8)$ representions for left and right movers to get the 10d fields, and then decompose with respect to $SO(7)$, or decompose the left and right states with respect to $SO(7)$ representations and then tensor them. Both methods give the same result, so we may use any of them at will.

For type IIB theory, the massless 10d fields are the metric, 2-form and dilaton, G, B, ϕ; two gravitinos and two spin 1/2 fields $\psi_{\mu\alpha}$, ψ_α; the scalar axion, a 2-form and a self-dual 4-form, a, $\tilde{B}$, A_4^+. We have the following set of 9d massless states (See table 35 [124] for tensor products in $SO(7)$):

$$
\begin{array}{cccccccccc}
\text{NS-NS} & & & & & & & & & \\
8_V,\, 8_V & \rightarrow & 8_V \times 8_V = & 35_V & + & 28_V & + & 1 & & \\
\downarrow & & & \downarrow & & \downarrow & & \downarrow & & \\
7+1,\, 7+1 & \rightarrow & 7 \times 7 = & 27 & + & 21 & + & 1 & & G_{\mu\nu},\, B_{\mu\nu},\, \phi \\
 & & 7 \times 1 + 1 \times 7 = & 7 & + & 7 & & & & G_{9\mu},\, B_{9\mu} \\
 & & 1 \times 1 & 1 & & & & & & G_{99} \\
\text{R-NS} & & & & & & & & & \\
8_C,\, 8_V & \rightarrow & 8_C \times 8_V = & 56_S & + & 8_S & & & & \\
\downarrow & & & \downarrow & & \downarrow & & & & \\
8,\, 7+1 & \rightarrow & 8 \times 7 = & 48+8 & & \downarrow & & & & \psi_{\mu\alpha},\, \psi_{9\alpha} \\
 & \rightarrow & 8 \times 1 = & & & 8 & & & & \psi_\alpha \\
\text{NS-R} & & & & & & & & & \\
8_V,\, 8_C & \rightarrow & 8_V \times 8_C = & 56_S & + & 8_S & & & & \\
\downarrow & & & \downarrow & & \downarrow & & & & \\
7+1,\, 8 & \rightarrow & 7 \times 8 = & 48+8 & & \downarrow & & & & \psi_{\mu\alpha},\, \psi_{9\alpha} \\
 & \rightarrow & 1 \times 8 = & & & 8 & & & & \psi_\alpha \\
\text{R-R} & & & & & & & & & \\
8_C,\, 8_C & \rightarrow & 8_C \times 8_C = & 1 & + & 28_C & + & 35_C & & \\
\downarrow & & & \downarrow & & \downarrow & & \downarrow & & \\
8,\, 8 & \rightarrow & 8 \times 8 = & 1 & + & 7+21 & + & 35 & & a,\, \tilde{B}_{9\mu},\, \tilde{B}_{\mu\nu},\, A_{9\mu\nu\rho}
\end{array}
$$

Here $\mu = 2, \ldots, 8$ runs in the seven non-compact directions transverse to the light-cone.

For type IIA theory, the massless 10d fields are the metric, 2-form and dilaton, G, B, ϕ; two gravitinos and two spin 1/2 fields $\psi_{\mu\alpha}$, ψ_{α}, $\psi_{\mu\dot{\alpha}}$, $\psi_{\dot{\alpha}}$; a 1-form and a 3-form A_1 C_3. We have the following set of 9d massless states (See table 35 [124] for tensor products in $SO(7)$):

$$
\begin{array}{ccccccccc}
\text{NS-NS} & & & & & & & & \\
8_V,\, 8_V & \to & 8_V \times 8_V = & 35_V & + & 28_V & + & 1 & \\
\downarrow & & & \downarrow & & \downarrow & & \downarrow & \\
7+1,\, 7+1 & \to & 7 \times 7 = & 27 & + & 21 & + & 1 & G_{\mu\nu},\, B_{\mu\nu},\, \phi \\
 & & 7\times 1 + 1 \times 7 = & 7 & + & 7 & & & G_{9\mu},\, B_{9\mu} \\
 & & 1 \times 1 & 1 & & & & & G_{99} \\
\text{R-NS} & & & & & & & & \\
8_C,\, 8_V & \to & 8_C \times 8_V = & 56_S & + & 8_S & & & \\
\downarrow & & & \downarrow & & \downarrow & & & \\
8,\, 7+1 & \to & 8 \times 7 = & 48+8 & & \downarrow & & & \psi_{\mu\alpha},\, \psi_{9\alpha} \\
 & \to & 8 \times 1 = & & & 8 & & & \psi_{\alpha} \\
\text{NS-R} & & & & & & & & \\
8_V,\, 8_S & \to & 8_V \times 8_S = & 56_C & + & 8_C & & & \\
\downarrow & & & \downarrow & & \downarrow & & & \\
7+1,\, 8 & \to & 7 \times 8 = & 48+8 & & \downarrow & & & \psi_{\mu\alpha},\, \psi_{9\alpha} \\
 & \to & 1 \times 8 = & & & 8 & & & \psi_{\alpha} \\
\text{R-R} & & & & & & & & \\
8_C,\, 8_S & \to & 8_C \times 8_S = & 8_V & + & 56_V & & & \\
\downarrow & & & \downarrow & & \downarrow & & & \\
8,\, 8 & \to & 8 \times 8 = & 1+7 & & 21+35 & & & A_9,\, A_\mu,\, C_{9\mu\nu},\, C_{\mu\nu\rho}
\end{array}
$$

Several observations are in order:

• Notice that there is one additional scalar besides G_{99} (which defines the compactification radius), namely A_9. It would be interesting to describe the compactification for an arbitrary background of this field. Unfortunately, it is not known how to couple RR fields to the worldsheet 2d theory, so we will be unable to do this. In later sections, in the compactification of several dimensions, there appear additional scalars arising from the NS-NS sector. For these it is known how to couple the background to the 2d theory, and the latter is exactly solvable (still a free theory), so we will be able to describe the compactification in a general background of these fields, in the complete string theory.

• Notice that both type II theories lead to the same 9d massless spectrum. In particular, notice that chirality of type IIB theory is lost in toroidal compactification, since there is no chirality in 9d. Notice also that the origin of the 9d fields in the RR sector is very different from the 10d viewpoint in the IIA and IIB theories. The low energy effective theory for the massless modes in either case is described by 9d supergravity with 32 supercharges (which is a unique theory).

• The generalization to compactification to lower dimensions (here we refer to square tori, and trivial B-field background, see later for non-triival cases) is very easy. At the massless level, one simply decomposes the representations 8_V, 8_S, 8_C with respect to the surviving Lorentz group, and then tensors them together. In particular it is possible to see that compactification to 4d on $\mathbf{T^6}$ leads to the field content of $\mathcal{N}=8$ 4d supergravity. Notice that again this theory is non-chiral, so it is not useful to describe the real world. The large amound of susy in lower dimensions is related to the fact that compactification on tori does not break any supersymmetry. This will motivate to discuss more involved compactifications in later sections (e.g. Calabi-Yau compactification).

• There is no point (besides $R=0$ or $R=\infty$) at which states become light. At $R\to\infty$ we have a tower of states of zero winding and arbitrary momentum which become very light. This corresponds to the decompactification limit of the theory. As $R\to 0$ we instead have a tower of states of zero momentum and arbitrary winding which become light. It is natural to think that this corresponds to the decompactification limit of a dual theory, where momentum is the original winding, etc, just as in the bosonic string theory. We study this in next section

11.2.2 T-duality for type II theories

Recall from the bosonic theory that the effect of T-duality is to relate a theory compactified on a circle of radius R with a theory compactified on a circle of radius $R'=\alpha'/R$, in such a way that states of momentum, winding (k,w) are mapped to states of momentum, winding $(k',w')=(w,k)$. Equivalently, starting with a 2d field theory of left- and right-moving bosons $X_L(\sigma+t)$, $X_R(\sigma-t)$, with a spacetime geometry spanned by $X(\sigma,t)=X_L+X_R$, T-duality related it to a theory on a spacetime geometry spanned by $X'^9(\sigma,t)=X_L^9-X_R^9$, $X^\mu(\sigma,t)=X_L^\mu+X_R^\mu$.

In type II theory we also have the 2d fermions. In order to be consistent with 2d susy, we require that the T-dual theory is described also by the fermion field $\psi'^9(\sigma,t)=\psi_L^9(\sigma+t)-\psi_R^9(\sigma-t)$.

Hence, T-duality acts as spacetime parity on the right-movers. It is then intuitive that at the level of the spacetime spectrum, it will flip the chirality of the R groundstate, exchanging $8_C \leftrightarrow 8_S$. Namely, it flips the GSO projection on the right movers. Hence, starting with type IIB theory compactified on radius R the T-dual will describe type IIA theory compactfied on radius $R' = \alpha'/R$. This is T-duality for type II theories. Notice that it implies that the spectrum of massless fields at *generic* radius must be the same for both theories; the full spectrum is the same only for R, R' related by the T-duality relation.

The flip in the GSO projection can be derived more explicitly as follows. Recall that to build the R groundstate one forms the linear combinations of fermion zero modes

$$A_a^{\pm} = \psi_0^{2a} \pm i\psi_0^{2a+1} \tag{11.5}$$

So T-duality acts as $A_4^{\pm} \leftrightarrow A_4'^{\mp}$. In the original theory, one defines a state $|0\rangle$ satisfying $A_a^-|0\rangle = 0$ and the states surviving the GSO are e.g.

$$|0\rangle \quad , \quad A_{a_1}^+ A_{a_2}^+|0\rangle \quad , \quad A_1^+ A_2^+ A_3^+ A_4^+|0\rangle \tag{11.6}$$

In the T-dual theory, one would define a state $|0\rangle'$ by ${A_a}'^-|0\rangle' = 0$. In terms of the original operators we have $A_a^-|0\rangle' = 0$ for $a = 1, 2, 3$ and $A_4^+|0\rangle' = 0$. Hence we have

$$|0\rangle' = A_4^+|0\rangle \tag{11.7}$$

This implies that the $(-1)^F$ eigenvalue of $|0\rangle'$ is opposite to that of $|0\rangle$. This implies the GSO projection is opposite in the T-dual. Indeed, the surviving states (11.6) read, in the T-dual

$$A_a^+|0\rangle \quad , \quad A_{a_1}^+ A_{a_2}^+ A_{a_3}^+|0\rangle \tag{11.8}$$

From the viewpoint of the T-dual theory, we are choosing the opposite GSO projection.

It is easy to check the effect that T-duality has on the 10d fields, by comparing the 9d spectra. For instance, for bosonic fields

$$\begin{array}{rcl} IIA & \stackrel{T}{\leftrightarrow} & IIB \\ G_{\mu\nu}, B_{\mu\nu} & \leftrightarrow & B_{\mu\nu}, G_{\mu\nu} \\ A_9, A_\mu & \leftrightarrow & a, \tilde{B}_{9\mu} \\ C_{9\mu\nu}, C_{\mu\nu\rho} & \leftrightarrow & \tilde{B}_{\mu\nu}, A_{9\mu\nu\rho} \end{array} \tag{11.9}$$

The beautiful conclusion of T-duality is that IIA and IIB theories are much more intimately related than expected. In fact, they can be regarded as different limits of a unique theory, namely type II compactificaion on $\mathbf{S^1}$ in the limits of $R \to 0$ and $R \to \infty$.

11.2.3 Compactification of several dimensions

In this section we study compactification on a d-dimentional torus $\mathbf{T^d}$. These compactified theories contain more additional scalar fields, which correspond to 10d fields with some internal indices. Hence the vaccuum expectation value of these scalars correspond to specifying the backgrounds for the metric and other fields in the internal manifold.

We are interested in studying the set of possible toroidal compactifications, that is, the set of vevs that these scalar fields can acquire. This is called the moduli space of (toroidal) compactification. Unfortunately, it is not known how to quantize the 2d theory exactly if backgrounds for RR fields are turned on. So we will restrict to turning on backgrounds for the NS-NS fields, namely the metric and 2-form [1]

We describe $\mathbf{T^d}$ by periodic coordinates $x^i \simeq x^i + 2\pi R$, and define its geometry by a constant metric tensor G_{ij}. We also introduce a background for the 2-form, B_{ij}, which must be constant so as not to induce cost in energy (for constant B, its field strength vanishes).

The light-cone gauge-fixed action for an arbitrary metric background reads (see equation after (27) in lecture on quantization of closed bosonic string)

$$L_G = -p^+\partial_t x^-(t) + \frac{1}{4\pi\alpha'}\int_0^\infty d\sigma\, G_{ij}\,(\partial_t X^i\partial_t X^j - \partial_\sigma X^i\partial_\sigma X^j\,) \quad (11.10)$$

where we have used $p^+ = \frac{\ell}{2\pi\alpha'}g_{\sigma\sigma}$, and set $\ell = 2\pi\alpha' p^+$, so $g_{\sigma\sigma} = 1$.

To this we must add the term that describes the interaction of the string with the B-field, which reads

$$L_B = \frac{1}{4\pi\alpha'}\int_0^\infty d\sigma\, \epsilon^{ab}\, B_{ij}\,\partial_a X^i\partial_b X^j = \frac{1}{2\pi\alpha'}\int_0^\infty d\sigma\, B_{ij}\,\partial_t X^i\partial_\sigma X^j \quad (11.11)$$

[1]The moduli spaces including RR backgrounds can be studied in the supergravity approximation; we postpone this discussion to coming lectures, since the analysis is most useful to study non-perturbative properties of string theory.

In total we have

$$L \;=\; \frac{1}{2\pi}\int_0^\infty d\sigma\;\left[\frac{1}{2\alpha'}\,G_{ij}\,(\partial_t X^i\partial_t X^j\;-\;\partial_\sigma X^i\partial_\sigma X^j\,)\;+\;\frac{1}{\alpha'}\,B_{ij}\,\partial_t X^i\partial_\sigma X^j\right] \tag{11.12}$$

The presence of the backgrounds and the periodicity of the coordinates x^i do not modify the oscillator piece for the 2d bosons. We are already familiar with this fact for the metric background, from our experience with circle compactifications. For backgrounds of the B-field, this follows because the lagrangian term in L_B is a total derivative

$$\epsilon^{ab}\partial_a X^i\partial_b X^j B_{ij} = \partial_a(\epsilon^{ab}X^i\partial_b X^j B_{ij}) \tag{11.13}$$

so it is insensitive to the 2d local dynamics, and feels only the topology of the 2d field configuration (namely, the winding number).

Thus it is enough to work with the zero oscillator number piece in the mode expansion of the 2d bosons. In a sector of momenta and winding $k_i, w^j \in \mathbf{Z}$ this reads

$$X^i(\sigma,t)\;=\;x_0^i\;+\dot{x}^i\,t\;+\frac{2\pi R}{\ell}\,w^i\,\sigma \tag{11.14}$$

where $\dot{x}^i$ will be related to k_i below. Plugging this ansatz into the lagrangian, we get

$$L\;=\;\frac{\ell}{2\pi}\,\left[\frac{1}{2\alpha'}\,G_{ij}\,(\dot{x}^i\dot{x}^j\;-\;(\frac{2\pi R}{\ell})^2\,w^i\,w^j\,)\;+\;\frac{1}{\alpha'}\,B_{ij}\,\dot{x}^i\,\frac{2\pi R}{\ell}\,w^j\right] \tag{11.15}$$

The canonical momentum conjugate to x^i is

$$p_i\;=\;\frac{\partial L}{\partial \dot{x}^i}\;=\;\frac{\ell}{2\pi\alpha'}\,(G_{ij}\dot{x}^j + B_{ij}w^j\frac{2\pi R}{\ell}) \tag{11.16}$$

It is quantized in units of $1/R$, namely $p_i = k_i/R$. This leads to

$$\dot{x}^i\;=\;\frac{G^{ij}}{p^+}\,(\,\frac{k_j}{R}\;-\;\frac{R}{\alpha'}B_{jl}w^l\,) \tag{11.17}$$

and

$$X^i(\sigma,t)\;=\;x_0^i\;+\;\frac{G^{ij}}{p^+}\,(\,\frac{k_j}{R}\;-\;\frac{R}{\alpha'}B_{jl}w^l\,)\,t + \frac{R}{\alpha' p^+}\,w^i\sigma \tag{11.18}$$

Splitting between the left and right movers, we have

$$\begin{aligned} X^i_L(\sigma + t) &= \frac{x^i_0}{2} + \frac{p_L}{2p^+}(t+\sigma) \\ X^i_R(\sigma - t) &= \frac{x^i_0}{2} + \frac{p_R}{2p^+}(t-\sigma) \end{aligned} \tag{11.19}$$

with

$$\begin{aligned} p_{L,i} &= \frac{k_i}{R} + \frac{R}{\alpha'}(G_{ij} - B_{ij})\, w^j \\ p_{R,i} &= \frac{k_i}{R} + \frac{R}{\alpha'}(-G_{ij} - B_{ij})\, w^j \end{aligned} \tag{11.20}$$

and mass formulae read

$$\begin{aligned} M_L^2 &= \frac{2}{\alpha'}(N_B + N_F + E_0) + \frac{p_L^2}{2} \\ M_R^2 &= \frac{2}{\alpha'}(\tilde{N}_B + \tilde{N}_F + \tilde{E}_0) + \frac{p_R^2}{2} \end{aligned} \tag{11.21}$$

Narain lattice

The 2d-dimensional lattice of momenta (p_L, p_R) has two very special properties. It is even with respect to the Lorentzian (d, d) signature scalar product

$$(p_L, p_R) \cdot (p'_L, p'_R) = \alpha' \sum_i (p^i_L p_{L,i} - p^i_R p_{R,i}) = 2 \sum_i (k^i w'_i + w^i k'_i) \in \mathbf{Z} \tag{11.22}$$

and it is self-dual. These two properties ensure that the 1-loop partition function for the theory is modular invariant. Namely, the partition function has roughly speaking the structure

$$Z(\tau) = \ldots \sum_{(k,w)} q^{\alpha' p_L^2/2}\, \overline{q}^{\alpha' p_R^2/2} = \ldots \sum_{(p_L, p_R)} q^{\alpha' p_L^2/2}\, \overline{q}^{\alpha' p_R^2/2} \tag{11.23}$$

It is easy to see that the even and self-duality properties ensure that this is invariant under $\tau \to \tau + 1$ and $\tau \to -1/\tau$, resp. So each choice of background fields determines a (lorentzian) even and self-dual lattice of momenta (p_L, p_R). This is the so-called Narain lattice.

Conversely, any choice of ((d, d) lorentzian) even and self-dual lattice $\Gamma_{d,d}$ can be used to *define* a consisten modular invariant toroidal compactification

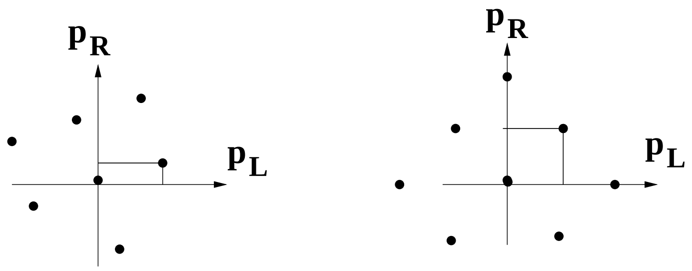


Figure 11.1: Heuristic picture of the relation between lattices and physical theories. Although the two lattices are related by a rotation in 2d spcace, the physics is sensitive to the independent values of p_L and p_R, and therefore not invariant under the rotation. Physics is not invariant under the mathematical isomorphism that relates the two lattices. The rotation parameters encode the background fields.

of type II theory, by simply using the vectors in the lattice to provide the sectors of momenta (p_L, p_R) in the theory.

This description, first introduced by Narain [58] in the heterotic context, is useful to provide a complete classification of all possible toroidal compactifications (which correspond to free worldsheet theories). Hence they allow to compute the moduli space of such compactifications, as follows.

A general theorem in mathematics states that all possible (d, d) lorentzian even self-dual lattices are isomorphic, namely any two such lattices differ by an $SO(d, d)$ rotation. This does not mean that there is a unique physical compactification, because the physics is not invariant under arbitrary $SO(d, d)$ transformations. In particular, the spacetime mass of a state with momenta (p_L, p_R) depends on $p_L^2 + p_R^2$, which is only $SO(d) \times SO(d)$ invariant. This is illustrated in figure 11.1. Hence, physically different theories are classified by elements in the coset $SO(d, d)/[SO(d) \times SO(d)]$. This is (almost, see below) the moduli space of compactifications. Note that it has dimension d^2.

It would be interesting to be able to provide an interpretation of a compactification defined by these abstract lattices, in terms of background fields as those introduced above. In fact, the number of background fields is also $d(d + 1)/2$ (for G_{ij}) plus $d(d - 1)/2$ (for B_{ij}), namely a total of d^2. This suggests that any abstract lattice corresponds to a particular choice of background fields.

In fact we can be even more specific: The background fields themselves are the rotation parameters in $SO(d, d)/[SO(d) \times SO(d)]$. For instance, it is easy to show that the lattice of momenta for generic B_{ij}

$$(p_{L,i}, p_{R,i}) \;=\; (\,\frac{k_i}{R} \,+\, \frac{R}{\alpha'}\,(G_{ij}-B_{ij})\,w^j\,;\,\frac{k_i}{R} \,+\, \frac{R}{\alpha'}\,(-G_{ij}-B_{ij})\,w^j\,) \qquad (11.24)$$

are related to the lattice of momenta for $B_{ij}=0$

$$(p_{L,i}, p_{R,i}) \;=\; (\,\frac{k_i}{R} \,+\, \frac{R}{\alpha'}\,G_{ij}\,w^j\,;\,\frac{k_i}{R} \,-\, \frac{R}{\alpha'}\,G_{ij}\,w^j\,) \qquad (11.25)$$

by the rotation matrix

$$M_B \;=\; \begin{pmatrix} \delta_i^j - \frac{1}{2}B_i^j & \frac{1}{2}B_i^j \\ -\frac{1}{2}B_i^j & \delta_i^j + \frac{1}{2}B_i^j \end{pmatrix} \qquad (11.26)$$

which is in $SO(d,d)$ because $M_B \;=\; \exp\frac{1}{2}\begin{pmatrix} -B & B \\ -B & B \end{pmatrix}$. Similarly, the momenta for generic G_{ij} can be related to the momenta for cubic metric $G_{ij}=\delta_{ij}$ via an $SO(d,d)$ rotation

$$M_G \;=\; \begin{pmatrix} \cosh S & \sinh S \\ \sinh S & \cosh S \end{pmatrix} = \exp\frac{1}{2}\begin{pmatrix} 0 & S \\ S & 0 \end{pmatrix} \qquad (11.27)$$

where S_{ij} is a symmetric matrix.

From either viewpoint we reach the conclusion that the moduli space of compactifications with these backgrounds is $SO(d,d)/[SO(d)\times SO(d)]$. In fact, this statement needs some refinement. In the description in terms of abstract lattices, it is clear that there may exist some finite $SO(d,d)$ transformations, not in $SO(d)\times SO(d)$, which leave the lattice Γ invariant as a whole, although acting non-trivially on the individual points (p_L, p_R). Since the lattice defines the physics, we should mod out by those transformations. They correspond to rotation matrices with integer entries, and generate a group denoted $SO(d,d;\mathbf{Z})$. Therefore the complete moduli space is

$$SO(d,d)/[SO(d)\times SO(d)\times SO(d,d;\mathbf{Z})] \qquad (11.28)$$

These latter transformations act nontrivially on the winding and momentum quantum numbers, and also relate theories with different backgrounds. They include large diffeomeorphisms of $\mathbf{T^d}$, large gauge transformations of the B_{ij}, and also T-dualities (sign flips of right-moving momenta). For this reason, $SO(d,d;\mathbf{Z})$ is often called the T-duality group.

Some observations are in order

• States in the theory must form representations of the T-duality group: Since it leaves the theory invariant, there must be sets of states which are shuffled among themselved by the action of the symmetry. They thus lie in representations of the group. Representations of $SO(d,d;\mathbf{Z})$ are easy to construct from representations of $SO(d,d)$ by restriction. To give one example of this discussion, the d states $k_i = 1$, $w^j = 0$ and the d states $k_i = 0$, $w^j = 1$ form a $2d$-dimensional representaion of $SO(d,d;\mathbf{Z})$, which is the representation obtained from restriction of the vector representation of $SO(d,d)$.

• Again, we recall that toroidal compactifications contain more moduli than those discussed here. The inclusion of the additional backgrounds leads to large moduli spaces. They cannot be computed in full-fledged string theory, but can be computed in the supergravity approximation (which is reliable since the large amount of supersymmetry protects the structure of moduli space to a large extent).

• Finally, there will be enlarged duality groups, which act nontrivially on the states and on the backgrounds. A novelty, to be studied in later lectures, is that these enlarged duality groups act nontrivially on the string coupling, and therefore relate weakly coupled and strongly coupled regimes of string theory. The corresponding duality multiplets therefore contains perturbative string states (such as strings with momentum and winding) and non-perturbative states (the so-called branes) Hence dualities provide an extremely useful tool to study non-perturbative phenomena in string theory.

11.3 Heterotic superstrings

In the discussion we follow section 11.6 of [71]

11.3.1 Circle compactification without Wilson lines

This is the simplest compactification, although not the most generic one (additional background fields, Wilson lines, are turned on in later sections). We simply take spacetime to be $M_9 \times \mathbf{S^1}$ (so we make one coordinate periodic, $x^9 \simeq x^9 + 2\pi R$) and turn on no background for the 10d gauge fields. As usual, the compactification only modifies the theory by the inclusion of winding sectors, and the restriction to quantized momenta in the compact direction. Therefore, different sector of the theory will be labelled by left and right moving momenta

$$p_{L,R} = \frac{k}{R} \pm \frac{R}{\alpha'} w \tag{11.29}$$

as well as the internal 16d lattice left moving momenta P^I in the $E_8 \times E_8$ or $Spin(32)/\mathbf{Z_2}$ lattices. Defining the internal left moving 16d dimensionful momenta $P_L = \sqrt{2/\alpha'}\, P$, the mass formulae are given by

$$\begin{aligned} M_L^2 &= \frac{P_L^2}{2} + \frac{p_L^2}{2} + \frac{2}{\alpha'} (N_B - 1) \\ M_R^2 &= \frac{p_R^2}{2} + \frac{2}{\alpha'} (\tilde{N}_B + \tilde{N}_F + \tilde{E}_0) \end{aligned} \tag{11.30}$$

The spectrum of massless states at a generic radius (in particular at large radius) is the $p_L = p_R = 0$ sector. This corresponds to $k = w = 0$, hence we recover the zero modes of the (field theory) KK reduction from 10d to 9d. States are just the group theory decomposition of the massless states in 10d. We have

$$
\begin{array}{ccccccccc}
\text{NS} & & & & & & & & \\
8_V,\ \alpha^i_{-1}|0\rangle & \rightarrow & 8_V\times 8_V = & 35_V & + & 28_V & + & 1 & \\
\downarrow & & & \downarrow & & \downarrow & & \downarrow & \\
7+1,\ 7+1 & \rightarrow & 7\times 7 = & 27 & + & 21 & + & 1 & G_{\mu\nu},\ B_{\mu\nu},\ \phi \\
 & & 7\times 1 + 1\times 7 = & 7 & + & 7 & & & G_{9\mu},\ B_{9\mu} \\
 & & 1\times 1 & 1 & & & & & G_{99} \\
\text{R} & & & & & & & & \\
8_C,\ \alpha^i_{-1}|0\rangle & \rightarrow & 8_C\times 8_V = & 56_S & + & 8_S & & & \\
\downarrow & & & \downarrow & & \downarrow & & & \\
8,\ 7+1 & \rightarrow & 8\times 7 = & 48+8 & & \downarrow & & & \psi_{\mu\alpha},\ \psi_{9\alpha} \\
 & \rightarrow & 8\times 1 = & & & 8 & & & \psi_\alpha \\
\text{NS} & & & & & & & & \\
8_V,\ \alpha^I_{-1}|0\rangle & \rightarrow & 8_V\times 1 = & 8_V & & & & & \\
\downarrow & & & \downarrow & & & & & \\
7+1,\ 1 & \rightarrow & & 7+1 & & & & & A^I_\mu,\ A^I_9 \\
8_V,\ |P_I\rangle_{P^2=2} & \rightarrow & 8_V\times 1 = & 8_V & & & & & \\
\downarrow & & & \downarrow & & & & & \\
7+1,\ 1 & \rightarrow & & 7+1 & & & & & A_{P,\mu},\ A_{P,9} \\
\text{R} & & & & & & & & \\
8_C,\ \alpha^I_{-1}|0\rangle & \rightarrow & 8_C\times 1 = & 8_C & & & & & \\
\downarrow & & & \downarrow & & & & & \\
8,\ 1 & \rightarrow & & 8 & & & & & \psi^I \\
8_C,\ |P_I\rangle_{P^2=2} & \rightarrow & 8_C\times 1 = & 8_C & & & & & \\
\downarrow & & & \downarrow & & & & & \\
8,\ 1 & \rightarrow & & 8 & & & & & \psi_P
\end{array}
$$

The first set of states is the gravity multiplet of 9d supergravity with 16 supersymmetries. The second set of states are 9d vector supermultiplets with respect to 16 supersymmetries, namely 9d gauge bosons, gauginos and real scalars in the adjoint of the gauge group, which is $E8\times E_8$ or $SO(32)$. Hence the 10d gauge group from the internal lattice is unbroken. In addition, there is the usual $U(1)^2$ gauge group arising from the familiar KK mechanism from the 10d graviton and B-field.

The generalization to lower dimensions is very easy, one simply needs to decompose the fields with respect to representations of the corresponding Lorentz group. Notice that in any of these compactifications chirality is lost. In particular, compactifications to 4d lead to theories with 4d $\mathcal{N}=4$

supersymmetry, which are automatically non-chiral.

Notice that in the above construction (i.e. without gauge field backgrounds) the pattern of enhance gauge symmetries at special values of R is exactly like in bosonic string theory. That is, the generic $U(1)^2$ gauge symmetry from the graviton and B-field enhance to $SU(2)^2$ at $R = \sqrt{\alpha'}$. Notice that there are no values of R for which the enhancement of the group involves both the $U(1)^2$ and the original 10d group. This will be different when we include Wilson lines.

Finally, we would like to mention that the $E_8 \times E_8$ and $SO(32)$ heterotic theories are self-T-dual. The $E_8 \times E_8$ heterotic theory on a circle of radius R is equivalent (up to relabeling of k and w) to the $E_8 \times E_8$ heterotic theory on a circle of radius $R' = \alpha'/R$ (and similarly for the $SO(32)$ heterotic theory). This would suggest that the two heterotics are not as intimately related as type IIA and IIB theories. We will see that they are: if one considers the more general case of compactifications with Wilson lines, there are T-dualities relating compactifications of the two heterotic theories.

11.3.2 Compactification with Wilson lines

The compactifications discussed above are not the most general circle compactifications. Note that the resulting 9d theory had additional scalars besides G_{99}, namely the scalar fields A_9^a in the adjoint of the gauge group. A vev for these scalars corresponds to turning on backgrounds for the internal components of the gauge fields, the so-called Wilson lines. In this section we discuss Wilson lines, first in the context of field theory, then in the context of heterotic string theory.

11.3.3 Field theory description of Wilson lines

Consider the following toy model of compactification from 5d to 4d. Consider a gauge theory with gauge group G in a spacetime $\mathbf{M_4} \times \mathbf{S^1}$, with $\mathbf{S^1}$ parametrized by the periodic coordinate $x^4 \simeq x^4 + 2\pi R$.

We also turn on a constant backbround for the internal component of the gauge bosons A_4^a. Locally, this is pure gauge, namely it could be gauge away, but the gauge parameter would not be a single-valued in $\mathbf{S^1}$ and thus would not define a global function. For instance, for $G = U(1)$, the gauge background can be locally gauge away with a gauge transformation

$$A_\mu \to A_\mu + \partial_\mu \lambda \;\text{ with }\; \lambda = -\langle A_4 \rangle x^4 \tag{11.31}$$

and λ is not globally well defined on $\mathbf{S^1}$.

The gauge non-triviality of the gauge background can be encoded in the gauge-invariant object, called the Wilson line, defined by

$$W^a = \exp i \int_{\mathbf{S^1}} A^a = \exp(2i\pi R A_4^a) \tag{11.32}$$

Notice that A_4^a is periodic with period $1/R$. It is convenient to define $\tilde{A}_4^a = 2\pi R A_4^a$ which has period 1.

From the 4d viewpoint, the Wilson lines or gauge backgrounds of this kind are interpreted as giving a vacuum expectation value to the 4d fields A_4^a, which are 4d scalars transforming in the adjoint of the gauge group.

This makes it clear that, using global transformations in the gauge group, one can always diagonalize the hermitian matrix of vevs. This means that one can always rotate within the gauge group to a configuration where the gauge backgrounds are non-zero only for Cartan generators. We will denote the gauge background in this basis by A_4^I, with $I = 1, \ldots, \operatorname{rank} G$. This is a vector of Wilson line vevs.

We are interested in obtaining the spectrum of light 4d fields. To obtain them we should expand the 5d action around the background defining the compactification (namely, the circle geometry and the gauge background). The 5d action for the gauge fields roughly reads

$$S_{5d} \;=\; \int_{M_4 \times \mathbf{S^1}} \operatorname{tr} \mathcal{F}_{MN} \mathcal{F}^{MN} \tag{11.33}$$

with

$$\mathcal{F}_{MN} \;=\; \partial_{[M} \mathcal{A}_N + [\mathcal{A}_M, \mathcal{A}_N] \quad ; \quad \mathcal{A}_M = \sum_a A_M^a \, t^a \tag{11.34}$$

The terms $|[\mathcal{A}_M, \mathcal{A}_N]|^2$ in the compactification lead to 4d mass terms for gauge bosons $|[\mathcal{A}_\mu, \mathcal{A}_4]| \simeq tr(A_\mu^a A_\mu^a) W^2$ unless the generators associated with the gauge bosons commute with the generators associated with the gauge background. This is called the commutant of the subgroup where the gauge background was turned ont. To understand better which gauge bosons survive, we describe their generators in the Cartan-Weyl basis.

Gauge bosons of Cartan generators always have zero mass terms in 4d (since they always commute with the background, because it is embedded in Cartan generators as well). The rank of the 4d gauge group is the same as for the 5d group.

For non-Cartan generators, associated with some non-zero root α, the corresponding gauge boson survives in the massless sector if the commutator vanishes

$$[\langle H_I \rangle, E_\alpha] = \alpha_I A_4^I = 0 \tag{11.35}$$

Namely we obtain massless 4d gauge bosons for $\alpha \cdot A_4 = 0$. Recalling the periodicity in A_4^I, careful analysis leads to the slightly more relaxed $\alpha \cdot \tilde{A}_4 \in \mathbf{Z}$.

Recalling that the α_I are integer, and the A_4^I are continuous parameters, it is clear that generically the only surviving massless gauge bosons are the Cartan generators, generically the 4d group is broken to $U(1)^r$, with $r =$ rank G. For special choices of Wilson line (i.e. at particular points in Wilson line moduli space) we will obtain enhanced non-abelian gauge symmetries. For instance, for zero Wilson lines the 4d group equal to G. Turning on small wilson lines starting from a point of enhanced symmetry, breaks the gauge group. From the viewpoint of the 4d theory this is understood as a Higgs effect due to the scalars in the adjoint of the enhanced gauge group.

To give a simple example, consider $G = U(n)$, and consider that the Wilson line along x^4 corresponds to $(\mathcal{A}_4^I) = (0, \ldots, 0, a)$. For generic a, the only elements of $SU(n)$ that preserve the background (commute with the Cartan with Wilson line) are the $U(n-1)$ rotations in the first $n-1$ entries, times the total trace $U(1)$. The unbroken group is $U(n-1) \times U(1)$.

There is an alternative description of what fields remain massless in the 4d theory in the presence of Wilson lines, which is valid not just for gauge bosons but for any 5d field ψ charged under the 5d gauge group. Recalling that in a gauge theory all derivatives must be promoted to covariant derivatives, involving the gauge field, and that derivatives are related to momenta, it is clear that the natural momentum in the fifth direction x^4 is not associated to ∂_4, but to

$$\begin{aligned} D_4 \psi &= \partial \psi + q_I A_4^I \psi \\ P_4 &= (k + q_I \tilde{A}_4^I)/R \end{aligned} \tag{11.36}$$

with $k \in \mathbf{Z}$.

The 4d mass of the KK modes of this field is given by $m^2 = P_4^2$, for varying k. Clearly we obtain 4d massless fields only if $q \cdot \tilde{A}_4 \in \mathbf{Z}$. This generalizes the condition on gauge bosons, which is recovered by recalling that the roots α^I are simply the charges of the gauge bosons under the corresponding Cartan generator.

Before concluding, we would like to mention how this generalizes to compactification of several dimensions, i.e. $\mathbf{T^d}$ compactifications. In this case, we can turn on gauge backgrounds along any of the internal directions, A_i^a. Now in order to turn on this background without any cost in vacuum energy (so that we are still describing a vacuum of the theory) we have to avoid that backgrounds in different directions contribute to the energy via the commutators $[\mathcal{A}_i, \mathcal{A}_j]$ in the higher dimensional gauge kinetic term. This implies that backgrounds in the different direction commute among themselves. (From the viewpoint of the 4d theory, it implies a conditions on the corresponding scalar vev, which is condition of minimization of the scalar potential). On the other hand, it means that the corresponding matrices (in the gauge indices) can be simultaneously diagonlized, i.e. the complete background can be rotated to the Cartan generators. Therefore, the most general configuration of Wilson lines corresponds to backgrouns A_i^I for the Cartan generators. Clearly, the basic rule is that we obtain massless fields for states with charge vector q^I satisfying $q \cdot \tilde{A}_i \in \mathbf{Z}$, for any $i = 1, \ldots d$. Namely, each Wilson line acts independently.

In later sections we will see how this effective field theory description arises in string theory, at least in the limit of large radii.

11.3.4 String theory description

Narain lattice

In order to discuss compactification with Wilson lines in string theory, is to couple the gauge background to the 2d worldsheet theory. Happily, in the presence of constant gauge backgrounds the 2d theory is still free, and so exactly solvable. The gauge backgrounds A_i^I in a T^d compactification can be seen to couple e.g. to the 2d bosons through a term

$$S_A = \int d^2\xi \, \epsilon^{ab} \, \partial_a X^i \, \partial_b X^I A_i^I \tag{11.37}$$

The complete action is quadratic, a free theory.

The canonical quantization of the complete lagragian in the presence of backgrounds G_{ij}, B_{ij} and A_i^I is discussed in [59]. This is analogous to our study of type II compactification on $\mathbf{T^d}$, but there is a subtlety in that the 2d fields X^I are constrained to be purely left moving. The use of Dirac method of quantization of constrained systems implies a subtle additional piece in the canonical momenta. Skipping the details, the result for the left and right moving momenta in this compactifications are given by

$$\begin{aligned} P_L^I &= \sqrt{\frac{2}{\alpha'}}\,(\,P^I\,+\,RA_i^I\,w^i\,) \\ p_{L,i} &= \frac{k_i}{R}\,+\,\frac{R}{\alpha'}\,(G_{ij}-B_{ij})\,w^j\,-\,P^I A_i^I\,-\,\frac{R}{2}\,A_i^I A_j^I\,w^j \\ p_{R,i} &= \frac{k_i}{R}\,+\,\frac{R}{\alpha'}\,(-G_{ij}-B_{ij})\,w^j\,-\,P^I A_i^I\,-\,\frac{R}{2}\,A_i^I A_j^I\,w^j \end{aligned} \tag{11.38}$$

The formulae are given by

$$\begin{aligned} M_L^2 &= \frac{P_L^2}{2}\,+\,\frac{p_L^2}{2}\,+\,\frac{2}{\alpha'}\,(N_B-1) \\ M_R^2 &= \frac{p_R^2}{2}\,+\,\frac{2}{\alpha'}\,(\tilde{N}_B+\tilde{N}_F+\tilde{E}_0) \end{aligned} \tag{11.39}$$

The lattice of momenta (11.38) is even with respect to the Lorentzian scalar product $P_L^I P_L^{I\prime} + p_L^i p'_{L,i} - p_R^i p'_{R,i}$, and self-dual. This ensures that the partition function for these theories is modular invariant for any choice of background fields, so they define consistent vacua of the theory.

As in type II compactifications, we are interested in the structure of the set of vacua of these theories, namely the moduli space for the scalars in the compactified theory. Following Narain, any $\mathbf{T^d}$ compactification can be defined in terms of an abstract $(16+d,d)$ lorentzian even and self-dual lattice $\Gamma_{16+d,d}$ of momenta. Mathematical theorems ensure that (p,q) lorentzian even self-dual lattices exist iff $p-q$ is a multiple of 8, which is fortunately satisfied in our case. Also, for $d>1$ all $(16+d,d)$ even self-dual lattices are isomorphic, up to a rotation in $SO(16+d,d)$. Again, this does not mean that all physical compactifications are equivalent, because the physics (e.g. the mass formulae) is invariant only under $SO(16+d)\times SO(d)$. Therefore, the set of inequivalent $\mathbf{T^d}$ compactifications of the theory is the coset $SO(16+d,d)/[SO(16+d)\times SO(d)]$.

This space has dimension $(16+d)d$, so a vacuum of the compactified theory is defined by $(16+d)d$ parameters. In fact, this is the number of parameters that define a background configurations, namely d^2 from G_{ij}, B_{ij} and $16d$ from the Wilson lines A_i^I. In fact, it is possible to see that these background fields are indeed the $SO(16+d,d)$ rotation parameters. Namely, the momenta (11.38) for generic values of B_{ij}, A_i^I are related to those for $B_{ij=0}$, $A_i^I = 0$

$$\begin{aligned} P_L^I &= \sqrt{\frac{2}{\alpha'}}\, P^I \\ p_{L,i} &= \frac{k_i}{R} + \frac{R}{\alpha'} G_{ij}\, w^j \\ p_{R,i} &= \frac{k_i}{R} - \frac{R}{\alpha'} G_{ij} w^j \end{aligned} \tag{11.40}$$

by the matrix

$$M_{B,A} = \begin{pmatrix} \delta_J^i & \sqrt{\frac{2}{\alpha'}} A_J^i & -\sqrt{\frac{2}{\alpha'}} A_J^i \\ -\sqrt{\frac{2}{\alpha'}} A_j^I & \delta_j^i - \frac{1}{2} B_j^i - \frac{\alpha'}{4} A_j^I A^{I,i} & \frac{1}{2} B_j^i = \frac{\alpha'}{4} A_j^I A^{I,i} \\ -\sqrt{\frac{2}{\alpha'}} A_j^I & -\frac{1}{2} B_j^i - \frac{\alpha'}{4} A_j^I A^{I,i} & \delta_j^i + \frac{1}{2} B_j^i = \frac{\alpha'}{4} A_j^I A^{I,i} \end{pmatrix} \tag{11.41}$$

which is an $SO(16+d,d)$ rotation since

$$M_{B,A} = \begin{pmatrix} 0 & \sqrt{\frac{2}{\alpha'}} A_J^i & -\sqrt{\frac{2}{\alpha'}} A_J^i \\ -\sqrt{\frac{2}{\alpha'}} A_j^I & -B_j^i & B_j^i \\ -\sqrt{\frac{2}{\alpha'}} A_j^I & -B_j^i & B_j^i \end{pmatrix} \tag{11.42}$$

As in type II, the momenta for generic G_{ij} are related to those for cubic metric by a rotation

$$M_G = \begin{pmatrix} \cosh S & \sinh S \\ \sinh S & \cosh S \end{pmatrix} = \exp \frac{1}{2} \begin{pmatrix} 0 & S \\ S & 0 \end{pmatrix} \tag{11.43}$$

As in type II, we should be careful in constructing the moduli space, since there may exist finite $SO(16+d,d)$ transformations which leave a lattice of momenta invariant, although acting non-trivially on individual states. These transformations form the group $SO(16+d,d;\mathbf{Z})$ and corresponds to large diffeomorphisms of $\mathbf{T^d}$, shifts on B_{ij}, A_i^I by whole periods, and T-dualities. Since theories related by these rotations are physically equivalent, the moduli space has really the structure

$$SO(16+d,d)/[SO(16+d) \times SO(d) \times SO(16+d,d;\mathbf{Z})] \tag{11.44}$$

This result will be useful in the discussion of non-perturbative dualities in compactifications of heterotic theories, etc, in later lectures.

Spectrum

At generic R the spectrum of light states is easily computed. For instance we obtain massless states from the decomposition of $8_V \times \alpha_{-1}^i|0\rangle$ and $8_C \times \alpha_{-1}^i|0\rangle$, which lead to the 4d $\mathcal{N} = 4$ supergravity multiplet. Notice that it includes gauge bosons arising from the 10d metric and 2-form with one internal index.

We also get massless states from the decomposition of $(8_v + 8_C) \times \alpha_{-1}^I|0\rangle$, they correspond to 4d $\mathcal{N} = 4$ $U(1)^{16}$ vector multiplets. Finally, states with nonzero 16d momentum lead to massless states if $p_L = p_R = 0$, $P_L^2 = 4/\alpha'$. This can only be achieved in the $w^i = 0$ sector where

$$\begin{aligned} P_L^I &= \sqrt{\frac{2}{\alpha'}}\, P^I \\ p_{L,i} &= \frac{(k_i - P\cdot\tilde{A}_i}{R} \\ p_{R,i} &= \frac{(k_i - P\cdot\tilde{A}_i}{R} \end{aligned} \tag{11.45}$$

$$\tag{11.46}$$

So massless states correspond to $P^2 = 2$, $P \cdot \tilde{A}_i \in \mathbf{Z}$. This result, valid for generic R (and thus also for large R) reproduces the field theory analysis, as should be the case. These modes correspond to the KK reduction of the 10d $\mathcal{N} = 1$ vector multiplets in the presence of Wilson lines. For generic Wilson lines the non-abelian gauge bosons do not survive and the 4d gauge symmetry is simply $U(1)^{16}$.

On the other hand, by tunning some backgrounds, it is possible to achieve situations where some vector in the lattice of momenta satisfies

$$P_L^2 + p_L^2 = 4/\alpha' \tag{11.47}$$

leading to some enhancement of the gauge symmetry breaking due to states $(8_V + 8_C) \times |P_L, p_L\rangle$. One simple particular case is tunning the Wilson lines to zero.

Notice that in general the new massless states at enhances symmetry points involve non-zero spacetime winding and momentum. This means that they are charged under the $U(1)^{2d}$ gauge bosons arising from the 10d metric and 2-form, in addition to being charged under the $U(1)^{16}$ from the internal 16d 'space'. The complete non-abelian group gathers Cartan generators of very different origin in 10d language!. The general recipe is that any non-abelian (simply laced[2]) group of rank $\leq 16 + 2d$ can appear as the gauge group in a corner of moduli space of $\mathbf{T^d}$ compactifications.

As a final comment, let us mention that moving away from such points (of enhanced gauge symmetry) in moduli space corresponds to a Higgs effect from the viewpoint of the lower dimensional effective field theory. This is similar to what we saw for the bosonic theory.

T-duality of $E_8 \times E_8$ and $SO(32)$ circle compactification

The fact that the moduli space of e.g. $\mathbf{S^1}$ compactifications of heterotic string theory is connected implies that a single theory in 9d can receive two interpretations, as compactification of $E_8 \times E_8$ heterotic on a radius R with Wilson lines A_i^I, and as compactification of $SO(32)$ heterotic on a different radius R' with different Wilson lines $A_i^I{}'$. Both compactifications are physically equivalent, although look different in 10d language. They are hence related by T-duality transformation. In this section we study the simplest example of these T-dualities (we follow section 11.6 of [71]).

Consider compactification of $E_8 \times E_8$ and $SO(32)$ heterotic theories on $\mathbf{S^1}$'s of radii R and R' repectively, with $G_{99} = 1$, $G'_{99} = 1$. The momenta lattice read

$$\begin{aligned} P_L^I &= \sqrt{\frac{2}{\alpha'}}\,(\,P^I \,+\, RA^I\,w^i\,) \\ p_{L,R} &= \frac{k}{R} \pm \frac{R}{\alpha'}\,w \,-\, P\cdot A \,-\, \frac{R}{2}\,A\cdot A\,w \end{aligned} \tag{11.48}$$

and similarly for primed parameters. Consider the choice of Wilson lines

$$\begin{aligned} (\tilde{A}^I) &= (\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2};0,0,0,0,0,0,0,0) \\ (\tilde{A}^I{}') &= (1,0,0,0,0,0,0,0;1,0,0,0,0,0,0,0) \end{aligned} \tag{11.49}$$

for $E_8 \times E_8$ and $SO(32)$, resp.

The T-duality is the statement that these two theories are equivalent if $R = \alpha'/(2R)$. To show this one would have to see that the two Narain lattices are exactly the same. This can be done [60], but is a bit involved, so we will be happy by just showing the matching of some subsets of states.

[2] A group is simply laced if all its roots have length square equal to 2).

For instance, it is easy to see that in either case the gauge group is defined by the surviving non-zero root vectors

$$(\underline{\pm,\pm,0,0,0,0,0,0};0,0,0,0,0,0,0,0) \quad ; \quad (0,0,0,0,0,0,0,0;\underline{\pm,\pm,0,0,0,0,0,0})$$

satisfying $P \cdot A \in \mathbf{Z}$. They correspond to a group $SO(16) \times SO(16)$ in both cases.

We can also match other states. Let us consider states uncharged under the 10d gauge group, i.e. neutral under $SO(16)\times SO(16)$, which have $P_L = 0$ and so $P^I = -\tilde{R}A^I_j w^j$. Using the particular form of the vectors P^I for the lattices, it can be seen that this condition requires that w is even, $w = 2m$. Hence, the spacetime left and right moving momenta are

$$p_{L,R} = \frac{k}{R} \pm \frac{wR}{\alpha'} + \frac{wR}{2}\frac{A\cdot A}{R^2} = \frac{k+2m}{R} \pm \frac{2mR}{\alpha'} \qquad (11.50)$$

and similarly for primed quantities. Defining $\tilde{k} = k + 2m$, we get

$$\begin{aligned} p_{L,R} &= \frac{\tilde{k}}{R} \pm \frac{2mR}{\alpha'} \\ p'_{L,R} &= \frac{\tilde{k}'}{R'} \pm \frac{2m'R'}{\alpha'} \end{aligned} \qquad (11.51)$$

We see that the two theories are equivalent for $R' = \alpha'/(2R)$, $\tilde{k}' = m$, $m' = \tilde{k}$. $E_8 \times E_8$ heterotic theory and $SO(32)$ heterotic theory can be considered different (decompactification) limits of this 9d theory. We then have a picture similar to that of type II theories.

11.4 Toroidal compactification of type I superstring

In this section we study type I superstring compactified on a circle. Generalization to $\mathbf{T^d}$ is analogous and will be mentioned only briefly.

Recall thet type I theory is a theory of unoriented closed and open strings. We have the 10d massless fields G, B, ϕ, and $SO(32)$ gauge bosons (and superpartners). This field content is the same as for the $SO(32)$ heterotic, which means that in the large R regime the results (which are well described by field theory in this regime) will agree with those in heterotic theory. The string theory description, however, will be very different, and the stringy features, like gauge enhancement or T-duality will be very different.

Before entering the detailed discussion, let us point out that in a general toroidal compactification it is possible to turn on background for the RR 2-form B; however, it is not known how to couple such backgrounds to the 2d worldsheet theory. Hence, the only backgrounds we will be able to describe exactly in the string theory are metric and Wilson line backgrounds.

11.4.1 Circle compactification without Wilson lines

We start discussing the simplest case of compactification on a circle of radius R, with zero gauge background. We have to describe the closed and open string sector independently.

Closed string sector

The toroidal compactification of the closed sector of type I is simply the Ω projection of the toroidal compactification of type IIB theory. In type IIB theory on a circle, different sectors of the theory are characterized by the momentum and winding, k and w, which define the mode expansion of the compactified direction (for clarity we omit the index in X^9)

$$\begin{aligned} X_L(\sigma+t) &= \frac{x_0}{2}+\frac{p_L}{2p^+}+\frac{1}{\alpha' p^+}N_B \\ X_R(\sigma-t) &= \frac{x_0}{2}+\frac{p_R}{2p^+}+\frac{1}{\alpha' p^+}\tilde{N}_B \end{aligned} \tag{11.52}$$

The effect of Ω on k, w is easy to find out, by recalling that it maps X to X^Ω such that

$$X^\Omega(\sigma,t)=X(-\sigma,t) \tag{11.53}$$

This implies that $Omega$ acts by $x_0\to x_0$, $k\to k$, $w\to -w$.

Hence Ω-invariant states are linear combinations of states in opposite winding sectors, schematically $|w\rangle+|-w\rangle$. This implies that winding number is not a well defined quantum number for states in this theory. This will be a relevant point in understanding some features of the T-dual version.

In the $w=0$ sector, Ω relates states within this sector. This implies that we get the usual projection on the operator piece of the states; namely in the NSNS sector the states of the form

$$\psi^i_{-1/2}|w=0\rangle\otimes\tilde{\psi}^j_{-1/2}|w=0\rangle \tag{11.54}$$

survive only by taking the symmetrized product, exactly as in the original 10d theory. Indeed it is easy to check that the $w = 0$ sector gives the KK reduction of the massless fields in the original 10d theory.

In sectors of $w \neq 0$ (these are massive states, but we are interested in discussing them at this point), there exist Ω-invariant combinations of winding excitations of these states both in symmetrized and antisymmetrized prodcucts. For instance, in the NSNS sector the state

$$\psi^{[i}_{-1/2}|w\rangle \otimes \tilde{\psi}^{j]}_{-1/2}|w\rangle + \psi^{[i}_{-1/2}|-w\rangle \otimes \tilde{\psi}^{j]}_{-1/2}|-w\rangle \tag{11.55}$$

survives. It can be considered as a winding excitation of the field B_{ij} since it is in a sense left-right antisymmetric. Nevertheless it is invariant under Ω due to the additional action on winding number. The observation that winding excitations of Ω-odd 10d massless fields are Ω invariant will be relevant in the discussion of the T-dual picture.

In any event, the spectrum of states massless at generic R is obtained by the Ω-invariant states in the $k = 0$, $w = 0$ sector of the IIB theory. As expected, this is simply the zero modes of the KK reduction of the 10d $\mathcal{N} = 1$ supergravity multiplet.

Notice that since the parent IIB theory did not have any enhanced symmetries at special values of R, neither does the closed sector of type I theory.

Open string sector

(We start the discussion in compactfications without Wilson lines; inclusion of the latter will be discussed in later sections.)

A key difference between the compactification of open string sectors and closed string sectors is the absence of winding. As shown in figure 11.2, open strings can always unwind in a compact dimension. This agrees with the fact that winding was defined using the periodicity in σ for closed strings, and this does not exist in open strings.

Hence, the only effect of the circle compactification in the open string sector is that the internal momentum is now quantized and equal to k/R. Since there is no winding, compactification of open strings is very much like KK compactification in field theory.

We have the mode expansion for 2d bososn

$$X(\sigma, t) \; = \; x_0 + \frac{k}{Rp^+} + \text{oscillators} \tag{11.56}$$

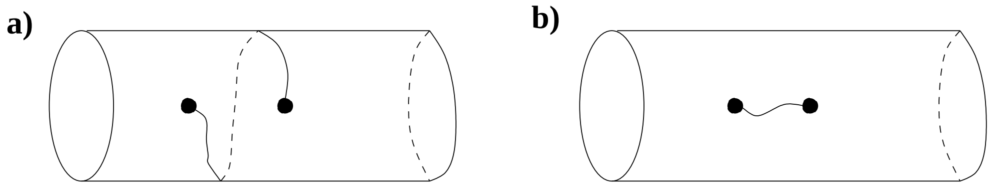


Figure 11.2: Open strings with NN boundary conditions in a compact direction cannot wind around it. String seemingly wrapped on the internal circle are in the same topological sector as strings with no winding.

leading to the mass formula

$$m^2 = \frac{1}{\alpha'}(N_B + N_F + E_0) + \frac{k^2}{R^2} \tag{11.57}$$

Thus massless states correspond to $k = 0$ and reproduce the zero modes of the KK reduction of the 10d massless fields. Namely $SO(32)$ gauge bosons, one real scalar in the adjoint representation, and fermion superpartners. States with non-zero k are the KK replicas of these zero modes. Again, there are no special values of R at which new states become massless.

11.4.2 T-duality

In this section we study the T-dual of the type I theory, also called type I' theory.

Closed string sector

Again, the closed string sector presents an infinite tower of states (with $k = 0$ and arbitrary w) which become light as $R \to 0$. This suggests the existence of a T-dual theory, which becomes decompactified in this limit. In this section we find out the structure of this T-dual theory, which is related to the original one by

Original	T-dual
R	$R' = \alpha'/R$
$k,\ w$	$k' = w,\ w' = k$
$X_L,\ \psi_L$	$X_L,\ \psi_L$
$X_R,\ \psi_R$	$X'_R = -X_R,\ \psi'_R = \psi_R$

(the action is only on the coordinate along the compact direction 9, on which we are T-dualizing).

In the **closed string sector**, the dual theory described by (X', ψ') corresponds to type IIA theory (since T-duality flips the right moving GSO projection) compactified on a circle of radius $R' = \alpha'/R$, and modded out by an orientifold projection. The orientifold action on X' can be obtained by reading the Ω action on left and right movers

$$X_L^\Omega(\sigma + t) = X_R(-\sigma - t) \quad ; \quad X_R^\Omega(\sigma - t) = X_L(-\sigma + t) \quad ; \qquad (11.58)$$

and constructing $X^{\Omega\prime} = X_L^\Omega - X_R^\Omega$ and $X' = X_L - X_R$. We obtain

$$X^{\Omega\prime}(\sigma, t) = X_L^\Omega(\sigma + t) - X_R^\Omega(\sigma - t) = X_R(-\sigma, t) - X_L(-\sigma + t) = -X'(-\sigma, t) \qquad (11.59)$$

Hence the T-dual is type IIA theory on a circle modde out by an orientifold action $\Omega\mathcal{R}$, where $\mathcal{R}$ is a geometric action $x^9 \to -x^9$. It is easy to verify that the action (11.59) on the mode expansion is to flip the momentum and leave winding invariant, as should be the case for the T-dual of Ω.

Recalling our lecture on unoriented strings, recall that we claimed that one can mod out a theory by Ω only if it is left-right symmetry (i.e. IIB theory). Here we are modding by $\Omega\mathcal{R}$ and this can be done only if the theory is left-right symmetric up to a GSO shift (i.e. IIA theory).

Notice that $\mathcal{R}$ has fixed points at two diametrically opposite point in the dual circle, see figure 11.3. These are regions where the orientation of a string can flip. They are 9-dimensional subpaces of 10d space, and are called orientifold 8-planes, O8-planes for short (they involve 8 spatial plus one time direction).

The existence of these special points implies that the compactification violates translation invariance. This is not strange, since staes in the original model did not have winding as a good quantum number; hence in the T-dual, momentum is not a good quantum number, so there are violations of translation invariance in the internal coordinate.

Finally, let us mention that states are in general linear combinations of states of the original theory in sectors of opposite internal momentum. In the $k = 0$ sector this implies the usual projection, and that only ΩR even states arise. However, in sectors of $k \neq 0$ there exist momentum excitations of fields which are ΩR odd in the 10d theory. This has the interesting consequence that such 10d fields are not identically vanishing in the model, but rather propagate in the 'bulk', away form the orientifold planes. The orientifold

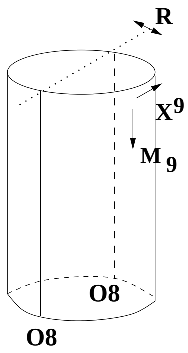


Figure 11.3: O8-planes in type I' theory.

projection impose the boundary condition that 10d ΩR odd fields vanish at the O8-plane location, and so lead to no zero modes. Hence, in the bulk the theory is still locally type IIA theory, and it is the O8-planes that project out part of the zero modes (although KK excitations survive).

Open string sector

We now study the open string sector in the T-dual version. The local 2d dynamics of the T-dual open string sector should be that of an (orientifold version of) type IIA theory. In particular, it implies that the interior of open string worldsheets propagates in 10d. However, since the original open string sector does not have winding number in x^9, the T-dual open string sector has no momentum in x^9. This implies that such fields propagate only in 9d.

The resolution to this seeming paradox can be understood by finding out the boundary conditions for the open strings in the T-dual. In the original theory we have Neumann boundary conditions at the open string endpoints

$$\begin{aligned} &\partial_\sigma X(\sigma,t)|_{\sigma=0,\ell}=0 \\ &\partial_\sigma X_L(\sigma+t)|_{\sigma=0,\ell}+\partial_\sigma X_R(\sigma-t)|_{\sigma=0,\ell}=0 \end{aligned} \tag{11.60}$$

This can be written as

$$\partial_t X_L(\sigma+t)|_{\sigma=0,\ell}-\partial_t X_R(\sigma-t)|_{\sigma=0,\ell}=0 \tag{11.61}$$

Namely, in terms of the T-dual coordinate $X=X_L-X_R$

$$\partial_t X'(\sigma,t)|_{\sigma=0\ell}=0 \tag{11.62}$$

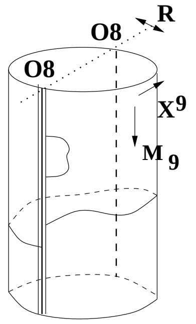


Figure 11.4: Open strings in type I' have endpoints at a fixed position in x^9, although their 'inside' can still move in 10d.

These are Dirichlet boundary conditions (the corresponding open strings are said to have DD boundary conditions in x^9). They imply that the open string endpoints cannot move from a fixed value of the coordinate x^9, so the open string states are forced to move in 9d only. However the inside of the open string can still move in 10d. See figure 11.4.

One may question whether this is consistent. For instance, the open string sector is not translational invariance in x^9, but neither is the underlying closed string sector, so this is not worrisome. Another issue is that we obtained Neumann boundary conditions as some correct boundary conditions to recover the familiar equations of motion for the 2d theory in the inside of the open string worldsheet. In fact, we can check that Dirichlet boundary conditions do the job as well. Recall that the variation of the Polyakov action is

$$\begin{aligned} \delta S_P &= -\frac{1}{2\pi\alpha'}\int_\Sigma d^2\xi\, g^{ab}\partial_a X^\mu \partial_b \delta X^\mu = \\ &= -\frac{1}{2\pi\alpha'}\int_{-\infty}^{\infty} de\,(g^{ab}\delta X^\mu \partial_b X_\mu)|_{\sigma=0}^{\sigma=\ell} + \frac{1}{2\pi\alpha'}\int_\Sigma d^2\xi\,\delta X_\mu\, g^{ab}\partial_a\partial_b X^\mu \end{aligned} \quad (11.63)$$

Dirichlet boundary conditions in x^9 imply that $\delta X^9 = 0$ at $\sigma = 0, \ell$. Hence DD boundary conditions on x^9 and NN boundary conditions on the remaining coordinates ensure that the first term in the second line vanishes and we recover the correct 2d dynamics in the interior of the open strings.

It is interesting to notice that the mode expansion for the T-dual coordinate $X'(\sigma,t)$ contains a winding term and no momentum

$$X'(\sigma,t) \;=\; \frac{2\pi R'}{\ell} w'\sigma + \text{osc.} \tag{11.64}$$

which is indeed allowed by DD boundary conditions. Pictorially, existence of winding for open strings with endpoints stuck at points in x^9 is manifest in figure 11.4. Notice that the endpoints of all open strings are necessarily located at the same point in $\mathbf{S^1}$. This can be seen directly from the above

$$X'(\sigma=\ell,t) - X'(\sigma=0,t) = 2\pi R\, w' \tag{11.65}$$

so the open strings stretch whole periods of x^9, such that endpoints always lie at $x^9 = 0$. This is true regardless of the Chan-Paton indices carried by the string. The presence of wilson lines in the original picture will modify this last fact, as we show later on.

A very intuitive picture, which becomes even more useful in more complicated situations (like with non-trivial Wilson lines in the original picture), is to consider that the model contains some objects, spanning the 9d hyperplane at $x^9 = 0$, called D8-branes, and on which open strings are forced to end. In fact, the precise picture is that there exist one such D8-brane for each possible value of the Chan-Paton index (32 D8-branes for the T-dual of type I). An open string endpoint with Chan-Paton index a must end on the a^{th} D8-brane. In the present situation, all 32 D8-branes are sitting at the same location in x^9.

The open string spectrum is easy to recover in this language. In the massless sector, we have open strings with all possible combinations of Chan-Paton factors (i.e. ending on the 32 D8-branes in all possible ways). This would lead to a 9d $U(32)$ vector multiplet with respect to the 16 unbroken supersymmetries. Since the open strings are sitting on top of an orientifold plane, we have to keep ΩR invariant states, leading to a 9d $SO(32)$ vector multiplet with respect to the 16 unbroken supersymmetries.

Notice that this gauge sector propagates in a 9d subspace of spacetime, while gravity and other fields still propagate in 10d. The possibility of constructing models of this kind has led to the brane-world idea, the proposal that perhaps the Standard Model that we observe is embedded in a brane which spans a subpace in a full higher dimensional spacetime. This would lead to the existence of extra dimensions which are detectable only using gravitational experiments. We will learn more about branes, and model building with them in later lectures.

11.4.3 Toroidal compactification and T-duality in type I with Wilson lines

As in heterotic theories, upon compactification there exist 9d scalars tranforming in the adjoint representations of the gauge group. Their vevs parametrize the possibility of turning on constant backgrounds for the internal components of the gauge fields. In this section we study the modifications they introduce for type I.

Clearly the closed string sector is insensitive to the presence of Wilson lines, since it contains states neutral under the gauge symmetry. The only modifications occur in the open string sector. To describe them, we need to couple the gauge background to the 2d theory. This is easily done by recalling the rule that an open string endpoint with Chan-Paton a has charge ± 1 under the $U(1)$ gauge boson arising in the sector of aa open strings. This implies that the worlsheet action must be modified by a boundary term

$$\Delta S = \int dt \, - iq_a A_i^a \, \partial_t X^\mu \tag{11.66}$$

Before the orientifold projection, there are 32 $U(1)$ gauge bosons, which are paired by the orientifold action. In terms of this parent $U(32)$ original theory, the most general wilson line consistent with the Ω action is

$$(A_i^a) = \frac{1}{2\pi R}(\theta_1, \theta_2, \ldots, \theta_{16}; -\theta_1, -\theta_2, - \ldots, \theta_{16}) \tag{11.67}$$

After the orientifold action, the surviving Cartans are linear combinations of the above; in terms of the $U(1)^{16}$ Cartan subalgebra of $SO(32)$ the wilson line is described by

$$(A_i^I) = \frac{1}{2\pi R}(\theta_1, \theta_2, \ldots, \theta_{16}) \tag{11.68}$$

Although the latter expression is more correct, it is sometimes more intuitive to use (11.67) to display the Chan-Patons and their orientifold images explicitly.

The Wilson line has the only effect of shifting the internal momentum, as discussed above in field theory terms. Namely, for an open string in the ab Chan-Paton sector (and so, with charges $(+1, -1)$ under $U(1)_a \times U(1)_b$, we have

$$p = \frac{k}{R} + \frac{\theta_a - \theta_b}{2\pi R} \tag{11.69}$$

here we are using the notation (11.67), so $\theta_{a+16} = -\theta_a$. The spacetime mass formula for these states is

$$m^2 \;=\; (\frac{k}{R} + \frac{\theta_a - \theta_b}{2\pi R})^2 + \frac{1}{\alpha'}(N_B + N_F - 1) \qquad (11.70)$$

For generic R and θ_a, the gauge group is broken to $U(1)^{16}$, since only aa states are able to lead to massles modes. When several, say N eigenvalues θ_a coincide and are not zero or π, then there are additional massless fields, leading to $U(N)$ gauge bosons and superpartners. Finally, when n eigenvalues vanish or are equal to π, the gauge symmetry is $SO(2n)$.

The moduli space of compactifications is difficult to obtain, and there is no analog of the Narain lattice. Hence, without further ado, we turn to the discussion of T-duality.

T-duality

The T-dual closed string sector in still given by type IIA theory on a circle, modded out by $\Omega\mathcal{R}$. The T-dual of the open strings is slightly modified by the Wilson lines. By simply mapping the mode expansion of the original into the mode expansion of the T-dual, we find that the dual coordinate has shifted winding

$$X'(\sigma,t) \;=\; \text{const.} + \frac{2\pi R'}{\ell} w'\sigma + \frac{2\pi R'}{\ell}\frac{(\theta_a - \theta_b)}{2\pi}\sigma \,+\, \text{osc.} \qquad (11.71)$$

This implies that the open string endpoints of ab strings are at different locations in x^9

$$X'(\sigma=\ell,t) - X'(\sigma=0,t) \;=\; 2\pi R'\, w' \,+\, \theta_a R' \,-\, \theta_b R' \qquad (11.72)$$

The mass formula for ab strings is

$$m^2 = \frac{R}{\alpha'}\,(w + \frac{\theta_a}{2\pi} - \frac{\theta_b}{2\pi})^2 \,+\, \frac{1}{\alpha'}(N_B + N_F - E_0) \qquad (11.73)$$

In more intuitive terms, recall our description of an endpoint with Chan-Paton a as ending on the a^{th} D8-brane. What we have found is that $\theta_a R'$ is the location in x^9 of the a^{th} D8-brane. The ab open strings start on the a^{th} and end on the b^{th} D8-brane, so their length is $\theta_a R - \theta_b R$, modulo the period $2\pi R$. This stretching contributes to the mass of the corresponding state. See figure 11.5.

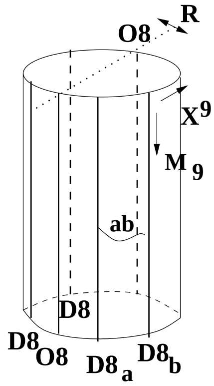


Figure 11.5: Open string endpoints in the T-dual of type I with Wilson lines are fixed on D8-branes at different positions in the circle. Their stretching is determined by the location of the D8-branes.

The D8-brane picture makes the gauge symmetry enhancements clear. Generically the D8-branes are located at different positions, so the generic gauge symmetry is $U(1)^{16}$ (since only aa strings have zero stretching). When several, say N θ_a's coincide, several D8-branes overlap, and the corresponding ab strings are massless, leading to $U(N)$ gauge symmetries. Finally, if N θ_a are zero or π, D8-branes and their orientifold images coincide on top of an O8-plane, leading to $SO(2N)$ gauge symmetry.

It is interesting to re-interpret the RR tadpole cancellation conditions in the T-dual language. In this case, the crosscap diagrams are located on top of the O8-planes, and in a sense compute the RR charge of these objects (the strength of their coupling to the RR 9-form (dual to the original 10-form). The disk diagrams are located on top of the D8-branes, and compute the RR charge of these objects. RR tadpole cancellation condition corresponds to the requirement that the fluxlines of the RR 9-form have nowhere to go in the internal space, which is compact, so the total charge must vanish (see fig11.6). This is Gauss law in a compact space [3] . It is possible to compute these tadpoles as we did for type I, and obtain that each O8-plane has -16 times the charge of a D8-brane. Hence we have $2 \times (-16) + 32 \times 1 = 0$.

[3]Equivalentely, one can check that the KK reduction of the 9-form has a zero mode, which corresponds to a 9-form in 9d, which has no kinetic term. RR tadpole cancellation can be recovered as the consistency condition for its equations of motion. This description is more analogous (T-dual) to the one used in type I.

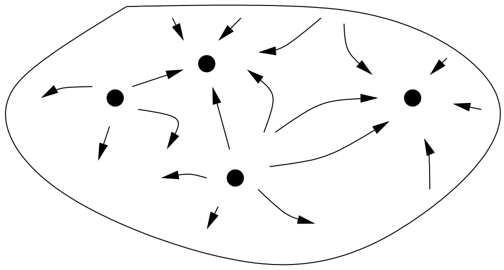


Figure 11.6: Schematic picture of the interpretation of RR tadpole cancellation as charge cancellation in a compact space.

We conclude with some relevant observations

• The generalization of this idea to further T-dualities is clear. In the closed string sector the orientifold action acquires an additional geometric piece inverting the T-dualized coordinate. Hence, in general we will find theories obtained from toroidal compactification of type IIA/B modded out by Ωg, where g is a geometric action flipping r coordinates, with r even/odd for IIB/IIA. This introduces 2^r O$(9-r)$-planes, which can be seen to carry $32/2^r$ units of RR charge. In order to cancel the RR tadpoles, we introduce 32 D$(9-r)$-branes, which can be at arbitrary locations, but respecting the $\mathbf{Z_2}$ symmetry imposed by g.

• The original type I theory also admits a description in terms of O-planes and D-branes. The Ω projection can be said to introduce an O9-plane (which fills spacetime completely), and the open strings (which can end anywhere in 10d space) can be said to end on D9-branes (which fill spacetime completely).

We should not worry too much about understanding all the details of D-branes at this point. Such objects will reappear in a different way in subsequent section. In fact they correspond to new non-perturbative states in type II string theory. This can be understood already in our picture: recalling that the bulk of spacetime is described by type IIA theory, if one takes the decompactification limit in which the O8-planes go off to infinity, keeping the D8-branes in the middle of the interval, we are roughly left with non-compact type IIA theory in the presence of D8-branes, see figure 11.7. This shows that there exist states in type IIA string theory which are not obtained as perturbative excitations of the type IIA string. Rather, this states should be regarded as a non-perturbative state, analogous in many respects to a soliton. We will come back to these states in later lectures.

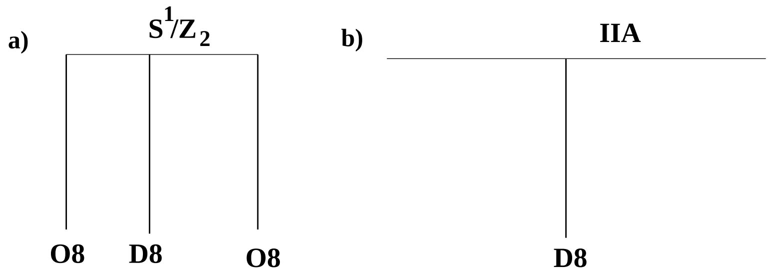


Figure 11.7: The decompactification limit of type I' keeping the D8-branes at finite distance produces type IIA theory with a topological defect (domain wall) given by the D-brane.

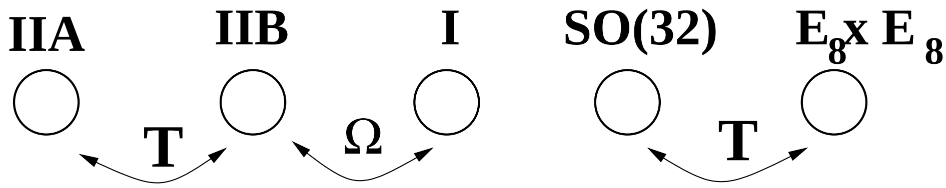


Figure 11.8: .

11.5 Final comments

Let us summarize this lecture by emphasizing that we have shown an extremely intimate relation between the different string theories, once we start compactifying them. See figure B.1.

This is all very nice, but we should recall that we started out studying string theory as a theory with the potential to unify the interactions we observe in Nature. The theories we have obtained have too much supersymmetry to allow for chirality, so they are quite hopeless as theories of our world. Therefore, we will turn to the study of other compactifications in subsequent lectures.

Chapter 12

Calabi-Yau compactification of superstrings. Heterotic string phenomenology

12.1 Motivation

We have seen that toroidal compactification leads to 4d theories at low energies. However, it is too simple to lead to anything realistic, similar to the Standard Model of Particle Physics. The fact that toroidal compactification does not break any of the supersymmetries of string theory implies the 4d theories are non-chiral. We are missing an essential ingredient of Particle Physics.

Thus we have to consider more general compactifications with background geometry $\mathbf{M_4} \times \mathbf{X_6}$, where $\mathbf{X_6}$ is a compact curved manifold [1]. Since the background metric is not flat, the worldsheet 2d theory is interacting, and not exactly solvable. Hence one usually works at leading order in the 2d expansion prameter, which is α'/r^2, where r is a curvature length scale in spacetime. This corresponds to working at low energies, in the supergravity limit, and is a good approximation if all curvature length scales of $\mathbf{X_6}$ are large compared with the string length. This is essentially a point particle limit, and the stringy physics will be hidden in the α' corrections, which are

[1] We also include in our ansatz that backgrounds for other bosonic fields are trivial, e.g. we do not consider compactifications with field strength fluxes for p-form fields, which only very recently have been considered in the literature.

very difficult to obtain.

It should be pointed out, though, that there exist some abstract exactly solvable 2d conformal field theories, known as Gepner models, which are proposed to describe (exactly in α') the physics of string theories on spaces of stringy size. Also, in next lecture we will study orbifolds, which are in a sense, simple versions of non-trivial spaces, which still lead to free 2d worldsheet theories (with sectors of non-trivial boundary conditions).

12.1.1 Supersymmetry and holonomy

We are interested in compactifications which preserve some 4d supersymmetry. Compactifications breaking all the supersymmetries would be very interesting but

- often contain instabilities, appearing as tachyonic fields in 4d.
- lead to a too large 4d cosmological constant to be of any phenomenological use to describe the real world.

Nevertheless, it is important to realize that assuming supersymmetry is also an oversimplification if one is interested in describing the real world, which is not exactly supersymmetric. Upon breaking supersymmetry (by some of the mechanisms in the market) the above two problems rearise [2].

Finally, it is possible to see that the conditions imposed on $\mathbf{X_6}$ by supersymmetry ensure that the background satisfies the supergravity equations of motion, it is a good vacuum of the theory. This can be found in the main reference for this lecture [61].

What are the conditions on $\mathbf{X_6}$ in order to have some unbroken 4d supersymmetry? Recall from our discussion of Kaluza-Klein reduction that 4d fields visible at low energies are zero modes, constant in the internal space. Similarly, gauge symmetries visible at low energies correspond to gauge transformations constant over the internal space. Analogously, supersymmetries unbroken in the low energy 4d physics correspond to (local) supersymmetry parameters (which are spinors $\xi(x^\mu, x^i)$ in $\mathbf{M_4} \times \mathbf{X_6}$) which are covariantly constant in $\mathbf{X_6}$ (with the connection inherited from the metric), i.e. $\nabla_{\mathbf{X_6}}\xi(x^i) = 0$.

Recalling now the discussion of the holonomy group of a Riemannian manifold, we can obtain a conditions on $\mathbf{X_6}$ to admit covariantly constant

[2]Yes, it is a bit disappointing that for the moment string theory has not given a strong proposal to solve the cosmological constant problem, despite many attempts.

spinors. Clearly, a covariantly constant spinor is a singlet under the holonomy group (of the spinor bundle with the spin connection), since it does not change under parallel transport around a closed loop. This implies that the holonomy group of a Riemannian manifold IX_6 leaving some 4d susy unbroken is not generic. The generic holonomy for a metric in a 6d manigold is $SO(6)$, and spinors transform in the representation 4 or $\overline{4}$ under it (depending on their chirality [3]), hence there is no singlet, and no covariantly constant spinor. For metrics of $SU(3)$ holonomy, spinors transform as $3+1$ or $\overline{3}+1$, hence there are components which are singlets under the holonomy group, corresponding to covariantly constant spinors. The decomposition of a susy parameter in 10d under the holonomy and 4d Lorentz groups follows from the following chain

$$\begin{array}{ccccc} SO(10) & \to & SO(6)\times SO(4) & \to & SU(3)\times SO(4) \\ 16 & & (4,2)+(4',2') & & (3,2)+(\overline{3},2')+(1,2)+(1,2') \end{array}$$

In the last column only the $SU(3)$ singlet components lead to 4d supersymmetries, while the others are broken by the compactification.

The surviving supersymmetries can also be verified by looking at the KK reduction of 10d gravitinos under the holonomy group. This is described by the following chain. The 10d gravitinos are in the say 56_S of $SO(8)$, which arises from a product $8_V \times 8_C$. Decomposing with respect to $SO(6)\times SO(2)$, we have $8_V \to 6_0 + 1_{\pm 1}$ and $8_C = 4_{1/2} + \overline{4}_{-1/2}$, where subindices denote the $SO(2)$ charges. We are interested in 4d gravitinos, which have spin 3/2 with respect to $SO(2)$; these fields are obtained from the product $1_{\pm1}\times(4_{1/2}+\overline{4}_{-1/2}$, and decompose under $SU(3)\times SO(2)$ as $1_{\pm1}\times(3_{1/2}+\overline{3}_{-1/2}+1_{1/2}+1_{-1/2})$. Clearly only the latter lead to 4d gravitinos unbroken by the compactification. It is possible to check that one 10d gravitino leads to one 4d gravitino if the holonomy group of the compactification manifold is $SU(3)$.

The generalization of the above statements to other dimensions is that compactification on a $2n$-dimensional manifold with $SU(n)\subset SO(2n)$ holonomy preserves some supersymmetry.

[3]The Lie algebras of $SO(6)$ and $SU(4)$ are the same, and the spinor representations of $SO(6)$ are the fundamental and antifundamental of $SU(4)$, so they are often written 4 and $\overline{4}$.

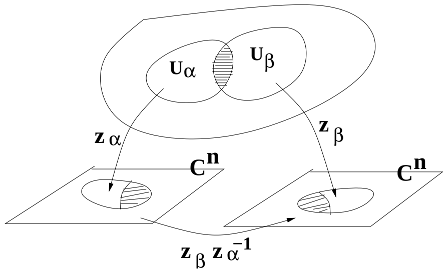


Figure 12.1: Charts covering a complex manifold.

12.1.2 Calabi-Yau manifolds

A $2n$-dimensional manifold admitting a metric with spin connection of $SU(n)$ holonomy is a Calabi-Yau manifold.

This definition is difficult to use in order to determine whether a manifold is Calabi-Yau, since in principle one needs an explicit construction of the metric. This is very difficult: in fact there is no known explicit metric for any (non-trivial) compact Calabi-Yau, explicit metrics are knonw only for a few examples of non-compact spaces. Happily the existence of a metric with this property is guaranteed for manifolds satisfying the following (simplest to check) topological conditions: the manifold must be Kahler and (its tangent bundle must) have vanishing first Chern class.

To understand better the meaning of these conditions, we need some background information on complex differential geometry.

An n-dimensional complex manifold is a topological space M, together with a holomporphic atlas, i.e. a set of charts $(U_\alpha, z_{(\alpha)})$ where $z_{(\alpha)}$ are maps from U_α to some open set in $\mathbf{C^n}$, such that i) the U_α cover M, ii) on $U_\alpha \bigcap U_\beta$, the map

$$z_{(\beta)} \circ z_{(\alpha)}^{-1} : z_{(\alpha)}(U_\alpha \bigcap U_\beta) \longrightarrow z_{(\beta)}(U_\alpha \bigcap U_\beta) \qquad (12.1)$$

is holomorphic (namely $\partial z_{(\beta)}/\partial \overline{z}_{(\alpha)} = 0$. See figure 12.1.

Notice that a complex n-dimensional manifold can always be regarded as a real $2n$-dimensional differential manifold, by simply splitting the complex coordinates into its real and imaginary parts. On the other hand,a real $2n$-dimensional manifold M can be regarded as an n-dimensional complex manifold only if it admits a globally defined tensor of type $(1,1)$, $J_m^n dx^m \otimes \partial_n$ satisfying

i) $J_m^n J_n^l = -\delta_m^l$

(this is used to define local complex coordinates $dz^i = dx^i + iJ_l^i dy^l$ and $d\overline{z}^i = dx^i - iJ_l^i dy^l$)

ii) The Niejenhuis tensor vanishes

$$N_{ij}^k = \partial_{[j} J_{i]}^k - J_{[i}^p J_{j]}^q \partial_q J_p^k = 0 \tag{12.2}$$

which ensures that the local complex coordinates have holomorphic transition functions. Such a J is called a complex structure [4].

Notice that a given real differential manifold can admit many complex structures. A familiar example is provided by the 2-torus, which admits a one (complex) dimensional family of complex structures, parametrized by a complex number τ; the two real coordinates x, y can be combined to form a complex coordinates via $dz = dx + \tau dy$.

In a complex manifold, p-forms and their cohomology classes (and p-chains and their homology classes) can be refined according to their number of holomorphic and antiholomorphic indices [5]. For instance, the 3-cohomology group splits as

$$H^3(M) = H^{(3,0)}(M) + H^{(2,1)}(M) + H^{(1,2)}(M) + H^{(0,3)}(M) \tag{12.3}$$

where $H^{(p,q)}$ corresponds to forms with p holomorphic and q antiholomorphic indices (spanned by a basis $dz^{i_1} \wedge \ldots \wedge dz^{i_p} \wedge d\overline{z}^{j_1} \wedge \ldots \wedge dz^{j_p}$. The dimensions of the $H^{(p,q)}$ are denoted $h_{p,q}$ and known as Hodge numbers; although to define them we have introduced a complex structure, they do not depend on the particular complex structure chosen, so they are topological invariants of M.

A metric in a complex manifold is called hermitian if it is of the form

$$ds^2 = g_{i\overline{j}} dz^i d\overline{z}^{\overline{j}} \tag{12.4}$$

namely has non-zero components only for mixed indices. Such metric can be used to lower one index of the complex structure tensor and thus define the $(1,1)$ form

$$J = g_{i\overline{j}} dz^i \wedge d\overline{z}^{\overline{j}} \tag{12.5}$$

[4]Manifolds with tensors J satisfying i) but not ii) are called almost complex, and J is called almost complex structure.

[5]In fact, this can be done even for almost complex manifolds.

A hermitian metric is called Kahler metric if the associated $(1,1)$-form satisfies

$$dJ = 0 \tag{12.6}$$

The $(1,1)$-form is known as Kahler form. A manifold which admits a Kahler metric is called a Kahler manifold; this is a topological property of the manifold.

Notice that the Kahler form defines a non-trivial cohomology class in $H^{(1,1)}(M)$. It defines a cohomology class because it is closed. We can show that the class is non-trivial because (12.5) implies

$$\int_M J \wedge \ldots \wedge J = \int_M \sqrt{\det g}\; dz^1 d\overline{z}^1 \ldots dz^n d\overline{z}^n = \mathrm{Vol}(M) \tag{12.7}$$

which would be vanishing if J is exact (since $J = dA$ would imply $\int J \ldots J = \int d(AJ \ldots J) = 0$).

The Kahler form is very interesting since it characterizes the overall volume of the manifold M. In particular, α' corrections are in fact weighted by the adimensional parameter α'/r^2, where r is an overall size determined by the Kahler form.

Returning to the issue of holonomy, the crucial property of Kahler manifolds is that the Christoffel connection induced by the Kahler metric leads to a parallel transport that does not mix holomorphic and antiholomorphic indices. This implies that the holonomy group is in a $U(n)$ subgroup of $SO(2n)$, as is manifest e.g. by splitting the basis of tangent space in holomorphic and antiholomorphic elements

$$(\partial_{z^1}, \ldots, \partial_{z^n}; \partial_{\overline{z}^1}, \ldots, \partial_{\overline{z}^n}) \tag{12.8}$$

The $U(1)$ part of the holonomy can be seen to be associated to the Ricci tensor, so the manifold must admit a Kahler and Ricci-flat metric to have $SU(n)$ holonmy. A necessary topological condition for this is that the first Chern class $c_1(R)$ of the tangent bundle is trivial. Calabi conjectured this to be also a sufficient condition, as was finally proved by Yau (hence the name Calabi-Yau for such spaces).

Yau's theorem states that, given a complex manifold with $c_1(R) = 0$ and Kahler metric g with Kahler form J, there exists a unique Ricci-flat metric g' with Kahler form J' in the same cohomology class. It provides, as promised, a topological way of characterizing manifolds for which a $SU(n)$

holonomy metric exists (without constructing it explicitly). This facilitates the classification and study of Calabi-Yau spaces, in fact tables with many hundreds of such spaces exist in the literature.

Yau's theorem also provides a characterization of the parameters that determine the $SU(n)$ holonomy metric. For a given differential manifold M we should

i) specify the parameters that define a complex structure on this real manifold to make it a complex manifold. This set of parameters spans what is called the compelx structure moduli space, and can be computed to have (complex) dimension $h_{2,1}(M)$.

ii) for fixed complex structure, specify the parameters which define the Kahler class. This set of parameters is known as Kahler moduli space, and clearly has (real) dimension $h_{1,1}(M)$.

The complete moduli space of Calabi-Yau metrics in a given differential manifold M is (locally) the product of these.

We would like to point out that the condition for supersymmetry which we have used is valid to lowest order in α'. In particular, one can imagine that there could be higher order α' corrections that modify the 'equation of motion' condition Ricci=0. However, there are diverse arguments (see [62]) showing that in differential manifolds, satisfying the topological conditions of being Kahler and have zero first Chern class, there exists some underlying 2d interacting field theory which is conformal exactly in α'. In other words, the leading α' proposal for the metric can be consistently completed to an α' exact one.

The Calabi-Yau condition implies certain structure of Hodge number. For 6d manifolds admitting a metric of holonomy $SU(3)$ (and not in a proper subgroup like $SU(2)$), often referred to as Calabi-Yau threefolds, they read

$$\begin{array}{ccccccc} & & & h_{0,0} & & & \\ & & h_{1,0} & & h_{0,1} & & \\ & h_{2,0} & & h_{1,1} & & h_{0,2} & \\ h_{3,0} & & h_{2,1} & & h_{1,2} & & h_{0,3} \\ & h_{3,1} & & h_{2,2} & & h_{3,1} & \\ & & h_{3,2} & & h_{2,3} & & \\ & & & h_{3,3} & & & \end{array} \qquad \begin{array}{ccccccc} & & & 1 & & & \\ & & 0 & & 0 & & \\ & 0 & & h_{1,1} & & 0 & \\ 1 & & h_{2,1} & & h_{2,1} & & 1 \\ & 0 & & h_{1,1} & & 0 & \\ & & 0 & & 0 & & \\ & & & 1 & & & \end{array}$$

where equality of some Hodge numbers is due to duality between $H^{(p,q)}$ and $H^{(3-p,3-q)}$. Due to its shape, this diagram is known as Hodge diamond.

We conclude with some examples. In one complex dimension, the only compact Calabi-Yau space actually has trivial holonomy, it is the 2-torus. In two complex dimensions, there is only one topological space admitting $SU(2)$ holonmy metrics, known as K3 (complex) surface. Although a lot is known about the topology of this space, no explicit metrics are known. In three complex dimensions, there exist many compact Calabi-Yau spaces. One of the simplest is the quintic, which can be described as the (complex) hypersurface

$$f_5(z_1, \ldots, z_5) = 0 \tag{12.9}$$

in $\mathbf{P_5}$, the (four) complex (dimensional) projective space [6]. Here $f_5(z_1, \ldots, z_5)$ denotes a degree 5 polynomial (so that it is homogeneous and well-defined in $\mathbf{P_5}$). The general such polynomial (up to redefinitions) depends on 101 complex parameters, which determine the complex structure of the Calabi-Yau. Also, there is one Kahler parameter determining the overall size of $\mathbf{P_5}$ and hence of the quintic. Its Hodge diamond has therefore $h_{2,1} = 101$, $h_{1,1} = 1$.

12.2 Type II string theories on Calabi-Yau spaces

We now study what kind of theories arise from compactification of type II string theories on $SU(3)$ holonomy spaces.

12.2.1 Supersymmetry

Type II theories have two 10d gravitinos. Upon compactification on Calabi-Yau threefolds we obtain two 4d gravitinos, which corresponds to 4d $\mathcal{N} = 2$ supersymmetry. This is a non-chiral supersymmetry, so it appears for both IIA and IIB theories. The massless supermultiplets that may appear are:

i) the gravity multiplet, containing a graviton, a gauge boson (graviphoton), and two gravitinos of opposite chiralities

ii) the vector multiplet, containing a gauge boson, a complex scalar and a Majorana fermion, all in the adjoint representaion of the gauge group

[6] This is the set of points $(z_1, \ldots, z_5) \in \mathbf{C^5}$ with the equivalence relation $(z_1, \ldots, z_5) \simeq (\lambda z_1, \ldots, \lambda z_5)$ with $\lambda \in \mathbf{C} - \{\mathbf{0}\}$.

iii) the hypermultiplet, containing two complex scalars (in conjugate representations) and two Weyl fermions (in the same representation with opposite chiralities).

This structure makes it sufficient to determine the bosonic fields after compactification; the fermionic fields can be completed by using this multiplet structure.

12.2.2 KK reduction of p-forms

Since type II theories contain p-form fields in 10d, we need to know how to perform their KK reduction. A p-form in 10d $C_p(x^0,\ldots,x^9)$ can give rise to 4d q-forms via the ansatz

$$C_{\mu_1\ldots\mu_q m_1\ldots m_r}(x^0,\ldots,x^9) \;=\; C_{\mu_1\ldots\mu_q}(x^0,\ldots,x^3)\,A_{m_1\ldots m_r}(x^4,\ldots,x^9) \quad (12.10)$$

with $q+r=p$. The 4d q-form has a 4d mass given by the laplacian acting on the internal piece. The laplacian is read off from the kinetic term of p-forms, which is

$$\int dC\wedge *dC = \int (dC,dC) = \int (C,\Delta C) \quad (12.11)$$

and $\Delta = dd^\dagger + d^\dagger d$. Hence to get a massless 4d q-form we need to pick the internal r-form $A_{m_1\ldots m_r}(x^4,\ldots,x^9)$ to be a harmonic r-form in $\mathbf{X_6}$, namely $dA=0$, $d^\dagger A = 0$.

Since the number of linearly independent harmonic r-forms in IX_6 is $b_r(\mathbf{X_6})$, the dimension of $H^r(\mathbf{X_6},\mathbf{R})$, we obtain b_r independent q-forms in the KK reduction of the 10d p-form C_p.

That is, the ansatz for the zero mode of C_p is

$$C_{\mu_1\ldots\mu_q m_1\ldots m_r}(x^0,\ldots,x^9) \;=\; \sum_{\alpha=1}^{b_r} C^a_{\mu_1\ldots\mu_q}(x^0,\ldots,x^3)\,A^a_{m_1\ldots m_r}(x^4,\ldots,x^9) \quad (12.12)$$

The 4d q-form is often written as $\int_{\Sigma_a} C_p$, where r of the indices of C_p are integrated along the r-cycle Σ_a, dual of the r-form A_a.

We would like to emphasize the fact that out of a unique 10d field we have obtained several 4d fields with same quantum numbers. This arises simply because of the existence of several zero modes for a kinetic operator in the internal space. That is, several zero energy resonance modes of a 10d field in the 6d 'cavity' given by the internal space. As we will see later on, this beautiful mechanism is a possible origin of family replication in heterotic models reproducing physics similar to the Standard Model.

12.2.3 Spectrum

We now have enough tools to directly determine the spectrum of type IIA/B compactifications on Calabi-Yau threefolds with Hodge numbers $(h_{1,1}, h_{2,1})$. We just need to recall that the number of scalars obtained from the KK reduction of the metric is $h_{1,1}$ real scalars plus $h_{2,1}$ complex scalars. These arise because the metric depends on these numbers of complex and Kahler and complex structure parameters, so the internal kinetic operator for 10d gravitons should have the corresponding zero energy directions. It is important to note that Calabi-Yau threefolds do not have isometrical direction, thus the KK reduction of the 10d metric does not lead to 4d gauge bosons. Finally, p-forms are KK reduced as above. To simplify notation we denote Σ_a the non-trivial $(1,1,)$-cycles, $\tilde{\sigma}_a$ their dual $(2,2)$-cycles, Λ_b and $\tilde{\Lambda}_b$ the $(2,1)$- and $(1,2)$-cycles, and ω, $\tilde{o}mega$ the $(3,0)$- and $(0,3)$-cycles.

Recall that the bosonic fields for 10d type IIA are the graviton G, the NSNS 2-form B, the dilaton ϕ, and the RR 1-forms A_1 and 3-form C_3

IIA		Gravity	$h_{1,1}$ Vector	$h_{2,1}$ Hyper	Hyper
G	$\to$	$g_{\mu\nu}$	$h_{1,1}$	$2h_{2,1}$	
B	$\to$		$\int_{\Sigma_a} B$		c
ϕ	$\to$				ϕ
A_1	$\to$	A_1			
C_3	$\to$		$\int_{\Sigma_a} C_3$	$\int_{\Lambda_a} C_3,\ \int_{\tilde{\Lambda}_a} C_3$	$\int_\omega C_3,\ \int_{\tilde{\omega}} C_3$

Here c is the scalar dual to the 4d 2-form $b_{\mu\nu}$, i.e. $dc = *_{4d} db$. In total, we get the $\mathcal{N}=2$ 4d supergravity multiplet, $h_{1,1}$ vector multiplets (with abelian group $U(1)^{h_{1,1}}$) and $h_{2,1}+1$ hypermultiplets (neutral under the gauge group).

The bosonic fields for 10d type IIA are the graviton G, the NSNS 2-form B, the dilaton ϕ, and the RR 0-form a, 2-form $\tilde{B}$ and 4-form C_4^+ (with self dual field strength).

IIB		Gravity	$h_{2,1}$ Vector	$h_{1,1}$ Hyper	Hyper
G	$\to$	$g_{\mu\nu}$	$2h_{2,1}$	$h_{1,1}$	
B	$\to$			$\int_{\Sigma_a} B$	c
ϕ	$\to$				ϕ
a					a
$\tilde{B}_2$	$\to$			$\int_{\Sigma_a} \tilde{B}$	$\tilde{c}$
C_4^+	$\to$	$\int_\omega C_4^+$	$\int_{\Lambda_b} C_4^+$	$\int_{\Sigma_a} C_4^+$	

Note that the self duality $dC_4 = *dC_4$ reduces the number of independent integrals of C_4^+ that can be taken.

In total, we obtain the $\mathcal{N} = 2$ 4d supergravity multiplet, $h_{2,1}$ vector multiplets (with abelian gauge group) and $(h_{1,1}+1)$ hypermultiplets (neutral under the gauge group).

12.2.4 Mirror symmetry

Consider two Calabi-Yau threefolds $\mathbf{X}$ and $\mathbf{Y}$, such that $(h_{1,1}, h_{2,1})_{\mathbf{X}} = (h_{2,1}, h_{1,1})_{\mathbf{Y}}$. Then the low energy spectrum of type IIA on $\mathbf{X}$ and type IIB on $\mathbf{Y}$ are the same.

This suggest more that a coincidence. The mirror symmetry proposal is that for each Calabi-Yau threefold $\mathbf{X}$ there exists a mirror threefold $\mathbf{Y}$ such that type IIA string theory on $\mathbf{X}$ is exactly equivalent to type IIB string theory on $\mathbf{Y}$. This of couse implies the above relation between their Hodge numbers, but much more, since the claim implies equivalence of the two theories to all orders in α', i.e. including stringy effects (there are proposal for equivalence also to all orders in the spacetime string coupling constant).

There is a lot of evidence in favour of this proposal. For instance, classification of large classes of Calabi-Yau threefols show that they appear in pairs, for each $\mathbf{X}$ there is some $\mathbf{Y}$, with the right relation of Hodge numbers. Obviously, this is necessary but not sufficient for mirror symmetry. Nevertheless it is a compelling piece of evidence.

More convincing is the explicit construction of two different Calabi-Yau geometries starting from a unique 2d interacting conformal field theory, by two different geometric interpretation of the 2d fields. See [78].

The mirror symmetry proposal has very interesting implications. It implies an exact matching of the complex structure moduli space of $\mathbf{X}$ with the Kahler moduli space of $\mathbf{Y}$ (with the Kahler parameters complexified by the addition of scalars arising from B-fields), exactly in α'. This has led to remarkable predictions in *mathematics*, as follows. A non-renormalization theorem of $\mathcal{N} = 2$ 4d supersymmetry ensures that the structure (metric) of the vector multiplet moduli space is independent of scalars in hypermultiplets, and vice versa. Recall that α' corrections are controlled by a Kahler parameter, which for type IIB(IIA) is a hypermultiplet (vector multiplet) scalar. This implies that in the compactification of type IIB on $\mathbf{Y}$ the vector multiplet moduli space, i.e. the complex structure moduli space, does not suffer α' corrections, and the result obtained in the supergravity approximation is α' exact. Mirror symmetry proposes that this is exactly the vector multiplet moduli space of type IIA on the mirror $\mathbf{X}$; this is the Kahler moduli space of $\mathbf{X}$, and it suffers from α' corrections. Mirror symmetry is giving us a tool to resum all the α' corrections to the metric in the Kahler moduli space of IIA on $\mathbf{X}$ via its equivalence with the complex structure moduli space of IIB on $\mathbf{Y}$, which is exactly computable from classical geometry in supergravity. The α' corrections on the Kahler moduli space of IIA on $\mathbf{X}$ are interesting, because a non-renormalization theorem ensures that there are no perturbative (in the α' expansion) corrections; on the other hand, there are non-perturbative (in the α' expansion) corrections, due to worldsheet instantons: these are processes mediated by configurations where the closed string wraps around a holomorphic 2-cycle in $\mathbf{X}$. Mirror symmetry allows to compute these contributions from the mirror, and to extract from this the number of holomorphic 2-cycles in the Calabi-Yau threefold $\mathbf{X}$. These numbers are very difficult to compute from other mathematical means, and easily derived from mirror symmetry. Hence mirror symmetry has attracted the attention of many mathematicians.

12.3 Compactification of heterotic strings on Calabi-Yau threefolds

In this section we study the more interesting (and difficult) compactification of heterotic theory on Calabi-Yau threefolds. They will lead to models with potential phenomenological application, in the sense that they are similar to the physics of Elementary Particles we observe in Nature.

Notice that since we work in the supergravity approximation, heterotic $SO(32)$ and type I compactifications will be very similar. Also both heterotics require the same tools for this compactification, hence (for historical reasons, and also because they lead to nicer models with the particular ansatz we make (standard embedding)), we center on compactifications of the $E_8 \times E_8$ heterotic.

12.3.1 General considerations

The original massless 10d fields of the theory are the metric G, the 2-form B, the dilaton ϕ, and the gauge bosons A^a in $E_8 \times E_8$, plus the fermion superparters of all these. We compactify the corresponding supergravity theory on $M_4 \times \mathbf{X_6}$. Clearly, the condition that we get some unbroken 4d supersymmetry, in particular some 4d gravitino, implies that $\mathbf{X_6}$ must be a Calabi-Yau threefold. We see that starting with a single 10d gravitino we will end up with a single 4d gravitino, namely the 4d theory has $\mathcal{N}=1$ supersymmetry. This is very nice, since it is a low enough degree of supersymmetry to allow for chiral fermions. On the other hand, we know that $\mathcal{N}=1$ supersymmetry is considered one of the most promising extensions beyond the Standard Model.

One difference of heterotic compactifications, compared with type II compactifications, are the presence of the 10d nonabelian gauge fields. Hence in the compactification there is the possibility of turning on a non-trivial background for their internal components $A_m(x^4,\ldots,x^9)$. More formally, we need to specify not just a compactification manifold, but also a gauge bundle (a principal G-bundle, with $G \subset E_8 \times E_8$) over the internal space $\mathbf{X_6}$. Such bundles are also constrained in order to lead to unbroken 4d susy in the gauge sector of the theory (see below).

Before discussing the bundles in more detail, let us wonder whether we really need non-trivial bundles, or else compactifications with trivial gauge bundle are consistent. The answer is that such compactifications are inconsistent if the Calabi-Yau is non-trivial (i.e. is not a six-torus). To see this, recall the Green-Schwarz terms in the 10d action, that we mentioned in the discussion of heterotic (or type I) 10d anomalies. In particular, there is a term of the form

$$\int_{10d} B_6 \wedge (\mathrm{tr}\, F^2 - \mathrm{tr}\, R^2) \tag{12.13}$$

where F and R are the curvatures of the gauge and tangent bundle, and B_6 is the dual to the NSNS 2-form, $dB_6 = *dB_2$. This leads to an action for B_6 which can be written

$$\int_{10d} H_3 \wedge dB_6 + \int_{10d} B_6 (\mathrm{tr}\, F^2 - \mathrm{tr}\, R^2) \tag{12.14}$$

where H_3 is the field strength for B_2. This leads to the equations of motion

$$dH_3 = \mathrm{tr}\, F^2 - \mathrm{tr}\, R^2 \tag{12.15}$$

Taking this equation in cohomology (both sides are closed), the left hand side is exact so corresponds to the zero class. We get

$$[\mathrm{tr}\, F^2] = [\mathrm{tr}\, R^2] \quad \text{namely} \quad c_2(E) = c_2(R) \tag{12.16}$$

the second Chern class of the gauge bundle must equal that of the tangent bundle. The latter is trivial only for the six-torus, so consistency of the equations of motion requires the internal gauge bundle to be non-trivial.

Thus we need to specify a connection in a non-trivial principal G-bundle to have a consistent compactification. The requirements on this connection in order to have unbroken 4d supersymmetry is that the curvatures obey the conditions

$$F_{ij} = 0 \quad ; \quad F_{\overline{i}\overline{j}} = 0 \quad ; \quad g^{i\overline{j}} F_{i\overline{j}} = 0 \tag{12.17}$$

Again, explicit solutions to these equations are difficult to find. However, there is a theorem (by Donaldson, Uhlenbeck and Yau) which guarantees the existence of a solution for gauge bundles satisfying the (simpler to check, since they are almost topological) conditions

i) The complexified vector bundle (with fiber given by the vector space of *complex* linear combinations of the basis vectors) is holomorphic (i.e. transition functions are holomorphic).

ii) The bundle is stable. This is a complicated to state condition, which in physics terms ensures that the gauge field configuration is stable against decay into product of bundles.

The classification or even the construction of stable holomorphic bundles over a Calabi-Yau is a difficult task even for mathematicians, so we will not say much about this.

Happily, there is a very natural gauge bundle that satisfies the above conditions, and can be used for any Calabi-Yau manifold, therefore leading to a 4d $\mathcal{N} = 1$ supersymmetric compactification. It amounts to taking the gauge bundle to be isomorphic to the tangent bundle, and the gauge connection to be the same, at each point, to the spin connection. This is called the standard embedding, or embedding the spin connection on the gauge degress of freedom.

Note that since $F = R$ it automatically satisfies the condition $c_2(F) = c_2(R)$. Also note that due to the Calabi-Yau property, the tangent bundle has holonomy $SU(3)$, so the non-trivial part of the gauge bundle is embedded in an $SU(3)$ subgroup of one of the E_8, i.e. $H = SU(3)$.

We emphasize that the standard embedding is just a possible choice of consistent gauge background in the heterotic compactification. Any other choice of bundle, with different structure group, etc, would lead to equally consistent models. In this lecture we however center on standard embedding models for simplicity.

12.3.2 Spectrum

Before entering the construction of the final 4d spectrum, recall the basic 4d $\mathcal{N}=1$ supermultiplets.

i) the gravity multiplet, containing the 4d metric and one gravitino

ii) the vector multiplet, containing the gauge bosons and the gauginos (Majorana fermions in the adjoint)

iii) the chiral multiplet, containing a complex scalar and a Weyl fermion, both in some representation of the gauge group.

With this information it will be enough to determine just the spectrum of bosons or of fermions.

The reduction of the 10d $\mathcal{N}=1$ sugra multiplet leads to the following bosonic fields in 4d

Het		Gravity	$h_{1,1}$ Chiral	$h_{2,1}$ Chiral	Chiral
G	$\to$	$g_{\mu\nu}$	$h_{1,1}$	$2h_{2,1}$	
B	$\to$		$\int_{\Sigma_a} B$		c
ϕ	$\to$				ϕ

Thus we get $h_{1,1}+h_{2,1}+1$ chiral multiplets, neutral under the gauge group.

In the compactification of the 10d $\mathcal{N}=1$ $E_8\times E_8$ vector multiplet, it is easy to identify the resulting 4d vector multiplets. This can be done by realizing that the gauge symmetries surviving in 4d are those gauge transformations in $E_8\times E_8$ which leave the background invariant. Thus the 4d gauge group is the commutant of the subgroup H with non-trivial gauge background turned on.

For the standard embedding $H=SU(3)$, embedded within one of the two E_8's. Thus, the other E_8 is untouched and survives in the 4d gauge group. About the E_8 on which we embed the $SU(3)$, the unbroken 4d gauge group by realizing that E_8 has a maximal rank subgroup $E_6\times SU(3)$ and we embed the gauge connection on the last factor. The adjoint representation of E_8 decomposes as (see below)

$$\begin{array}{ll} E_8 & \to E_6\times SU(3) \\ 248 & \to (78,1)+(1,8)+(27,3)+(\overline{27},\overline{3}) \end{array} \tag{12.18}$$

The generators commuting with $SU(3)$ must be singlets under it, so the unbroken 4d group is E_6 (times E_8).

To verify the above decomposition, recall that the generators of E_8 are 8 Cartans H^I and the non-zero roots

$$(\underline{\pm,\pm,0,0,0,0,0,0}) \quad ; \quad \frac{1}{2}(\pm,\pm,\pm,\pm,\pm,\pm,\pm,pm) \tag{12.19}$$

(with an even number of minus signs in the second set).

The decomposition (12.18) is as follows

$$\begin{array}{lll}
SU(3) & \rightarrow & H_1 - H_2 \ , \ H_1 + H_2 - 2H_3 \\
 & & (\underline{+,-,0}) \\
E_6 & & H_1 + H_2 + H_3 \ , \ H_4 \ , \ H_5 \ , \ H_6 \ , \ H_7 \ , \ H_8 \\
 & & (0,0,0,\underline{\pm,\pm,0,0,0}) \\
 & & \frac{1}{2}(+,+,+,\pm,\pm,\pm,\pm,\pm) \\
 & & \frac{1}{2}(-,-,-,\pm,\pm,\pm,\pm,\pm) \\
(27,3) & & (\underline{+,0,0},\pm,0,0,0,0,0) \\
 & & (\underline{-,-,0},0,0,0,0,0) \\
 & & \frac{1}{2}(\underline{+,-,-},\pm,\pm,\pm,\pm,\pm) \\
(\overline{27},\overline{3}) & & (\underline{-,0,0},\pm,0,0,0,0,0) \\
 & & (\underline{+,+,0},0,0,0,0,0) \\
 & & \frac{1}{2}(\underline{-,+,+},\pm,\pm,\pm,\pm,\pm)
\end{array} \tag{12.20}$$

Thus we get 4d $\mathcal{N}=1$ vector multiplets of $E_6 \times E_8$. This is very interesting, since the group E_6 has been considered as a candidate group for grand unification models. So in a sense, it is relatively close to the Standard Model (we simply point out that slightly more complicated models, with other structure group on the gauge bundle, can lead to gauge groups even closer to that of the Standard Model).

Finally, we need to discuss the spectrum of chiral multiplets. To obtain these it is more convenient to obtain the fermionic components that arise in the KK reduction of the 10d gaugino. Let us discuss the general idea of how to do this, before going to the particular case of the standard embedding. For simplicity, we center on the E_8 factor broken by the compactification, the corresponding gaugino transforms in the adjoint of the original gauge group E_8. In the breaking of the gauge group $E_8 \to H \times G_{4d}$, the adjoint of E_8 suffers a general decomposition

$$\begin{array}{ccc} E_8 & \to & H \times G_{4d} \\ 248 & & \sum_i (R_{H,i}, R_{G,i}) \end{array} \qquad (12.21)$$

The ansatz for the profile of the 10d gaugino in the KK reduction is of the form

$$\begin{aligned} \lambda_\alpha(x^0,\ldots,x^9) \;=\; & \textstyle\sum_i \Big(\xi_4^{R_{H,i}}(x^4,\ldots,x^9)\psi_{-1/2}^{R_{G,i}}(x^0,\ldots,x^3) + \\ & \xi_{\overline{4}}^{R_{H,i}}(x^4,\ldots,x^9)\psi_{1/2}^{R_{G,i}}(x^0,\ldots,x^3)\Big) \end{aligned} \qquad (12.22)$$

where ξ_4, $\xi_{\overline{4}}$ are spinors of opposite chiralities in the internal 6d and $\psi_{\pm 1/2}$ as spinors of opposite chiralities in 4d. The singlet component of ξ gives rise to the 4d gauginos.

As usual, the 4d mass of a chiral left handed 4d fermion $\psi_{-1/2}^{R_{H,i}}$ in the representation $R_{G,i}$ of G_{4d} is given by the eigenvalue of the kinetic operator on the corresponding internal wavefuncion $\xi_4^{R_{H,i}}$. This is the 6d Dirac operator for fermions in the 4, coupled to an H-bundle in the $R_{H,i}$ representation. Hence we obtain a left handed chiral 4d fermion in the $R_{G,i}$ for each solution of the equation

$$\not{D}_{R_{H,i}} \xi_4^{R_{H,i}} = 0 \qquad (12.23)$$

The number of fermions $n_{R_{G,i}}^-$ is hence the dimension of $\ker \not{D}_{R_{H,i}}$.

In general, the number of zero modes of $\not{D}_{R_{H,i}}$ is *not* given by a topological quantity of $\mathbf{X_6}$ or the bundle. The reason is that the KK reduction of the 10d gaugino can also lead to 4d right-handed chiral fermions in the representation $R_{H,i}$. The number $n_{R_{G,i}}^+$ of such zero modes is given by the dimension of $\ker \not{D}_{R_{H,i}}^\dagger$. Since two 4d chiral fermions of opposite 4d chiralities and in the same representation of the gauge group can couple to get a Dirac mass, they can disappear from the massless spectrum. This can be triggered by a small change of the geometry of the manifold or the gauge bundle, while staying in the same topological sector. Therefore, the individual numbers of massless chiral fermions $n_{R_{G,i}}^\pm$ are *not* topological. However, in all these processes of Dirac mass generation, the difference between the two two numbers is conserved. Indeed, it was known to mathematicians that the difference,

$$\text{ind}\not{D}_{R_{H,i}} = \dim \ker \not{D}_{R_{H,i}} - \dim \ker \not{D}_{R_{H,i}}^\dagger \qquad (12.24)$$

called the index of the Dirac operator (coupled to a suitable bundle) can be expressed in terms of characteristic classes of the tangent and gauge bundles

$$\text{ind}\, \not{D}_{R_{H,i}} = \int_{\mathbf{X^6}} ch(F)\sqrt{\hat{A}(R)} \tag{12.25}$$

where it is understood that one must expand the Chern character (computed in the representation $R_{H,i}$) and A-roof genus, and pick the degree 6 piece to integrate it.

This is satisfactory enough, since we expect that generically vector-like pairs of fermions pick up large masses, of the order of the cutoff scale (the string scale or compactfication scale) since there is no symmetry or principle that forbids it. Hence, the only fields that we see in Nature would be the unpaired chiral fermions (this is a version of Georgi's 'survival hypothesis')

Returning to the case of the standard embedding, we are interested in obtaining the (net) number or fermions in the 27 of E_6. Since the gauge connection is determined by the spin connection, the index theorem gives the number of such 4d fermions in terms of just the topology of $\mathbf{X_6}$. It can be shown that the index theorem gives

$$n_{27}^{-} - n_{27}^{+} = \xi(\mathbf{X}_6)/2 = h_{1,1} - h_{2,1} \tag{12.26}$$

where $\xi(\mathbf{X_6})$ is the so-called Euler characteristic of $\mathbf{X_6}$. Therefore we get this number of chiral multiplets in the 27 of E_6. This is very remarkable because in E_6 grand unification the Standard Models families arise from representaions 27, hence $\xi/2$ is the number of fermion families in this kind of compactification. As we discussed above, this is a beautiful geometric origin for the number of families, as they arise from different zero energy resonances of a 10d field in the internal space! (see figure 12.2).

This number can be quite large in simple examples. For instance, for the quintic Calabi-Yau we get a model with 100 families, far more than we would like. In any event, there exist Calabi-Yaus where this number is small, and can be even three.

Note that the KK reduction would also lead to other fields, like singlets of E_6 (arising from internal wavefunctions in the 8 of $SU(3)$). These can be obtained from the index theorem, although the topological invariants are much more difficult to compute, so we skip their discussion.

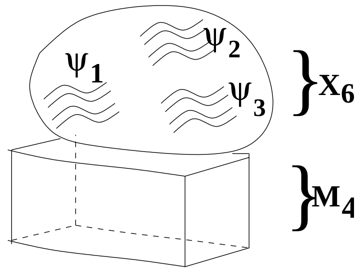


Figure 12.2: The replication of chiral families has a geometric origin in heterotic compactifications on Calabi-Yau spaces.

12.3.3 Phenomenological features of these models

Let us start by mentioning that far more realistic models have been constructed explicitly. In particular one can achieve smaller gauge groups, closer to the Standard Model one, by adding Wilson lines breaking E_6. All examples of heterotic compactification show some generic features, which can be considered as predicitions of this setup (although there exist other ways in which string theory can lead to something similar to the Standard Model, with different phenomenological features).

• The string scale must be around the 4d Planck scale. The argument is as follows. The 10d gravitational and gauge interactions have the structure

$$\int d^{10}x \, \frac{M_s^8}{g_s^2} \, R_{10d} \quad ; \quad \int d^{10}x \, \frac{M_s^6}{g_s^2} \, F_{10d}^2 \tag{12.27}$$

where M_s, g_s are the string scale and coupling constant, and R_{10d}, F_{10d} are the 10d Einstein and Yang-Mills terms. Upon Kaluza-Klein compactification on $\mathbf{X_6}$, these interactions reduce to 4d and pick up a factor of the volume V_6 of $\mathbf{X_6}$

$$\int d^4x \, \frac{M_s^8 \, V_6}{g_s^2} \, R_{10d} \quad ; \quad \int d^4x \, \frac{M_s^6 \, V_6}{g_s^2} \, F_{10d}^2 \tag{12.28}$$

From this we may express the experimental 4d Planck scale and gauge coupling in terms of the microscopic parameters of the string theory configuration

$$M_P^2 \, = \, \frac{M_s^8 \, V_6}{g_s^2} \, \simeq \, 10^{19} \, \text{GeV} \quad ; \quad \frac{1}{g_{YM}^2} \, = \, \frac{M_s^6 \, V_6}{g_s^2} \, \simeq \, \mathcal{O}(.1) \tag{12.29}$$

From these we obtain the relation

$$M_s = g_{YM}\, M_P \simeq 10^{18}\,\mathrm{GeV} \tag{12.30}$$

This large string scale makes string theory very difficult to test, since it reduces to an effective field theory at basically any experimentally accessible energy.

• This large cutoff scale makes the proton very stable, since in principle baryon number violating operators are suppressed by such large scale.

• Gauge and gravitational interactions have a similar coupling constant at the string scale, since they are controlled by the vev of the dilaton, which is universal. This is in reasonable good agreement with the renormalization group extrapolation of low energy couplings up in energy (assuming no exotic physics beyond supersymmetry in the intermediate energy region).

• The compactification scale cannot be too small. In order to avoid unobserved Kk replicas of Standard Model gauge bosons, the typical radius of the internal space should be much smaller than an inverse TeV. Other arguments about how the volume moduli modify the gauge couplings of string theory at one loop suggest that the compactification scale should be quite large to get weak gauge couplings. Usually one takes the compactification scale close to the string scale.

• The Yukawa couplings are given by the overlap integral of the internal wavefunctions of zero modes of the Dirac operator in $\mathbf{X_6}$. These are difficult to compute, in particular for the more realistic models which do not have standard embedding. So it is difficult to analyze the generic patterns of fermion masses at the string scale.

Finally, let us mention that this construction is very remarkable. We have succeeded in relating string theory with something very close to the observed properties of Elementary Particles. However, the setup has several serious problems, which are being addressed although no satisfactory solution exists for the moment

• How to break supersymmetry without generating a large cosmological constant?

• The models contain plenty of massless or very light fields, in particular the moduli that parametrize the background configuration. How to get rid of these?

• The vacuum selection problem. There is no criterion in the theory that tells us that a background is preferred over any other. Is the string theory that corresponds to our world special in any sense? Or is it a matter of chance or of anthropic issues that we see the world as it is?

Despite these open questions, we emphasize again the great achievement that we have reviewed today. We have provided a class of theories unifying gauge and gravitational interactions, and leading to 4d physics similar to the physics observed in Nature!

Chapter 13

Orbifold compactification

The basic reference for this lecture is [64]. See also [65].

13.1 Introduction

13.1.1 Motivation

We have seen that compactification on smooth Calabi-Yau spaces leads to very interesting 4d theories. However, they require quite a lot of geometrical tools, and the information one can extract is, in a sense, limited (because of the need to use the supergravity approximation (lowest order in α' expansion), and the difficulty in constructing explicit metrics, only topological quantities can be reliably obtained).

In this lecture we discuss orbifold compactifications. They share many of the features of compactification on smooth Calabi-Yau spaces (they can be regarded as compactifications on singular Calabi-Yau's), but are described by free 2d worldsheet theories. Hence, the quantization of the string can be carried out exactly in the α' expansion, and one can compute quantities explicitly, and including the stringy corrections. In this sense, orbifolds are (almost) as simple as toroidal compactifications, but have the advantage of leading to models with reduced supersymmetry. In this lecture we center on 6d orbifolds preserving 1/8 of the supersymmetries; namely i.e. leading to 4d $\mathcal{N}=2$ supersymmetry for type II theories or to 4d $\mathcal{N}=1$ supersymmetry for heterotic theories. The description of orbifolds of type I theory (also known as type IIB orientifolds) is more technical and is not discussed (left for the

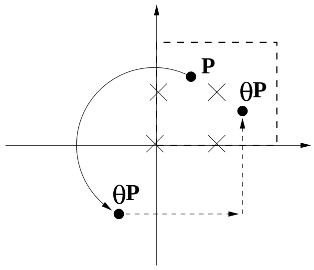


Figure 13.1: $\mathbf{T^2/IZ_2}$ orbifold. The 2-torus is shown as the 2-plane modded by discrete translations; hence the sides of the unit cell, shown in dashed lines, are identified. The rotation θ maps each point to its symmetric with respect to the origin. The action on the 2-torus is obtained by translating the points into the unit cell. Crosses represent points fixed under the action of θ on $\mathbf{T^2}$.

final projects).

13.1.2 The geometry of orbifolds

A toroidal orbifold (or just orbifold, for short) $\mathbf{T^6}/\mathbf{\Gamma}$ is the quotient space of $\mathbf{T^6}$ by a finite isometry group Γ, which acts with fixed points.

One simple example, before going to the 6d case, is the 2d orbifold $\mathbf{T^2}/\mathbf{Z_2}$. Consider a $\mathbf{T^2}$ parametrized by two coordinates x_1, x_2, with periodic identifications $x_i \simeq x_i + 1$, and consider the $\mathbf{Z_2}$ action generated by the symmetry $\theta : x_i \to -x_i$. The orbifold $\mathbf{T^2}/\mathbf{Z_2}$ is $\mathbf{T^2}$ with the identification of points related by the action of θ. This is shown in figure 13.1.

The action θ has fixed points, namely points with coordinates (x_1, x_2) equivalent to $(-x_1, -x_2)$ up to periodicities. Namely obeying

$$(-x_1, -x_2) = (x_1, x_2) + n(1, 0) + m(0, 1) \tag{13.1}$$

for some $n, m \in \mathbf{Z}$. There are four such points, with coordinates $(0, 0)$, $(0, 1/2)$, $(1/2, 0)$ and $(1/2, 1/2)$. These fixed points of the orbifold action descend to conical singularities in the quotient space. This can be seen by studying the local geometry near one of this points, which is a quotient space $\mathbf{R^2}/\mathbf{Z_2}$, and can be regarded as the space obtained by taking a piece of paper, cutting half of it, and glueing the two halves of the boundary to obtain a cone. This is shown in figure 13.2. The idea generalizes to more complicated higher-dimensional orbifolds.

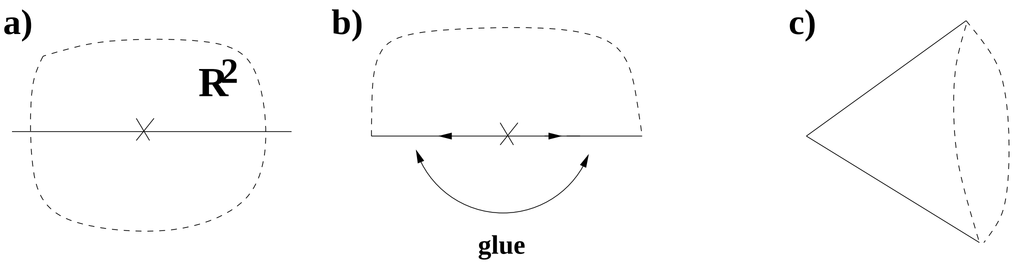


Figure 13.2: The quotient $\mathbf{R^2}/\mathbf{Z_2}$ has a conical singularity at the origin. This can be seen by starting with the 2-plane (a), keeping points in the upper half (b) (points in the lower half are their θ images, and performing the remaining θ identification in the horizontal boundary (c).

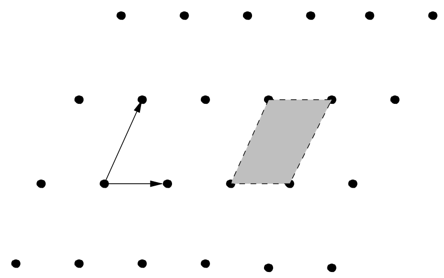


Figure 13.3: A 2d lattice, admitting a $\mathbf{Z_2}$ symmetry (reflection with respect to any point in the lattice. It is easy to cook up other 2d lattices with $\mathbf{Z_3}$ or $\mathbf{Z_4}$ symmetry.

Notice that to obtain a well-defined quotient, the discrete group must be a symmetry of the torus. This is most easily checked by regarding the d-dimensional torus as $\mathbf{R^d}$ modded out by translations in a lattice. The group Γ should be a symmetry of the lattice. Such groups are said to act crystallographically on the lattice, by analogy with crystallographic groups in solid state physics. An example of a 2d lattice is shown in figure 13.3.

A very popular example is the 4d orbifold $\mathbf{T^4}/\mathbf{Z_2}$, with the generator θ of $\mathbf{Z_2}$ acting by $x_i \to -x_i$ on the four coordinates of $\mathbf{T^4}$. The resulting quotient space is a singular limit of the Calabi-Yau space K3, with 16 singular points, locally of the form $\mathbf{R^4}/\mathbf{Z_2}$.

Clearly, one can form orbifold using other discrete groups. For instance,

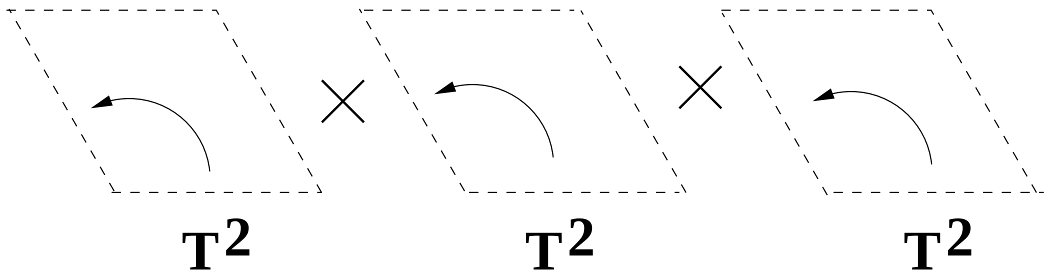


Figure 13.4: The $\mathbf{T^6}/\mathbf{Z_3}$ orbifold.

we will later on center on a 6d orbifold $\mathbf{T^6}/\mathbf{Z_3}$, where $\mathbf{T^6}$ is described by three complex coordinates z_i, with the periodic identifications $z_i \simeq z_i + 1$ and $z_i \simeq z_i + e^{2\pi i/3}$. The generator θ of $\mathbf{Z_3}$ is an order three action given by

$$\theta \; : \; (z_1, z_2, z_3) \to (e^{2\pi i/3} z_1, e^{2\pi i/3} z_2, e^{-4\pi i/3} z_3) \tag{13.2}$$

We have used $e^{-4\pi i/3}$ instead of $e^{2\pi i/3}$ for z_3 in order to stick to the convention (useful in later purposes) that the sum of the phases in the rotations add up to zero. The orbifold action is a simultaneous rotaion by 120 degrees in all three complex planes, as shown in figure 13.4. The action has 27 fixed points which are points where the coordinates z_i are either of the values 0, $(1 + e2\pi i/3)/3$, $(e2\pi i/3 + e^{4\pi i/3})/3$. Each point is locally of the form $\mathbf{C^3}/\mathbf{Z_3}$.

Although it is possible to construct orbifolds where Γ is a non-abelian discrete group, these are technically more involved and not specially illuminating. So in this lecture we center on abelian Γ, and in particular to cases $\Gamma = \mathbf{Z_N}$, generated by an action θ acting on three complex coordinates by

$$\theta \; : \; (z_1, z_2, z_3) \to (e^{2\pi i v_1} z_1, e^{2\pi i v_2} z_2, e^{2\pi i v_3} z_3) \tag{13.3}$$

with $(v_1, v_2, v_3) = (a_1, a_2, a_3)/N$ and $a_i \in \mathbf{Z}$ [1].

Orbifolds are not smooth manifolds, but are similar in many respects to manifolds. Indeed, removing the singular points they are manifolds. In fact one can define the holonomy group, and will be related to the amount of supersymmetry preserved by the compactification, just like for smooth manifolds. By parallel transporting a vector around closed loops which were closed in the torus, the holonomies generated are trivial, because the metric

[1] An additional condition $\sum_i a_i =$ even, is required for the quotient space to be spin.

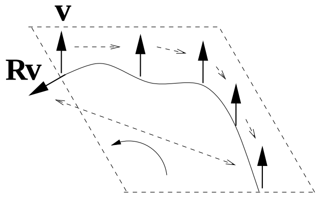


Figure 13.5: Holonomy on an orbifold: we start with the vector v and parallel transport it along a loop (closed up to the θ action); the vector ends up rotated by an action R which is isomorphic to θ.

on the torus is flat. However, there are loops in the quotient space that surround the singular points, and are closed in the quotient altough they are not closed in the 'parent' torus. The holonomies around those loops are non-trivial, and generate a holonomy group which is precisely Γ. This is shown for $\Gamma = \mathbf{Z_3}$ in figure 13.5.

This suggests that $2n$-dimensional orbifold preserving some supersymmetry should be defined by discrete groups Γ whose geometric action is in a subgroup of $SU(n)$. For 6d orbifolds with $\Gamma = \mathbf{Z_N}$ generated by the action (D.1), the condition is $v_1 \pm v_2 \pm v_3 = 0 \text{ mod } N$, for some choice of signs (the choice determines *which* susy (out of the many susys of the torus) is preserved). We will stick to orbifolds obeying the condition

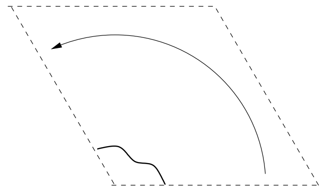
$$v_1 + v_2 + v_3 = 0 \text{ mod } N \tag{13.4}$$

These orbifolds are simple versions of Calabi-Yau manifolds.

One easily checks that the $\mathbf{T^4}/\mathbf{Z_2}$ and $\mathbf{T^6}/\mathbf{Z_3}$ examples above are supersymmetry preserving, while $\mathbf{T^2}/\mathbf{Z_2}$ is non-supersymmetric.

13.1.3 Generalities of string theory on orbifolds

One might think that a physical theory defined on an orbifold space could be singular, due to the bad geometric behaviour at the singular points. Interestingly, string theory on orbifold spaces is completely non-singular and well-behaved. This result follows from a very special set of states in string theory (twisted states), which arise due to the extended nature of strings (and would be absent in a theory of point particles).



$\mathbf{T^2/Z_3}$

Figure 13.6: Open string in a twisted sector in a $\mathbf{Z_3}$ orbifold.

To define string theory on an orbifold, we should regard the orbifold as a quotient of the torus by a symmetry. Therefore, string theory on the orbifold can be constructed by starting with string theory on the 'parent' torus, and imposing invariance under the discrete symmetry group, i.e. keeping only states which are invariant under the action of Γ (on the Hilbert space of string states). This sector is inherited from the spectrum of states in the toroidal compactification, and is called untwisted. Clearly it is described by a free 2d theory, because the metric is locally flat.

However, this is not the complete story. There exist additional closed string sectors arising from strings which are closed in the orbifold, but do not correspond to closed string in the 'parent' torus. They correspond to strings whose 2d fields have boundary conditions periodic, up to the action of some element $g \in \Gamma$, for instance

$$X(\sigma + \ell, t) = (gX)(\sigma, t) \tag{13.5}$$

this is shown in figure 13.6.

These sectors/states are known as twisted sectors/states. Notice that, these sectors are localized in the neighbourhood of fixed points, so in a sense are the sectors that carry the information that the orbifold space is not a torus, but has some curvature concentrated at those points. Note however, that the local 2d dynamics on the string is still the same as in the torus (since the inside of these strings still propagates in a flat metric), and all the non-triviality of the geometry enters simply in boundary conditions like (13.5). This remarkable feature allows to quantize the 2d theory exactly in α', although it describes propagation of strings in a non-trivial geometry. Note finally that twisted states exist because strings are extended objects, they would be absent in a theory of point particles.

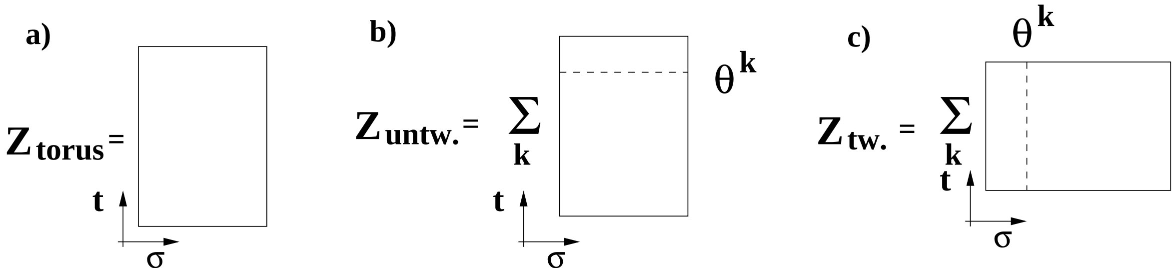


Figure 13.7: Modular invariance of string theory on orbifolds requires the existence of twisted sectors.

The complete spectrum of the string theory on the orbifold is given by the untwisted sector (states in the torus, projected onto Γ-invariant states), and twisted sectors (one per element of Γ and per fixed point of the element).

Modular invariance

We would like to make a short and qualitative comment (although the argument is also quantitatively correct) showing that twisted sectors are absolutely crucial in order to have a consistent modular invariant theory, i.e. a consistent worldsheet geometry. Hence, twisted states are crucial in maintaining the good properties of string theory (finiteness, unitarity, anomaly cancellation, etc), and making it smooth even in the presence of the singular geometry. In a sense, we may say that α' stringy effects (the very existence of twisted states) corrects the singular behaviour of the geometry and leads to smooth physics.

Let us describe the 1-loop partition function for the theory on $\mathbf{T^6}$ as a torus, parametrized by σ, t, as in figure 13.7a. In order to construct the theory on $\mathbf{T^6}/\mathbf{Z_N}$, let us insert a projector operator

$$P = \frac{1}{N}(1 + \theta + \ldots \theta^{N-1}) \tag{13.6}$$

in the t direction, which forces that only $\mathbf{Z_N}$-invariant states give a non-zero contribution to the partition function. See fig 13.7b. Since only $\mathbf{Z_N}$-invariant states propagate, this describes the partition function for the untwisted sector.

Now we can see that this contribution is not modular invariant. Let us rewrite it as a sum of contributions with insertions of θ^k in the t direction, and perform a modular transformation $\tau \to -1/\tau$, which exchanges σ and t.

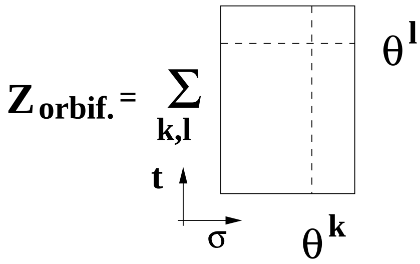


Figure 13.8: Modular invariant partition function for an orbifold.

We obtain a sum of amplitudes with insertions of θ^k in the sigma direction, see figure 13.7c. They correspond to closed strings which are periodic in σ, up to the action of θ^k; that is, they are twisted strings. Clearly, the complete modular invariant amplitude is as in figure 13.8, a sum over the untwisted and twisted sectors, with projector insertions in t to ensure that only $\mathbf{Z_N}$-invariant states propagate.

13.2 Type II string theory on $\mathbf{T^6/Z_3}$

Let us consider the above described $\mathbf{T^6/Z_3}$ orbifold, where the underlying $\mathbf{T^6}$ background is described by three complex coordinates $z_i \simeq z_i + R_i \simeq z_i + R_i e^{2\pi i/3}$, and zero NSNS B-field. Recall that the generator θ of $\mathbf{Z_3}$ acts by $\theta : z_i \to e^{2\pi i v_i}$ with $v = (1, 1, -2)/3$.

We describe the 2d worldsheet theory (in the light-cone gauge) by the following fields: Along the two real non-compact coordinates, we have 2d bosons X^2, X^3 and 2d fermions ψ^2, ψ^3; to describe the three complex dimensions in $\mathbf{T^6}$, we have 2d bosons Z^1, Z^2, Z^3 (and their conjugates $Z^{\bar{\imath}}$) and 2d fermions Ψ^1, Ψ^2, Ψ^3 (and their conjugates $\Psi^{\bar{\imath}}$). The action of θ on these 2d fields is

$$Z^i \to e^{2\pi i v_i} Z_i \quad ; \quad \Psi^i \to e^{2\pi i v_i} \Psi^i \qquad (13.7)$$

Let us consider the untwisted sector. The spectrum is obtained by simply taking the spectrum of the theory on $\mathbf{T^6}$ and keeping states invariant under the $\mathbf{Z_3}$ action. In the theory on $\mathbf{T^6}$, different sectors are labelled by the momentum and winding along the internal dimensions. For the corresponding

2d fields we have the following expansion

$$
\begin{aligned}
Z^i(\sigma,t) &= z_o^i + \frac{k_i}{R_i p^+}\, t + \frac{2\pi R_i}{\ell}\, w^i\,\sigma + \\
&= i\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \left[\frac{\alpha_n^i}{n}\, e^{-2\pi i\, n(\sigma+t)/\ell} + \frac{\tilde{\alpha}_n^i}{n}\, e^{2\pi i\, n(\sigma-t)/\ell} \right]
\end{aligned}
\qquad (13.8)
$$

where all the coefficients in the mode expansion (z_0, k, w, α's, $\tilde{\alpha}$'s) are complex, and the expansion for $Z^{\bar{i}}$ involve the complex conjugates.

The action of θ on the coefficient of the mode expansion are

$$
\begin{aligned}
z_0^i \to e^{2\pi i v_i} z_0^i \quad ; \quad k_i \to e^{2\pi i v_i} k_i \quad ; \quad w^i \to e^{2\pi i v_i} w^i \\
\alpha_n^i \to e^{2\pi i v_i} \alpha_i^n \quad ; \quad \tilde{\alpha}_n^i \to e^{2\pi i v_i}\, \tilde{\alpha}_n^i
\end{aligned}
\qquad (13.9)
$$

ans similarly for the 2d fermionic coordinates.

Untwisted states in the orbifold are obtained by taking suitable $\mathbf{Z_3}$ invariant linear combinations. For sectors of non-zero momentum and/or winding, such states are roughly of the form

$$
\mathcal{O}|k,w\rangle + (\mathcal{O}^{\theta})|\theta k, \theta w\rangle + (\mathcal{O}^{\theta^2})|\theta^2 k, \theta^2 w\rangle \qquad (13.10)
$$

where $\mathcal{O}$ is a generic sausage of operators, and superscript θ^k implies taking its image under θ^k. The zero momentum and winding sectoris not mixed with other sector by θ, so one is constrained to use only operators $\mathcal{O}$ which are directly $\mathbf{Z_3}$ invariant. The mass formula for all these states is given by the same expression as for $\mathbf{T^6}$.

We will be interested in massless states. As usual, they arise from the sector of zero momentum and winding, so the spectrum is obtained by constructing the left and right vacua, and applying left and right moving operators whose phase transformation under θ cancel each other.

Consider the massless states in the left moving NS sector. They are

State	$SO(8)$ weight	$\mathbf{Z_3}$ phase
$\psi^2_{-1/2}\|0\rangle$, $\psi^3_{-1/2}\|0\rangle$	$(0,0,0,\pm)$	1
$\Psi^i_{-1/2}\|0\rangle$	$(\underline{+},0,0,0)$	$e^{2\pi i/3}$
$\Psi^{\bar{i}}_{-1/2}\|0\rangle$	$(\underline{-},0,0,0)$	$e^{-2\pi i/3}$

The phase picked up by the different states [2] can be also described as $e^{2\pi i\, r\cdot v}$, where r is the above $SO(8)$ weight and $v = (v_1, v_2, v_3, 0)$.

For left handed states in the R sector (with GSO projection selecting the 8_C as vacuum), we have

$SO(8)$ weight	$\mathbf{Z_3}$ phase
$\frac{1}{2}(+,+,+,-)$	1
$\frac{1}{2}(-,-,-,+)$	1
$\frac{1}{2}(\underline{-,+,+},+)$	$e^{2\pi i/3}$
$\frac{1}{2}(\underline{+,-,-},-)$	$e^{-2\pi i/3}$

Performing the same computation for the right movers (with opposite GSO on the R sector, since we are working on type IIA), the massless untwisted states are

[2] This arises naturally if one bosonizes the internal 2d fermions into 2d bosons ϕ^i compactified on a lattice of $SO(8)$ weights. The phase rotation of the 2d fermions becomes a translation of the corresponding bosons, which carry a lattice momentum r.

Left ⊗ Right	$e^{2\pi i\, r\cdot v}$	$e^{2\pi i\, \tilde{r}\cdot}v$	4d field
NSNS			
$(0,0,0,\pm)\otimes(0,0,0,\pm)$	1	1	$G_{\mu\nu}$, $B_{\mu\nu}$, ϕ
$(\underline{+,0,0},0)\otimes(\underline{-,0,0},0)$	$e^{2\pi i/3}$	$e^{-2\pi i/3}$	$G_{i\bar{j}}$, $B_{i\bar{j}}$ =
$(\underline{-,0,0},0)\otimes(\underline{+,0,0},0)$	$e^{-2\pi i/3}$	$e^{2\pi i/3}$	9 cmplx scalars
NS-R			
Left ⊗ Right	$e^{2\pi i\, r\cdot v}$	$e^{2\pi i\, \tilde{r}\cdot}v$	4d field
$(0,0,0,\pm)\otimes\frac{1}{2}(+,+,+,+)$	1	1	$\psi_{\mu\alpha}$, ψ_α
$(0,0,0,\pm)\otimes\frac{1}{2}(-,-,-,-)$	1	1	4d gravitino and Weyl fermion
$(\underline{+,0,0},0)\otimes\frac{1}{2}(\underline{+,-,-},+)$	$e^{2\pi i/3}$	$e^{-2\pi i/3}$	9 spin 1/2 ...
$(\underline{-,0,0},0)\otimes\frac{1}{2}(\underline{-,+,+},-)$	$e^{-2\pi i/3}$	$e^{2\pi i/3}$	...Weyl fermions
R-NS			
Left ⊗ Right	$e^{2\pi i\, r\cdot v}$	$e^{2\pi i\, \tilde{r}\cdot}v$	4d field
$\frac{1}{2}(+,+,+,-)\otimes(0,0,0,\pm)$	1	1	$\psi_{\mu\alpha}$, ψ_α
$\frac{1}{2}(-,-,-,+)\otimes(0,0,0,\pm)$	1	1	4d gravitino and Weyl fermion
$\frac{1}{2}(\underline{-,+,+},+)\otimes(\underline{-,0,0},0)$	$e^{2\pi i/3}$	$e^{-2\pi i/3}$	9 spin 1/2
$\frac{1}{2}(\underline{+,-,-},-)\otimes(\underline{+,0,0},0)$	$e^{-2\pi i/3}$	$e^{2\pi i/3}$	Weyl fermions
RR			
Left ⊗ Right	$e^{2\pi i\, r\cdot v}$	$e^{2\pi i\, \tilde{r}\cdot}v$	4d field
$\frac{1}{2}(+,+,+,-)\otimes\frac{(}{+},+,+,+)$	1	1	Gauge boson
$\frac{1}{2}(+,+,+,-)\otimes\frac{(}{-},-,-,-)$	1	1	A_μ and
$\frac{1}{2}(-,-,-,+)\otimes\frac{(}{+},+,+,+)$	1	1	cmplx scalar
$\frac{1}{2}(-,-,-,+)\otimes\frac{(}{-},-,-,-)$	1	1	C_{123}, $C_{\overline{123}}$
$\frac{1}{2}(\underline{-,+,+},+)\otimes\frac{1}{2}(\underline{+,-,-},+)$	$e^{2\pi i/3}$	$e^{-2\pi i/3}$	9 Gauge bosons
$\frac{1}{2}(\underline{+,-,-},-)\otimes\frac{1}{2}(\underline{-,+,+},-)$	$e^{-2\pi i/3}$	$e^{2\pi i/3}$	9 $C_{i\bar{j}\mu}$

Notice that there are two 4d gravitinos, signalling $\mathcal{N}=2$ 4d supersymmetry. Recalling the structure of the corresponding supermultiplet, the above fields are easily seen to gather into the supergravity multiplet ($G_{\mu\nu}$, the two $\psi_{\mu\alpha}$ and A_μ), one hypermultiplet (the two ψ_α, ϕ and the scalar dual to $B_{\mu\nu}$), and 9 vector multiplets (scalars $G_{i\bar{j}}$, $B_{i\bar{j}}$, Weyl fermions in RNS and NSR, gauge bosons $C_{i\bar{j}\mu}$).

Let us now consider the twisted sector. As mentioned above, there is one such sector per non-trivial element in $\mathbf{Z_3}$ and per fixed point. The twisted states at each fixed point are similar, so we simply obtain 27 replicas of the content in one of them. Finally one can check that states in the θ^2 twisted sector correspond to the antiparticles of states in the θ twisted sector (it is easy to see graphically that states in oppositely twisted sectors can annihilate into the vacuum). So we just compute the latter.

In the θ twisted sector, we impose boundary conditions of the kind

$$Z^i(\sigma+\ell,t) \;=\; e^{2\pi i v_i}\, Z^i(\sigma,t) \;+\; 2\pi R_i\, n^i \tag{13.11}$$

(where n^i is a vector in the two-torus lattice Λ_i). That is, the string is closed up to the rotational and translational identification in the toroidal orbifold. Similarly for the 2d fermions. Using the general mode expansion

$$\begin{aligned} Z^i(\sigma,t) \;&=\; z_0^i + \frac{p^i}{p^+}\,t \;+\; \frac{2\pi R_i}{\ell}\,n^i\,\sigma\;+ \\ &\quad +\sum_{\nu_i}\frac{\alpha^i_{\nu_i}}{\nu_i}\,e^{-2\pi i\,\nu_i(\sigma+t)/\ell}+\sum_{\tilde\nu_i}\frac{\tilde\alpha^i_{\tilde\nu_i}}{\tilde\nu_i}\,e^{2\pi i\,\tilde\nu_i(\sigma-t)/\ell} \end{aligned} \tag{13.12}$$

(and similarly for 2d fermions) the boundary conditions impose that the zero mode sits at a fixed point

$$z_0^i = e^{2\pi i v_i} z_0^i \bmod 2\pi R_i\,\Lambda_i \tag{13.13}$$

that the momentum p^i and winding w^i vanish, and that the moddings of oscillators are shifted by $\pm v_i$. Indeed, we have the oscillators

$$\begin{array}{ccccccc} \alpha^i_{n-v_i} & ; & \tilde\alpha^i_{n+v_i} & ; & \alpha^{\bar i}_{n+v_i} & ; & \tilde\alpha^{\bar i}_{n-v_i} \\ \Psi^i_{n+\rho-v_i} & ; & \tilde\Psi^i_{n+\rho+v_i} & ; & \Psi^{\bar i}_{n+\rho-v_i} & ; & \tilde\Psi^{\bar i}_{n+\rho+v_i} \end{array} \tag{13.14}$$

with $\rho = 1/2, 0$ for NS and R fermions.

The fractional modding of the oscillator modifies the vacuum energies. In the notes on type II superstring we used the familiar regularization by an exponential, and derived the relation

$$\frac{1}{2}\sum_{n=0}^{\infty}(n+\alpha) = -\frac{1}{24}+\frac{1}{4}\alpha(1-\alpha) \tag{13.15}$$

for $\alpha \geq 0$. Vacuum energies for orbifold follow from application of this formula.

We should now construct the Hilbert space of twisted states and impose the $\mathbf{Z_3}$ projection. Centering on left movers, the mass formula is given by

$$M_L^2 = \frac{2}{\alpha'}(N_B+N_F+E_0) \tag{13.16}$$

with $E_0 = -1/6$ in the NS sector and $E_0 = 0$ in the R sector.

In the NS sector, we define the vacuum as annihilated by all positive modding oscillators, and build the Hilbert space by applying negatively modded oscillators to it (and respecting the GSO projection). In the R sector, there are no fermion zero modes in the internal directions, just in the two non-compact ones. The vacuum is two-fold degenerate, and the GSO projection selects one of them as the only massless state. At the massless level, the states are

Sector	State	Mass	$r+v$	$e^{2\pi i(r+v)\cdot r}$
NS	$\Psi^3_{-1/6}\|0\rangle$	$m^2=0$	$(\frac{1}{3},\frac{1}{3},\frac{1}{3},0)$	1
R	$A_1^+\|0\rangle$	$m^2=0$	$(-\frac{1}{6},-\frac{1}{6},-\frac{1}{6},-\frac{1}{2})$	1

where we have labelled the states by a vector $r+v$, which is useful in determining the $\mathbf{Z_3}$ phase picked up by the state [3].

Working similarly with the right moving sector (with opposite GSO in the R sector, since we are in IIA), we can construct the massless physical states

Sector	$r+v\otimes\tilde{r}-v$	$SO(2)$
NSNS	$(\frac{1}{3},\frac{1}{3},\frac{1}{3},0)\otimes(-\frac{1}{3},-\frac{1}{3},-\frac{1}{3},\frac{1}{3},0)$	0
NSR	$(\frac{1}{3},\frac{1}{3},\frac{1}{3},0)\otimes(\frac{1}{6},\frac{1}{6},\frac{1}{6},-\frac{1}{2})$	$-1/2$
RNS	$(-\frac{1}{6},-\frac{1}{6},-\frac{1}{6},-\frac{1}{2})\otimes(-\frac{1}{3},-\frac{1}{3},-\frac{1}{3},0)$	$-1/2$
RR	$(-\frac{1}{6},-\frac{1}{6},-\frac{1}{6},-\frac{1}{2})\otimes(\frac{1}{6},\frac{1}{6},\frac{1}{6},-\frac{1}{2})$	-1

It is important to recall that right movers have an opposite shift in the modding of oscillators (hence we label the states are $\tilde{r}-v$.

Together with states in the θ^2 twisted sector (antiparticles), we obtain one 4d $\mathcal{N}=2$ vector multiplet per fixed point. They give rise to independent $U(1)$ gauge symmetries (no non-abelian enhancement).

The total spectrum of type IIA theory on the $\mathbf{T^6}/\mathbf{Z_3}$ orbifold is: the 4d $\mathcal{N}=2$ gravity multiplet, one hypermultiplet and $9+27=36$ (abelian) vector multiplets.

[3] In the bosonized formulation, twisted states have momentum in a shifted lattice, so the notation $r+v$ is more natural.

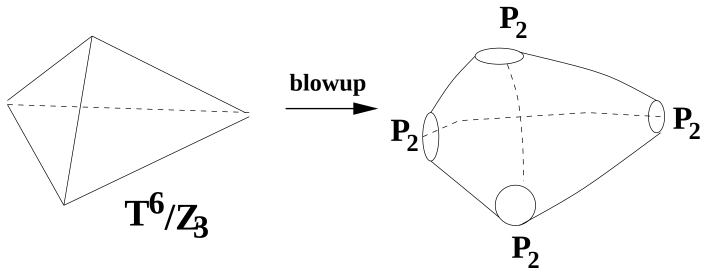


Figure 13.9: The singular orbifold $\mathbf{T^6}/\mathbf{Z_3}$ is a particular limit of a smooth Calabi-Yau in the limit in which some $\mathbf{P_2}$ 4-cycles go to zero size. The smooth Calabi-Yau is called the blowup of the orbifold.

13.2.1 Geometric interpretation

This spectrum is very much like the spectrum on a compactification on a smooth Calabi-Yau with Hodge numbers $(h_{1,1}, h_{2,1}) = (36, 0)$.

Indeed, mathematicians know that the singular space $\mathbf{T^6}/\mathbf{Z_3}$ can be regarded as a particular limit of a smooth Calabi-Yau, in the limit in which 27 4-cycles collaps to zero size (This is a singular limit in the geometric sense, but is completely smooth in string theory, due to twisted states, namely to α' effects).

In other words, the singular space $\mathbf{T^6}/\mathbf{Z_3}$ can be continuously smoothed to a non-singular space, preserving the Calabi-Yau property. This is done by the procedure known as blowing-up the singular point; roughly, this amounts to removing the 27 singular points of the orbifold and replacing them by a suitable 4-cycle, which for the singularities at hand must be a $\mathbf{P_2}$, the two (complex) dimensional projective space [4]. see figure 13.9. The resulting space is Kahler and has vanishing first Chern class, so it admits a $SU(3)$ holonomy metric. The smooth spaces are characterized by moduli which control the size of the $\mathbf{P_2}$'s, so the singular orbifold is geometrically recovered at the point of moduli space corresponding to zero sizes. Of course this limit is beyond the reach of the supergravity approximation, which is not valid for so small lentgh scales. Happily, the singular limit is nice enough so that we can quantize string theory exactly in that regime.

The homology of the resulting smooth space can be computed as follows:

[4] This is the set of points $(z_1, z_2, z_3) \in \mathbf{C^3}$ with the identification $(z_1, z_2, z_3) \simeq \lambda(z_1, z_2, z_3)$ with $\lambda IC - \{0\}$.

Before blowing up the homology was given by the homology of cycles in $\mathbf{T^6}$ invariant under the $\mathbf{Z_3}$ (what mathematicians call the equivariant homology) which leads to Hodge numbers $(h_{1,1}, h_{2,1}) = (9,0)$. To these we must add the homology of cycles associated to the $\mathbf{P_2}$'s, which appear after blowup. Each $\mathbf{P_2}$ has one 2-cycle and no 3-cycle inside it, so their contribution to Hodge numbers is $(27,0)$. Therefore we see that the homology of the smooth blowup $\tilde{\mathbf{T}}^6/\mathbf{Z_3}$ is $(36,0)$.

Thus, string theory is clever enough to 'know' that the singular orbifold belongs to a continuous family of smooth spaces with Hodge numbers $(36,0)$, and thus gives the right spectrum in the orbifold space.

The above geometric interpretation allows a geometric interpretation for the twisted sector fields in string theory. Indeed, denoting Σ the 2-cycle inside the collapsed $\mathbf{P_2}$ at each singularity, we interpret: the two real scalars correspond to the geometrical size of $\mathbf{P_2}$ (i.e. a metric modulus) and to $\int_\Sigma B$; the gauge boson corresponds to $\int_\Sigma C_3$.

It is important to emphasize that the philosophy of the geometric interpretation of the orbifold spaces also exists for other orbifolds (although the cycles arising upon blowing up are in general more involved). It is in this precise sense that orbifolds are very similar to Calabi-Yau spaces (in fact, they are CY's at a particular point in moduli space) but far more tractable.

13.3 Heterotic string compactification on $\mathbf{T^6/Z_3}$

13.3.1 Gauge bundles for orbifolds

Compactification of heterotic string on orbifolds is very similar to type II. The main difference is that noew we have the left moving internal bosons X^I, and we have the freedom of choosing a non-trivial action of Γ on them. For $\Gamma = \mathbf{Z_N}$ a simple choice is to require that the generator θ acts as a shift [5] $X^I \to X^I + V^I$, where NV is a vector in the internal 16d lattice $\Lambda_{\text{int.}}$.

Using the relation of $\mathbf{T^6/Z_3}$ with the singular limit of a smooth Calabi-Yau threefold, the above embedding of $\mathbf{Z_N}$ in the gauge degrees of freedom corresponds, from the geometric viewpoint, to using a non-trivial gauge bun-

[5] One may thing that it is more natural to use a rotation of the X^I instead of the above shift. In fact, both options are related by conjugation of the rotation to the Cartan subalgebra. More manifestly, the shift in the bosonic coordinates is equivalent to a rotation of the 32 internal fermions in the fermionic description of the heterotic.

dle in the compactification. In fact, just as for Calabi-Yau compactification, it is not consistent to choose $V = 0$. Indeed, modular invariance imposes the constraint

$$N(V^2 - v^2) = \text{even} \tag{13.17}$$

(this arises from requiring invariance under $\tau \to \tau + N$, which imposes a constraint on the contributions from the unpaired right moving fermions and left moving internal bosons).

A natural choice of gauge shift, although there exist other consistent ones, is to take V to be a copy of v. For instance, we center on the $\mathbf{Z_3}$ orbifold of the $E_8 \times E_8$ heterotic string, so we take

$$V = (\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}, 0, 0, 0, 0, 0) \times (0, 0, 0, 0, 0, 0, 0, 0) \tag{13.18}$$

Clearly this is the equivalent of the standard embedding which we studied for smooth Calabi-Yau threefolds.

13.3.2 Computation of the spectrum

The computation of the spectrum is easy as for the type II theories. In the untwisted sector we need to take the states of the theroy on $\mathbf{T^6}$ and keep those invariant under $\mathbf{Z_3}$. In heterotic theory the only additional ingredient is to realize that states with internal momentum P^I pick up a phase $e^{2\pi i\, P\cdot V}$ under the action of θ. At the massless level, we have the following massless right and left moving states

Right

	r	$e^{2\pi i\, r\cdot v}$		r	$e^{2\pi i\, r\cdot v}$
NS	$(0,0,0,\pm)$	1	R	$\frac{1}{2}(+,+,+,-)$	1
	$(\underline{+,0,0},0)$	$e^{2\pi i/3}$		$\frac{1}{2}(-,-,-,+)$	1
	$(\underline{-,0,0},0)$	$e^{-2\pi i/3}$		$\frac{1}{2}(\underline{-,+,+},+)$	$e^{2\pi i/3}$
				$\frac{1}{2}(\underline{+,-,-},-)$	$e^{-2\pi i/3}$

Left

State	θ phase	State $\vert P\rangle$	$e^{2\pi i\, P\cdot V}$
$\alpha^2_{-1}\vert 0\rangle$	1	E_8'	1
$\alpha^3_{-1}\vert 0\rangle$	1	$E_6 \times SU(3)$	1
$\alpha^i_{-1}\vert 0\rangle$	$e^{2\pi i/3}$	$(3, 27)$	$e^{2\pi i/3}$
$\alpha^{\bar{i}}_{-1}\vert 0\rangle$	$e^{-2\pi i/3}$	$(\overline{3}), \overline{27})$	$e^{-2\pi i/3}$
$\alpha^I_{-1}\vert 0\rangle$			

The decomposition of the $E_8 \times E_8$ roots with respect to the $E_6 \times SU(3) \times E_8$ is exactly as in the lecture on Calabi-Yau compactification, from which the phases $e^{2\pi i\,P\cdot V}$ are easily obtained.

Glueing the left and right moving states in a $\mathbf{Z_3}$ invariant fashion we get

Sector	State	4d Field
NS	$(0,0,0,\pm)\otimes\alpha_{-1}^{2,3}\|0\rangle$	$G_{\mu\nu}$, $B_{\mu\nu}$, ϕ
	$(0,0,0,\pm)\otimes[\,E_6\times SU(3)\times E_8'\,]$	Gauge bosons
	$(\underline{+,0,0},0)\otimes[(\overline{3},\overline{27})]$	Complex scalars
	$(\underline{-,0,0},0)\otimes[(3,27)]$	Complex scalars
R	$\pm\frac{1}{2}(+,+,+,-)\otimes\alpha_{-1}^{2,3}\|0\rangle$	4d gravitino, Weyl spinor
	$\pm\frac{1}{2}(+,+,+,-)\otimes[\,E_6\times SU(3)\times E_8'\,]$	Gauginos
	$\frac{1}{2}(\underline{-,+,+},+)\otimes[(\overline{3},\overline{27})]$	Weyl spinors
	$\frac{1}{2}(\underline{+,-,-},-)\otimes[(3,27)]$	Weyl spinors

In total, we get the 4d $\mathcal{N}=1$ supergravity multiplet, vector multiplet with gauge group $E_6\times SU(3)\times E_8'$, one neutral chiral multiplet, and 3 chiral multiplets in the $(3,27)$. Note that the spinors in the conjugate representation have also opposite chirality, so they are their antiparticles.

In the θ twisted sector, the only new ingredient is that the 16d internal momenta P are shifted by V. This follows from the boundary conditions for the internal coordinates in a twisted sector

$$X_L^I(\sigma+t+\ell) = X^I(\sigma+t) + P^I + V^I \tag{13.19}$$

(with P^I is a winding/momentum in $\Lambda_{\rm int.}$. Upon imposing it on the corresponding mode expansion

$$X_L(\sigma+t) = \frac{P_\theta^I}{2p^+} + i\sqrt{\frac{\alpha'}{2}}\sum_\nu \alpha_n^I\, e^{-2\pi i\, n(\sigma+t)/\ell} \tag{13.20}$$

we obtain the promised relation $P_\theta^I = P^I + V^I$, and the oscillators are integer-modded.

The right moving sector behaves as in type II. The left-moving spacetime mass is

$$M_L^2 = \frac{2}{\alpha'}\,(N_B + \frac{(P+V)^2}{2} + E_0) \tag{13.21}$$

with $E_0 = -1 + 3\times\frac{1}{2}\frac{1}{3}\frac{2}{3} = -\frac{2}{3}$.

We have the right moving massless states

Sector	$r+v$	$e^{2\pi i(r+v)\cdot r}$
NS	$(\frac{1}{3},\frac{1}{3},\frac{1}{3},0)$	1
R	$(-\frac{1}{6},-\frac{1}{6},-\frac{1}{6},-\frac{1}{2})$	1

The massless left moving massless states are

Osc.	P	$P+V$
$N_B=0$	$(-,-,0,0,0,0,0,0)$	$(-\frac{2}{3},-\frac{2}{3},-\frac{2}{3},0,0,0,0,0)$
	$(0,0,+,\underline{\pm,0,0,0,0})$	$(\frac{1}{3},\frac{1}{3},\frac{1}{3},\underline{\pm,0,0,0,0})$
	$\frac{1}{2}(-,-,+,\pm,\pm,\pm,\pm,\pm)$	$(-\frac{1}{6},-\frac{1}{6},-\frac{1}{6},\pm\frac{1}{2},\pm\frac{1}{2},\pm\frac{1}{2},\pm\frac{1}{2},\pm\frac{1}{2})$
$N_B=1/3$	$(0,0,0,0,0,0,0,0)$	$(\frac{1}{3},\frac{1}{3},-\frac{2}{3},0,0,0,0,0)$
	$(\underline{-,0},+,0,0,0,0,0)$	$(\underline{-\frac{2}{3},\frac{1}{3}},\frac{1}{3},0,0,0,0,0)$

where we have ignored the momentum in the second E_8 piece of $\Lambda_{\text{int.}}$ since it is zero for all these states.

The $N_B = 0$ states transform in the representation $(1, 27)$ under $SU(3) \times E_6$. All of them pick up a phase $e^{2\pi i(P+V)\cdot V} = 1$. The states with $N_B = 1/3$ transform in the representation $(\overline{3}, 1)$ under $SU(3) \times E_6$. There are three of them correspoding to the oscillators $\alpha^1_{-1/3}$, $\alpha^2_{-1/3}$ and $\alpha^3_{-1/3}$. They also pick up a total phase 1 under θ, with the oscillator phase compensating the phase from the internal momentum.

Glueuing the left and right moving states is now straightforward. The result is one chiral multiplet in the $(1, 27)$ and three chiral multiplets in the $(\overline{3}, 1)$ per fixed point. Note that the θ^2 sector contains the antiparticles of these.

In total, the massless spectrum is given by the 4d $\mathcal{N} = 1$ supergravity multiplet, $E_6 \times SU(3) \times E_8'$ vector multiplets, the dilaton chiral multiplet, and the following charged chiral multiplets

$$3(3, 27; 1) + 27(1, 27; 1) + 27 \times 3\,(\overline{3}, 1; 1) \tag{13.22}$$

This is remarkable, since it corresponds to an E_6 grand unificiation theory with 36 fermion families. Although not realistic, it is remarkable that we can obtain an explicit construction of string theory models with features similar to those of the Standard Model.

There an important point we would like to mention. Notice that $SU(3)$ has potential chiral anomalies (E_6 is always automatically non-anomalous).

The anomalies however vanish because the spectrum contains as many chiral multiplets in the 3 as in the $\overline{3}$. Note that for this to be true it is essential that twisted sectors are included in the theory! Hence this is a simple example where we see that string theory requires the presence of twisted sectors for consistency. Incidentally, we point out that the story of anomaly cancellation in 4d is even richer in models with $U(1)$ factors in the gauge group, since mixed anomalies involve a 4d version of the Green-Schwarz mechanism. we leave this discussion for the interested reader.

Notice that the above spectrum is roughly (looking just at the number of E_6 representations) that corresponding to compactification on a smooth Calabi-Yau with Hodge numbers $(36, 0)$ and gauge bundle specified by the standard embedding. This agrees with the geometric interpretation of $\mathbf{T^6/Z_3}$ we described in type II. It is interesting to notice that in this case the fields in twisted sectors that correspond to resolving the singularity are the states with $N_B = 1/3$. They not only blow up the singularities but also deform the gauge bundle (and break the gauge factor $SU(3)$). On the other hand, the states with $N_B = 0$ correspond to deformations of the gauge bundle (break the gauge group) preserving the singular geometry (these states do not carry any index of the internal space). See [66] for a nice discussion of moduli space of local versions of this orbifold.

13.3.3 Final comments

In conclusion, we see how easily and systematically one can construct compactifications on orbifolds. These have the advantage that they allow explicit string theory models, exact in α', while keeping the rich and interesting dynamics of reduced supersymmetry.

These constructions have many advantages:

• The low energy effective action is computable including α' corrections, which include the effects of massive string states. This kind of corrections can be important, for instance, in the computation of threshold effects to the unification of gauge coupling constants.

• The classification and construction of heterotic models is very systematic (and easy to program on a computer), hence allows for searching phenomenologically interesting models.

• There are many generalizations of the basic construction we have described: inclusion of Wilson lines, other orbifold groups. A less intuitive extension is that of asymmetric orbifolds [67], where one considers modding the left and right movers with different orbifold action, being careful to ensure modular invariance. These have the interesting feature that many moduli are frozen at fixed values (typically corresponding to self-dual points with respect to the T-duality group). They are however too technical to be discussed here.

The lesson to take home is that orbifolds allow to construct compactifications of full-fledged string theory (and not just supergravity) with interesting features, even close to those of Particle Physics.

Chapter 14

Non-perturbative states in string theory

Some useful references for this lecture are [86, 104, 70].

14.1 Motivation

We have studied the main properties of string theory within the framework of perturbation theory. We have uncovered very interesting formal properties of the theory, and potential applications for model building of unified theories of gauge and gravitational interactions.

In the following lectures we start reviewing several results of the recent years on the structure of string theory beyond perturbation theory. This is important i) to obtain information perhaps eventually leading to a non-perturbative formulation of string theory, and ii) to determine non-perturbative effects which may be important even at weak coupling.

In particular in this lecture we describe certain important non-perturbative states in string theory (the so-called p-branes), their properties, and their implications for string theory at the non-perturbative level (for instance, duality properties, etc).

14.2 p-branes in string theory

Non-perturbative states are states in the theory which do not have a perturbative description, i.e. they do not correspond to oscillation states of the string. Given that there is no definition of string theory beyond perturbation theory, the main question is how to look for non-perturbative states.

The main tool to do so is to use the low energy effective theory to construct them. The form of the supergravity effective actions, for large enough number of supersymmetries, is fixed by supersymmetry up to some order in the number of derivatives. Therefore it is valid even at finite coupling, if the energy densities involved are not too large (low energies). We can thus construct field configurations solving the supergravity equations of motion, with the structure of a localized core and asymptoting to flat space. These solutions describe classical excitations over the vacuum of the theory, which is given by flat space. It is useful to regard them as the field background created by a source sitting at the core of the solution. Unfortunately, supergravity is just an effective theory, and is clearly not enough to provide us with a microscopic description of these objects.

First there is the approximation of taking the lowest order in $alpha'$. Solutions will be reliable when the curvature lengths are larger than the string length. Second, there is the approximation of describing the solutions at leading order in g_s. However, some reliable information can be extracted from supergravity for some particular classes of solutions. This is the topic of this lecture.

In particular we will center on solutions which preserve some supersymmetry (and correspond to the so-called BPS states), and on properties of the solutions which are protected by supersymmetry. Before entering this discussion, let us describe the different kinds of objects we will deal with.

Detour on q-form gauge fields and charges

To describe them in a unified way, it will be useful to introduce, for each $(p+1)$-form field C_{p+1} in the theory, with field strength $(p+2)$-form H_{p+2}, the corresponding dual $(7-p)$-form C_{7-p} with field strength $(8-p)$-form H_{8-p}, defined by $H_{8-p} = *H_{p+2}$.

An object with p spatial dimensions sweeps out a $(p+1)$-dimensional subspace W_{p+1} of spacetime as it evolves in time. Such object is said to be electrically charged under C_{p+1} if the theory contains a coupling $Q\int_{W_{p+1}} C_{p+1}$. The terms containing C_{p+1} in the action are

$$\int_{10d} H_{p+1}\wedge *H_{p+1} + Q\int_{W_{p+1}} C_{p+1} = \int_{10d} C_{p+1}\wedge d*H_{p+1} + Q\int_{10d} C_{p+1}\wedge \delta(W_{p+1}) \quad (14.1)$$

where $\delta(W_{p+1})$ is the Poincare dual to the cycle W_{p+1}, bump $(9-p)$-form with support on W_{p+1}. The equation of motion reads

$$dH_{8-p} = Q\delta(W_{p+1}) \tag{14.2}$$

This implies that the flux of H_{8-p} around a $(8-p)$-sphere surrounding the object in the transverse $(9-p)$-dimensional space is

$$\int_{S^{8-p}} H_{8-p} = \int_{B^{9-p}} dH_{8-p} = Q \int_{B^{8-p}} \delta(W_{p+1}) = Q \tag{14.3}$$

where B^{9-p} is the interior of the $(8-p)$-dimensional sphere. Similarly, an object with $(7-p)$-dimensional volume W_{7-p} is charged magentically under C_{p+1} if it satisfies

$$\int_{S^{p+2}} H_{p+2} = Q' \tag{14.4}$$

Notice that this implies that the object couples electically to the dual potential C_{7-p}.

14.2.1 p-brane solutions

The main examples of elementary [1] are the D-branes, the NS fivebranes, and the fundamental strings.

The Dp-brane

This solution exists in type IIB theory for p odd, in type IIA theory for p even, and in type I theory for $p = 1, 5$; this kind of solution does not exist for heterotic theories.

The solutions (see section 14.8 in [71]) have the form (for $p \leq 6$, so as to have flat space asymptotics)

$$\begin{aligned} ds^2 &= Z(r)^{-1/2}\eta_{\mu\nu}dx^\mu dx^\nu + Z(r)^{1/2}dx^m dx^m \\ e^{2\phi} &= Z(r)^{(3-p)/2} \\ Z(r) &= 1 + \frac{\rho^{7-p}}{r^{7-p}} \quad ; \quad \rho^{7-p} = g_s Q \alpha'^{(7-p)/2} \\ H_{8-p} &= \frac{Q}{r^{(8-p)}} d(vol)_{S^{8-p}} \end{aligned} \tag{14.5}$$

where $\mu = 0, \ldots, p$, $m = p+1, \ldots, 9$, $r = \sum_m |x^m|^2$, and $d(vol)_{S^{8-p}}$ is the volume form of the $(8-p)$-sphere of unit volume.

[1] in the sense that they carry charge under just one p-form field

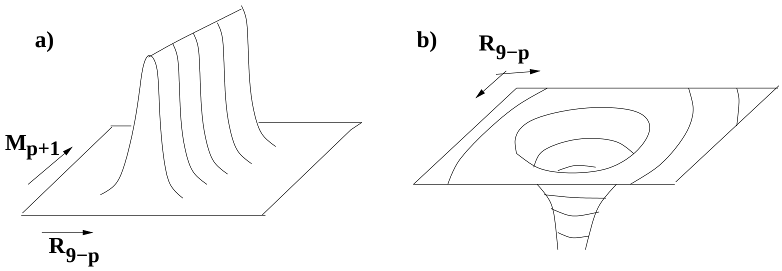


Figure 14.1: Two pictures of the p-brane as a lump of energy. The second picture shows only the transverse directions, where the p-brane looks like point-like.

The above solution has a core given by a flat $(p+1)$ dimensional plane at $r=0$ and asymptotes to flat 10d space. See figure 14.1. The core describes an object electrically charged under the RR field C_{p+1}, with charge proportional to Q. This is very remarkable, since there is no perturbative state in string theory charged under RR fields.

It is possible to compute the tension and charge using standard ADM techniques in gravitational systems, and get the result

$$T_p^2 = \frac{\pi}{g_s^2 \kappa_{10}^2} (4\pi^2\alpha')^{3-p} \quad ; \quad \mu_p^2 = \frac{\pi}{\kappa_{10}^2} (4\pi^2\alpha')^{3-p} \tag{14.6}$$

Notice that the tension is inversely proportional to the string coupling, so the state is non-perturbative, and is often referred to as soliton.

The solution is invariant under half of the supersymmetries of the vacuum of the theory. It described a so-called BPS state. This implies the particular relation between the tension and charge of the object, as we discuss below.

The fluctuations of the supergravity fields around the soliton background contain a sector of fluctuations which are localized on the $(p+1)$-dimensional volume of the soliton core. Since the soliton leaves 16 unbroken supersymmetries, these fluctuations must arrange into supermultiplets of the corresponding $(p+1)$-dimensional supersymmetry. In fact, for Dp-branes in type II theory, they form a $U(1)$ vector multiplet of 16 susys in $(p+1)$-dimensions (e.g. for a type IIB D3-brane, a vector multiplet of 4d $\mathcal{N}=4$ supersymmetry); this contains a $U(1)$ gauge boson, $(9-p)$ real scalars, and a set of fermion superpartners. On the other hand, for type I D-branes, the spectrum of fluctuations is more complidated and will be discussed in later lectures, using a simpler microscopic description.

These fluctuations localized on the soliton volume can be thought of as field living on the brane world-volume. Moreover, their dynamics is related to the dynamics of the soliton. For instance, the scalars on the brane volume are goldstone bosons of translational symmetries of the vacuum, broken by the presence of the soliton. As such, the vevs of these $(9-p)$ scalars parametrize the location of the brane in transverse $(9-p)$-dimensional space. A fluctuation leadint to non-constant profile for these scalars describes a fluctuation where the brane volume is no longer flat. The low energy effective action of these $(p+1)$-dimensional fields (which is basically the Maxwell action and kinetic terms for the scalars and fermions) is an effective action for the dynamics of the brane.

There exist also multi-soliton solutions, where the field configuration has several cores, localized at different positions x_a^m in the transverse space. The interactions between the different soliton cores cancel as a consequence of the BPS conditions, namely the gravitational attraction cancels agains their 'Coulomb' repulsion due to their (equal sign) RR charges. Thus these static configurations are solutions of supergravity. They are described by a bakcground (14.5), with

$$Z(r) \;=\; 1+\sum_a \frac{\rho^{7-p}}{|x^m - x_a^m|^{7-p}} \tag{14.7}$$

and a more complicated form for H_{8-p}, with the property that integrated over any $(8-p)$ sphere surrounding $x^m = x_a^m$ gives Q.

The analysis of certain properties (e.g. the analysis of fluctuations around the soliton background) of these multisoliton configurations is reliable only if the inter-soliton distances are larger than the string length.

We would like to conclude by emphasizing that at weak coupling there exists a microscopic description for Dp-branes, which will be the topic of next lecture. The above facts and many other will be derived from this microscopic description.

The NS5-brane

This 5-brane solution exists for type IIA and type IIB theories, and also for heterotic theories; type I theory does not contain such states.

For type II theories, the solution (see page 182 in [71]) is of the form

$$ds^2 \;=\; \eta_{\mu\nu}dx^\mu dx^\nu + Z(r)dx^m dx^m$$

$$\begin{aligned} e^{2\phi} &= Z(r) = g_s^2 + \frac{Q}{2\pi^2 r^2} \\ H_3^{NSNS} &= *_{6789} d\phi \end{aligned} \tag{14.8}$$

For heterotic theories, it has a similar expression, since the background does not excite the 10d gauge fields.

The solution describes a $(5+1)$-dimensional core, namely a 5-brane. It is electrically charged under the NSNS 6-form dual to the NSNS 2-form. Namely it is magnetically charged under the latter. The tension and charge of the object can be computed to be

$$T_{NS5} = \frac{2\pi^2\alpha'}{g_s^2\kappa_{10}^2} Q \quad ; \quad Q_{NS5} = \frac{2\pi^2\alpha'}{\kappa_{10}^2} Q \tag{14.9}$$

The solution is invariant under half of the supersymmetries of the vacuum, and so describes a BPS state. This implies the above manifest relation between the tension and charge of the object.

The spectrum of fluctuations localized on the brane volume fill out supermultiplets under the unbroken supersymmetries. For the type IIA NS5-brane, they form a 6d $\mathcal{N} = (2,0)$ tensor multiplet (containing a 2-form with 6d self-dual field strength, 5 real scalars, and 2 Weyl fermions); for the type IIB NS5-brane, they form a 6d $\mathcal{N} = (1,1)$ vector multiplet (containing a gauge boson, 4 real scalars and 2 Weyl fermions); for the $E_8 \times E_8$ heterotic, they form a 6d $\mathcal{N} = 1$ tensor multiplet (containing a self-dual 2-form, 1 scalar and 1 Weyl fermion) and hypermultiplet (containing 4 scalars and one Weyl fermion); for the $SO(32)$ heterotic, one 6d $\mathcal{N} = 1$ vector multiplet (with one gauge boson, and one Weyl fermion), one neutral hypermultiplet and 29 hypermultiplets charged under the 10d gauge group (this will more easily determined in later lectures).

Other properties of the solution are analogous to those of D-brane. For instance, the existence of multi soliton solutions, or the interpretation of fluctuations as 6d fields describing the dynamics of the brane. An important difference, however, is that there is no known microscopic description for NS5-branes at weak coupling. One intuitive explanation of this is that the effective couplig constant $g_{eff} = e^{\phi}$ grows at the core of the soliton, no matter how small the asymptotic coupling g_s is.

Fundamental string

In addition to the above objects, there exist supergravity solutions preserving half of the supersymmetries, and describing 1-branes electrically

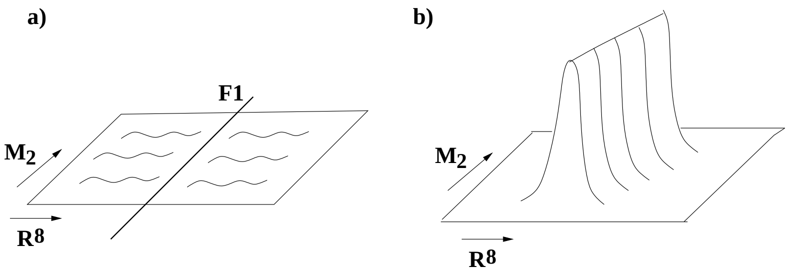


Figure 14.2: An infinitely extended fundamental string is a source for supergravity fields. The field configuration it excites is a solution of the supergravity equations of motion, which corresponds to the 1-brane like configuration. The two are simply different descriptions of the same object.

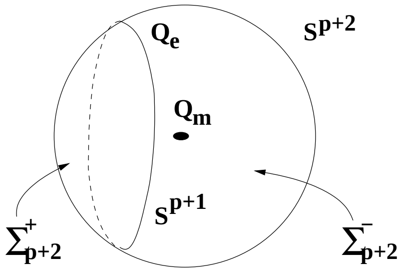
charged under the NSNS 2-form, and with tension $T_{F1} = (2\alpha')^{-1}$. This object is not non-perturbative, and has the same properties as a fundamental string with infinitely extended flat worldsheet. The natural proposal is that the supergravity solution is providing the field configuration excited by a large macroscopic fundamental string, so does not correspond to a new object. In this sense, the fundamental string is providing a microscopic description of the object we found in the 'rough' approximation of supergravity. See fig 14.2.

This object exists for type IIA, type IIB and heterotic theories. The reason why type I theory does not have a fundamental string sugra solution is that the type I string is not a BPS state. In fact, BPS states are necessarily stable, while the type I string can break.

14.2.2 Dirac charge quantization condition

Following an analysis similar to the discussion in section .1.1, we can show that in a quantum theory the electric and magnetic charges under a $p+1$ form C_{p+1} must satisfy a Dirac quantization condition.

Consider a p-brane charged electrically under C_{p+1}, i.e. the theory contains a term $Q_e \int_{W_{p+1}} C_{p+1}$ in the action. In the presence of a $(6-p)$-brane coupling magnetically under C_{p+1}, the flux of the dual field strength H_{p+2} over an $(p+2)$-sphere surrounding the $(6-p)$-brane in the transverse $(p+3)$-



dimensional space is

$$\int_{\mathbf{S^{p+2}}} H_{p+2} = Q'_m \tag{14.10}$$

Wrapping the p-brane over a $\mathbf{S^{p+1}}$ in the equator of the above $\mathbf{S^{p+2}}$, see figure 14.2.2, the phase in the path integral can be written as an intergral of H_{p+2} over a hemisphere. The change in the phase depending on which hemisphere one chooses is

$$Q_e \Delta \int_{\mathbf{S^{p+1}}} C_{p+1} = Q_e \left(\int_{\Sigma^+_{p+2}} H_{p+2} - \int_{\Sigma^-_{p+2}} H_{p+2} \right) = Q_e \int_{\mathbf{S^{p+2}}} H_{p+2} = Q_e Q'_m \tag{14.11}$$

In order to have a well-defined phase, we then need

$$Q_e Q'_m \in 2\pi \mathbf{Z} \tag{14.12}$$

If the theory contains dyonic objects, carrying electric and magnetic charges at the same time, consistency requires

$$Q_e Q'_m - Q_m Q'_e \in 2\pi \mathbf{Z} \tag{14.13}$$

At the level of supergravity these conditions are not visible. However, they should follow form any consistent microscopic description of these solitons (see lecture on D-branes). And they should hold in any consistent quantum theory, so we explicitly require them to hold in our theories.

14.2.3 BPS property

In analogy with the discussion in the field theory setup in section .1.2, the 10d supersymmetry algebras of the different string theories can be seen to admit extensions by central charges, which is this case are tensorial. The supersymmetry algebras have the structure

$$\{Q_\alpha^A, Q_\beta^{B\,\dagger}\} = -2\delta^{AB} P_\mu \Gamma^\mu_{\alpha\beta} - 2i Z^{AB}_{\mu_1\ldots\mu_{p+1}} (\Gamma^{\mu_1} \ldots \Gamma^{\mu_{p+1}})_{\alpha\beta} \qquad (14.14)$$

The operators $Z_{\mu_1\ldots\mu_{p+1}}$ are central charges, in the sense that they commute with the Q's and P_μ's, but behave as tensors with respect to the generators of the Lorentz group. They commute with the hamiltonian, hence are moduli-dependent multiples of the $(p+1)$-brane charge.

In a sector where just one of these central charges is non-zero, one can go to the rest frame of the corresponding state and derive a BPS bound for the tension of the corresponding p-brane object. Also, BPS states, i.e. states saturating the bound, belong to short representations of the supersymmetry algebra. This implies that they cannot cease to be BPS under continuous deformations of the theory, and also that the dependence of their tension with the moduli is exactly determined from the classical result (does not change by quantum corrections or otherwise).

The p-brane states studied in section (14.2.1) are BPS states, in this sense. This guarantees that, although they were constructed in the supergravity approximation, they exist in the complete theory (once α' and g_s corrections are included), and their properties, charge and tension are exactly knonw as function of the moduli.

Going through the list of string theories and brane states, the conclusion is that for any string theory, the theory contains states charged under all p-form gauge fields and their duals. These states have tension controlled by their charges, and are guaranteed to be stable (since there is no lighter state carrying those charges (it would violate the BPS bound)).

14.3 Duality for type II string theories

In this section we scratch the surface of the implications of the existence of these states in string theory. The main implication we would like to explore here is the existence of duality relations in string theory, which are analogous to the field theory duality in section .1.3. Our discussion is not complete, but just inspirational. We will return to the issue of duality in latter lectures.

14.3.1 Type IIB $SL(2,\mathbf{Z})$ duality

Ten-dimensional Type IIB supergravity has a classical $SL(2,\mathbf{R})$ invariance. It acts on the NSNS and RR 2-forms B, $\tilde{B}$ and the complex coupling $\tau = a + ie^{-\phi}$ (which takes values in the coset $SL(2,\mathbf{R})/\mathbf{U(1)}$ as

$$\begin{aligned} \tau &\rightarrow \frac{a\tau+b}{c\tau+d} \\ \begin{pmatrix} B \\ \tilde{B} \end{pmatrix} &\rightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} B \\ \tilde{B} \end{pmatrix} \end{aligned} \tag{14.15}$$

leaving the metric G (in the Einstein frame) and the 4-form A_4 fields invariant.

Clearly this continous symmetry cannot be a symmetry of the complete quantum theory, since it would rotate the charges continously, in contradiction with the fact that they must lie in a lattice by Dirac quantization condition. There is however plenty of evidence for the conjecture that a discrete $SL(2,\mathbf{Z})$ subgroup (defined by $a,b,c,d \in \mathbf{Z}$) is an exact symmetry of the complete string theory.

This remarkable proposal has the implication that there is a strong-weak duality between the theory at coupling g_s, $a=0$ and the theory at coupling $1/g_s$, $a=0$. Namely, the strong coupling regime of type IIB theory is equivalent to the perturbative weak coupling regime of a dual type IIB theory. Following the dependence of brane tensions as g_s changes it is possiblo to match the BPS states in both theories. For instance

IIB at g_s		IIB at $1/g_s$
F1	$\leftrightarrow$	D1
D1	$\leftrightarrow$	F1
NS5	$\leftrightarrow$	D5
D5	$\leftrightarrow$	NS5
D3	$\leftrightarrow$	D3

We see that starting at $g_s \simeq 0$, as g_s increases at goes to infinity the initial fundamental string becomes a D1-brane in the dual description, while the original D1 becomes light and turns into the fundamental, perturbative string in the dual description. The flow of BPS states is illustrated in figure 14.3.

This has the striking implication that the fundamental string is 'fundamental' only at weak coupling, while at finite coupling both the D1 and the

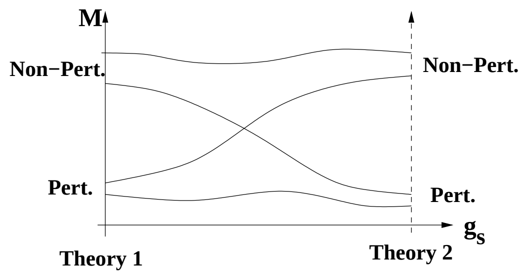


Figure 14.3: As a modulus (the dilaton vev) is changed, the original weakly coupled string theory becomes strongly interacting, and at infinite coupling it can be described as a weakly interacting *dual* theory. Perturbative and non-perturbative states are reshuffled in this interpolation.

F1 are both simply two BPS string-like objects, and at strong coupling the D1 is the one becoming the fundamental, perturbative object.

Indeed the situation is even more intriguing. The $SL(2,\mathbf{Z})$ symmetry predicts the existence of BPS strings with charges (p,q) under the two type IIB 2-forms, all forming an orbit of $SL(2,\mathbf{Z})$. These are easily constructed as supergravity solutions, by applying $SL(2,\mathbf{Z})$ transformations to the known F1 or D1 solutions (which correspond to $(p,q) = (1,0)$, $(0,1)$). At different points in the moduli space of the coupling τ, related to the perturbative limit by an $SL(2,\mathbf{Z})$ transformation, it is a different (p,q) string which becomes the perturbative object in the dual ($SL(2,\mathbf{Z})$ transformed) theory.

Since the symmetry relates theories which are equivalent, up to (very non-trivial) field redefinitions, the moduli space of physically distinct theories is $SL(2,\mathbf{R})/(\mathbf{U(1)}\times\mathbf{SL(2,Z)})$,

Duality relations in other 10d string theories will be studied in later lectures. We conclude this lecture by pointing out that the picture for type II theories is even more intricate as one lowers the dimension.

14.3.2 Toroidal compactification and U-duality

Let us consider compactification of type IIB theories on e.g. $\mathbf{T^6}$. The results for type IIA on $\mathbf{T^6}$ would be equivalent via T-duality, but the interpretation in terms of the original 10d theory is clearly different. It will be better understood in later lectures.

We are interested in studying non-perturbative states and duality properties of this theory (the case of other toroidal compactification is similar in many respects, see [86, 104]. We are interested in i) the moduli space of scalars ii) the 4d gauge fields, in particular 1-form gauge bosons iii) the BPS states preserving 1/2 of the supersymmetries iv) the duality group.

i) Let us determine the structure of the moduli space of scalars. In $\mathbf{T^6}$ compactifications of type IIB theory we have 36 scalars from the moduli G_{ij}, B_{ij}. These are known from the Narain lattice description to take values in the coset

$$\frac{SO(6,6)}{SO(6)\times SO(6)\times SO(6,6;\mathbf{Z})} \tag{14.16}$$

In addition, we have the scalars a, ϕ inherited from 10d, and which parametrize the coset

$$\frac{SL(2,\mathbf{R})}{U(1)\times SL(2,\mathbf{Z})} \tag{14.17}$$

In addition, we have 15 scalars $\tilde{B}_{ij}$, 15 scalars A^+_{ijkl} and two scalars, dual to the 4d 2-forms $B_{\mu\nu}$, $\tilde{B}_{\mu\nu}$. Overall we have 70 scalars, which in the supergravity approximation live in a coset locally of the form

$$E_7/SU(8) \tag{14.18}$$

where E_7 denotes the (non-compact) group generated by exponentiating the Lie algebra generated by generators of $SO(6,6)$ and $SL(2)$.

The supergravity effective action has a continuous symmetry E_7 acting non-trivially on the moduli space of scalars. As usual, classical supergravity is not sensitive to quantization conditions, and it will be only a subgroup of this which will be proposed to correspond to a full symmetry of the theory. This will come later on.

ii) The theory contains 56 4d 1-form fields. 24 of them are given by $B_{\mu i}$, $\tilde{B}_{\mu i}$ and their 4d duals; these transform in the representation $(12,2)$ of the classical global symmetry $SO(6,6)\times SL(2,\mathbf{R})$. The remaining 32 are given by 12 from $G_{\mu i}$ and their duals and 20 from [2] $A^+_{ijk\mu}$; these transform in the representation $(32,1)$ of $SO(6,6)\times SL(2,\mathbf{R})$. In total the 56 gauge bosons transform in the representation 56 of the classical symmetry E_7.

[2]Notice that A^+_4 has self-dual field strength in 10d.

iii) The elementary (in the sense that they carry at most one charge) BPS states carrying charged under gauge bosons are of different kinds

• We can have fundamental strings winding along any of the 6 directions in $\mathbf{T^6}$. We can also have D1-strings winding along any of these directions. These are charged under the fields $B_{\mu i}$, $\tilde{B}_{\mu i}$

• We can have 6 particle-like states in 4d from NS5-branes wrapped in all dimensions of $\mathbf{T^6}$ except one [3] Similarly we get 6 additional states from D5-branes wrapped in all dimensions of $\mathbf{T^6}$ except one. These are charged under the duals of $B_{\mu i}$, $\tilde{B}_{\mu i}$. The above 12 states plus these 12 transform in the $(12, 2)$ representation of the global symmetry $SO(6, 6) \times SL(2, \mathbf{R})$.

• KK momentum states. These are described by fundamental string states with momentum along some internal direction in $\mathbf{T^6}$. There are 6 basic states, charged under the 4d gauge fields $G_{\mu i}$.

• The corresponding states charged magnetically under $G_{\mu i}$ (i.e. charged electrically under their 4d duals) are Kaluza-Klein monopoles (also known as KK5-branes). The KK monopole configurations are discussed in appendix .2. These 6 states are labelled by $i = 1, \ldots, 6$ and have their isometrical direction along the i^{th} direction in $\mathbf{T^6}$ and volume spanning the remaining 5 directions in $\mathbf{T^6}$.

• Finally we have 20 additional states given by D3-branes wrapped on three internal directions in $\mathbf{T^6}$. The above 12 states plus these 20 transform in the representation $(32, 1)$ of $SO(6, 6) \times SL(2, \mathbf{R})$.

In total, these states transform in the representaion 56 of the classical symmetry group E_7

iv) These states must have quantized charges, so clearly the full continuous E_7 symmetry cannot be an exact symmetry of the complete theory. Rather, the proposal is that the discrete subroup of E_7 which leaves the 56-dimensional lattice of charges invariant is an exact symmetry of the quantum theory.

This is a simple generalization of thing we already know. In fact, the discrete duality group, denoted $E_7(\mathbf{Z})$, is the also the group of discrete transformations containing the T-duality group $SO(6, 6; \mathbf{Z})$ and the S-duality group

[3] To consider branes with some transverse compact circle, we can consider starting with an infinite transverse dimension, on which we place and infinite periodic array of branes (this is possible and static due to the BPS no-force condition), and then modding by discrete translations to obtain a circle.

$SL(2, \mathbf{Z})$. The global structure of the moduli space is

$$\frac{E_7}{SU(8) \times E_7(\mathbf{Z})} \tag{14.19}$$

All BPS states in the theory transform in representaions of the duality group $E_7(\mathbf{Z})$ (known as U-duality group).

This has remarkable implications. In particular there are infinite sets of points in moduli space which are equivalent to weakly coupled large volume compactifications of IIB on $\mathbf{T^6}$ once written in suitable dual terms. The perturbative parameter in these dual theories can be a complicated combination of the 70 scalars in the coset $E_7/SU(8)$, and not just a function of the dilaton. Moreover the string-like object which is becoming the fundamental string in this dual theory can be a complicated object, not just the F1 or the D1-string. In fact string-like objects also form a complicated representation (I think the 133) of $E_7(\mathbf{Z})$: we have the unwrapped F1, and D1, A D3-brane wrapped in two directions, D5-branes wrapped in four directions, etc. Any of these can become the fundamental string in one particualr corner of moduli space.

For the interested reader, let us simply point out that similar duality relations hold in toroidal compactifications of heterotic string theory. In fact, $\mathbf{T^6}$ comapctfifications lead to $\mathcal{N} = 4$ 4d theories, whose gauge sector is a generalization of the kind of theories in appendix .1, and have an $SL(2, \mathbf{Z})$ duality which corresponds to Montonen-Olive in the associated gauge field theory. We will rederive Montonen-Olive duality in later lectures, using D-branes to study gauge field theories.

14.4 Final comments

We have seen that string theory contains plenty of non-perturbative states. These are very important for the theory at finite coupling, and are in a sense on an equal footing with perturbative or fundamental objects in this regime (p-brane democracy). In fact, they can become the fundamental degrees of freedom in different corners in moduli space, and can be described as the fundamental strings in a suitable dual description.

We still do not have a microscopic description of string theory which is valid beyond perturbation theory, and which includes all these BPS states on an equal footing. What is clear anyway is that as soon as we go beyond the perturbative regime, string theory is no longer a theory of strings! and must also include other extended objects.

.1 Some similar question in the simpler context of field theory

A more detailed reference for this section is [73].

.1.1 States in field theory

We consider a well studied and simple 4d field theory, which is $\mathcal{N}=4$ supersymmetric $SU(2)$ gauge theory. The vector multiplets contain one gauge boson, four Majorana fermions and six real scalars in the adjoint. The scalar potential has the form $V(\phi)=|[phi^i,\phi^j]|$, so a generic vacuum is labelled by diagonal vevs of the form

$$\phi^i=\begin{pmatrix} v_i & 0 \\ 0 & -v_i \end{pmatrix} \tag{20}$$

We denote $v=\sum_i v_i^2$. A generic vev v breaks spontaneously the gauge symmetry $SU(2)\to U(1)$.

At low energies in one of these vacua, $E \ll g_{YM}v$ the effective theory is $\mathcal{N}=4$ susy $U(1)$ gauge theory, with action

$$S=\int_{4d}\frac{1}{g_{YM}^2}F\wedge *F+\theta\int_{4d}F\wedge F \tag{21}$$

The theory clearly contains states electrically charged under the gauge potential A; they are the massive gauge bosons. The mass of one such state with charge $n_e\in\mathbf{Z}$ is

$$M=|n_e|g_{YM}v \tag{22}$$

We can also look for non-perturbative states of the theory by constructing solutions to the equations of motion (see [72] for an introduction to solitons). Indeed the theory contains particle-like states known as 't Hooft-Polyakov monopoles, as we discussed in the introductory lectures. Such monopoles are

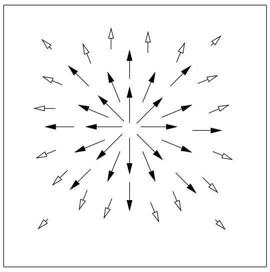


Figure 4: Picture of the hedgehog configuration for the Higgs field.

described by field configurations asymptoting as

$$\phi^i(\vec{x},t) \;\to\; \frac{v^i}{r}\,x^i \;+\mathcal{O}(1/r^2)$$
$$A^i(\vec{x},t) \;\to\; \frac{1}{r^2}\,x^i \;+\mathcal{O}(1/r^2) \qquad (23)$$

This is the so-called hedgehog configuration, shown in figure B.4. From the point of view of the low energy $U(1)$ theory, the field configurations are Wu-Yang monopoles of the kind studied in the differential geometry lecture.

These objects carry magnetic charge $n_m \in \mathbf{Z}$ under the gauge potential A, and their mass is

$$M = |n_m| v/g_{YM} \qquad (24)$$

(if the θ parameter is non-zero, they also carry an electric charge proportional to $q_e\theta n_m$). The mass of a general state with electric and magnetic charges (q_e, q_m) is given by

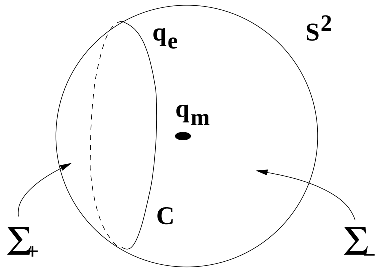
$$M^2 = v^2 \frac{1}{\Im\tau}|q_e + \tau q_m|^2 \qquad (25)$$

where $\tau = \theta + i/g_{YM}^2$. For $\theta = 0$ this gives

$$M = |v||g_{YM} q_e + \frac{1}{g_{YM}} q_m| \qquad (26)$$

Dirac charge quantization condition

This is a consistency condition on the possible set of charges in a theory with electric and magnetic charges. A particle with electric charge q_e moving



in a circle worldline C acquires a phase $\exp(iq_e \int_C A)$ in its path integral. In the presence of a particle carrying magnetic charge q'_m, the gauge potential is not globally well defined, so the above expression could be ambiguous, leading to an ill-defined wavefunction for the electric particle. Indeed, as shown in figure .1.1, the integral $\int_C A$ can be computed via Stokes theorem as $\int_\Sigma F$ over some surface Σ with $\partial\Sigma = C$. The result however can depend on the surface Σ chosen. For the two surfaces in the picture, the difference in the exponent of the phase is

$$\Delta q_e \int_C A = q_e(\int_{\Sigma_+} F - \int_{\Sigma_-} F) = q_e \int_{\mathbf{S^2}} F = q_e q'_m \tag{27}$$

where $\mathbf{S^2}$ is a surface that encloses the magnetically charged particle. In order to have a well-defined phase, we then need

$$q_e q'_m \in 2\pi\mathbf{Z} \tag{28}$$

This is Dirac quantization conditions, which constrains the charges in a theory with electric and magnetic objects.

If the theory contains dyonic particles, carrying electric and magnetic charges at the same time, consistency of the phase picked up by moving a particle of charges (q_e, q_m) in the presence of a particle of charge (q'_e, q'_m) requires

$$q_e q'_m - q_m q'_e \in 2\pi\mathbf{Z} \tag{29}$$

This implies that charges (q_e, q_m) must lie in a 2d discrete lattice. One can check that the charges of the above theory, which are of the form (q_e, q_m) with $q_e + iq_m = n_e + \tau n_m$, with $n_e, n_m \in \mathbf{Z}$, satisfy this constraint (zzz Warning: I was not careful about 2π's).

.1.2 BPS bounds

The general supersymmetry algebra for $\mathcal{N} = 4$ has the structure

$$\{Q^A_\alpha, Q^{B\,\dagger}_\beta\} = -2\delta^{AB} P_\mu \Gamma^\mu_{\alpha\beta} - 2iZ^{AB}\delta_{\alpha\beta} \tag{30}$$

where Q^A_α, $A = 1, \ldots, \mathcal{N}$ are the $\mathcal{N}$ supercharges ($\mathcal{N} = 4$ in our case) with a (Majorana) spinor index α. The Z^{AB} are operators that commute with the Q's, the P's and hence with the Hamiltonian. Thus they are conserved charges of the system, known as central charges, which are combination of the conserved gauge charges of the theory.

In a given state, Z^{AB} forms a real antisymmetric matrix, which can be brought to a block diagonal form with blocks

$$\begin{pmatrix} 0 & q_i \\ -q_i & 0 \end{pmatrix} \tag{31}$$

The supersymmetry algebra implies a bound on the mass of particle states in the sector of fixed (central) charges q_i. This is done as follows: take for simplicity a sector of equal charges $q_i = q$, we can go to the rest frame of the particle, where $(P_\mu) = (M, 0, 0, 0)$. Then the matrix $\{Q^A_\alpha, Q^{B\,\dagger}_\beta\}$, which is positive definite, is diagonal in blocks of the form

$$\begin{pmatrix} 2M & 2iq \\ -2iq & 2M \end{pmatrix} \tag{32}$$

This implies that the eigenvalues, which are $2(M \pm q)$ must be positive, so that we get a bound

$$M \geq |q| \tag{33}$$

This is known as BPS bound. States saturating this kind of bounds are called BPS states. They are special because they correspond to zero modes of the supercharge anticommutator matrix, and this implies that they are annihilated by some supercharges. This is equivalent to saying that BPS states are invariant under some supersymmetry transformations (generated by the corresponding supercharges). On the other hand, this implies that the supermultiplets to which these states belong are shorter than the generic supermultiplet.

This implies that upon continuous deformations of the theory (for instance including quantum corrections or threshold effects of the underlying high energy theory) BPS states cannot cease being BPS, since the number of fields in the supermultiplet cannot jump discontinuously. This also implies that, since the mass of the state is fixed by the supersymmetry algebra, it is exactly known, and does not suffer any correction form quantum loops or otherwise. Therefore, the classical result for the mass of a BPS state can be exactly extrapolated to strong coupling and other difficult regimes.

In our case above, it is possible to show that in a sector of electric and magnetic charges (q_e, q_m) the central charge for the superalgebra is of the above form

$$q_i = q = vg_{YM}(q_e + \tau q_m) \tag{34}$$

This allows to claim that the above discussed states are BPS and the masses (D.8) is exact.

.1.3 Montonen-Olive duality

The equations of motion for the $U(1)$ gauge theory are (for $\theta = 0$)

$$\begin{aligned} dF &= j_m \\ d * F &= j_e \end{aligned} \tag{35}$$

where j_e, j_m are the electric and magnetic charge currents. They have a global $SL(2,\mathbf{R})$ rotation invariance

$$\begin{pmatrix} *F \\ F \end{pmatrix} \rightarrow M \begin{pmatrix} *F \\ F \end{pmatrix} \quad ; \quad \begin{pmatrix} j_e \\ j_m \end{pmatrix} \rightarrow M \begin{pmatrix} j_e \\ j_m \end{pmatrix} \quad ; \quad M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad , \quad ad - bc = 1 \tag{36}$$

This also acts by rotating the charges (q_e, q_m), so it is able to exchange the roles of elementary electrically charged states and solitonic magnetic monopoles, i.e. of perturbative and non-perturbative states in the system. Indeed, for the theory (e.g. the energies of the states) to be invariant, $SL(2,\mathbf{R})$ must also act on the coupling constant τ by

$$\tau \rightarrow \frac{a\tau + b}{c\tau + d} \tag{37}$$

Since the charges must live in a discrete lattice due to the Dirac quantization condition, it is clear that the classical $SL(2,\mathbf{R})$ symmetry cannot be a symmetry of the full quantum theory. However, the subgroup $SL(2,\mathbf{Z})$ given by matrices M with $a, b, c, d \in \mathbf{Z}$ leaves the charge lattice invariant as a whole, and also is a symmetry of the mass formula (D.8). The Montonen-Olive duality proposal is that this $SL(2, \mathbf{Z})$ is an exact symmetry of the full quantum theory.

This symmetry has very non-trivial implications:

• It implies that BPS solitons must appear in orbits of $SL(2, \mathbf{Z})$. In particular this implies the existence of BPS dyonic states with charges $q_e + iq_m = n_e + \tau n_m$ for coprime n_e, n_m; this is the orbit containint the elementary electically charged states $(n_e, n_m) = (\pm 1, 0)$ and the basic magnetic monopoles $(n_e, n_m) = (0, \pm 1)$. Some of these dyonic states have been explicitly constructed [115].

• It implies that the theory at coupling g_{YM}, $\theta = 0$ has a completely equivalent description in terms of a theory with coupling $g'_{YM} = 1/g_{YM}$, $\theta' = 0$. One says that it is a strong-weak coupling duality. This implies that the strong coupling of the first theory is described by a weakly coupled theory in the dual side. The theory simplifies enormously in the limit of very strong coupling, which in principle looked like a very difficult regime!. The theory becomes simply perturbative Maxwell theory in terms of the dual elementary fields, which are the solitons of the initial theory.

• In fact, there is an infinite number of limits where the dynamics reduces to perturbative Maxwell theory in terms of a dual theory, which is related via an $SL(2, \mathbf{Z})$ transformation to the original one.

• These properties are a good toy model for the dualities in string theory. This has been our motivation for discussing this field theory example. In fact, we will see in later lectures that duality in string theory implies duality in field theory.

.2 The Kaluza-Klein monopole

Consider a D-dimensional theory with gravity, compactified on a circle, so that it corresponds to a vacuum of the form $M_{D-1} \times \mathbf{S^1}$. The Kaluza-Klein monopole is a purely metric configuration, which corresponds to an excited state of this theory, and exists if $D \geq 4$. It is described by a geometry $M_{D-4} \times X_{TN}$, where the so-called (multi)Taub-NUT space X_{TN} has the following metric

$$ds^2 = V(\vec{x})^{-1} d\vec{x}^{\,2} + V(\vec{x})(d\tau + \vec{\omega} \cdot d\vec{x})^2 \qquad (38)$$

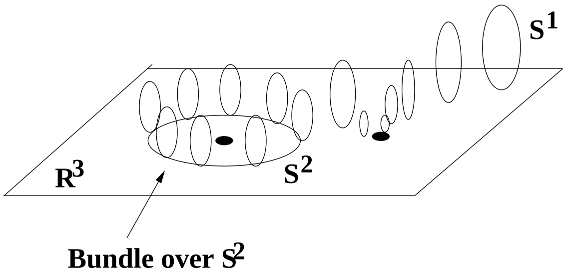


Figure 5: A picture of the multi-Taub-NUT space X_{TN}. It is a circle fibration over $\mathbf{R^3}$, with fiber asymptoting to constant radius at infinity, and degenerating to zero radius over the centers, shown as black dots. Around an $\mathbf{S^2}$ surrounding a center, the $\mathbf{S^1}$ fibrations defines a non-trivial $U(1)$ bundle with first Chern class 1.

with

$$\vec{\nabla} \times \vec{\omega} = V(\vec{x}) \quad ; \quad V(\vec{x}) = 1 + \sum \frac{1}{|\vec{x} - \vec{x}_a|} \tag{39}$$

The space X_{TN} is a fibration of $\mathbf{S^1}$ (parametrized by τ) over $\mathbf{R^3}$ (parametrized by $\vec{x}$), with the properties that (see figure 5)

i) the $\mathbf{S^1}$ in the fiber asymptotes to constant radius at infinity on the base $\mathbf{R^3}$. So it is a finite energy excitation of the vacuum $M_{D-1} \times \mathbf{S^1}$.

ii) the $\mathbf{S^1}$ denerates to zero radius at the location of the so-called centers $\vec{x} = \vec{x}_a$.

iii) The $\mathbf{S^1}$ fibered over an $\mathbf{S^2}$ in the base $\mathbf{R^3}$ surrounding a center, is a non-trivial $\mathbf{S^1}$ (or $U(1)$)bundle over $\mathbf{S^2}$ with first Chern class equal to 1. If the $\mathbf{S^2}$ surrounds k centers, the Chern class of the bundle of $\mathbf{S^1}$ over $\mathbf{S^2}$ is k. In fact, one can show that the mixed component of the Christoffel connection is exactly the gauge field of the Wu-Yang monopole studied in the lecture on differential geometry.

iv) This implies that the geometry carries a topological magnetic charge under the $D-1$ dimensional gauge boson $G_{\mu(\tau)}$. The sources of the charge are localized at the centers of the metric, which then behave as magnetic monopoles for this field. The configuration defined by Taub-NUT space is known as Kaluza-Klein monopole.

The above metric has $SU(2)$ holonomy (so can be though of as a non-compact Calabi-Yau in two complex coordinates) so it is invariant under half of the supersymmetries. It is a 1/2 BPS state. Its ADM tension is proportional to R^2/g_s^2, where R is the radius of the isometrical direction $\mathbf{S^1}$ parametrized by τ.

In circle compactifications of string theory, the resulting 9d object is Poincare invariant in six dimensions, and is localized in three dimensions. It is often called the Kaluza-Klein fivebrane. In toroidal compactficiations of several dimensions, one can have different BPS states given by the different choices of the circle in $\mathbf{T^d}$ chosen to correspond to the isometrical direction in X_{TN}.

Appendix A

D-branes

A.1 Introduction

In the previous lecture we used supergravity to obtain partial information on non-perturbative states in string theory. We could rely on the existence and certain properties (tension, charge) of some of these p-brane states, when they satisfy some BPS condition.

In this lecture we propose a *microscopic* description, valid at weak coupling, for some of these solitoon (those we called Dp-branes), explicitly in terms of the underlying string theory. This description allows to recover the results we found in the supergravity approximation, and to describe several others (exactly in α'). Indeed, the study of D-branes from several viewpoints is one of the most active topics in string theory nowadays.

Let us emphasize that the microscopic description we are going to propose cannot be derived from out macroscopic description from the supergravity viewpoint. Rather, the microscopic description will show that the object we describe microscopically is a source of the supergravity fields with the same properties of the objects in the previous lecture.

A.2 General properties of D-branes

From the supergravity viewpoint, we introduced some solitonic solutions, the Dp-branes. They exist for p even in type IIA theory, for p odd in type IIB theory and for $p = 1, 5$ in type I. They are described by a gravitational background; fluctuations of the theory around the soliton solution are localized

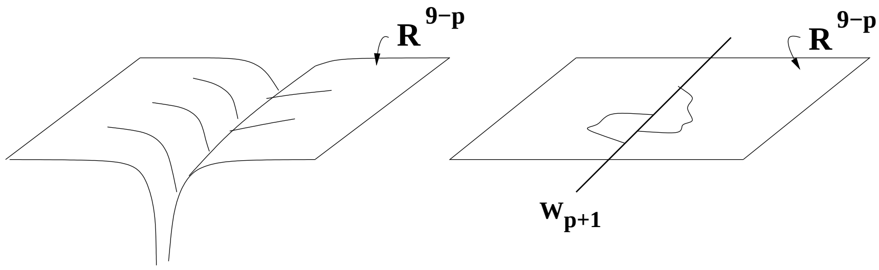


Figure A.1: Fluctuations of the theory around a Dp-brane sugra solution can be described in stringy language as open strings with ends on a $(p+1)$-dimensional surface, located at the core of the topological defect.

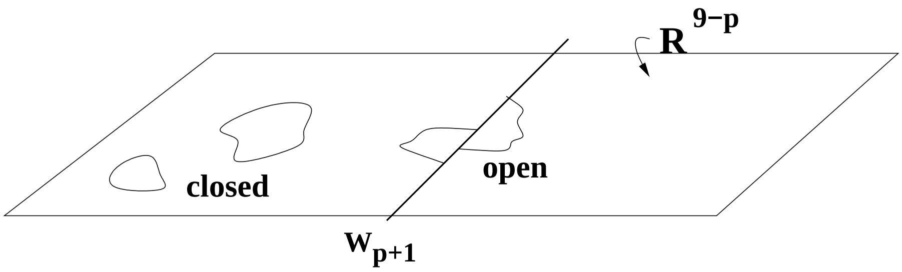


Figure A.2: String theory in the presence of a Dp-brane. The closed string sector describes the fluctuations of the theory around the vacuum (gravitons, dilaton modes, etc), while the sector of open strings describes the spectrum of fluctuations of the soliton.

on the $(p+1)$-dimensional volume of the soliton core.

The stringy description of Dp-branes, at weak coupling, is as follows. They are described as $(p+1)$-dimensional planes W_{p+1} in flat space, with the prescription that the theory in its presence contains open strings, with endpoints on the $(p+1)$-dimensional plane W_{p+1}. See figure A.1.

Equivalently, the fluctuations of the string theory around the topological defect are microscopically described as open strings ending on its $(p+1)$ volume.

A complementary point of view, relating the microscopic description with the supergravity solution, is that interactions of the $(p+1)$-dimensional plane with the closed string modes (via the open string modes on the brane) imply the plane is a source of the graviton, dilaton and RR fields, which creates a

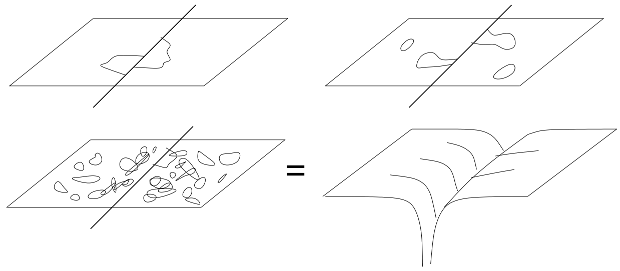


Figure A.3: A Dp-brane interacts with closed strings via open strings, creating an effective background which describes the backreaction of the D-brane tension and charge on the configuration.

background as that described by the supergravity solution. See figure A.3. Recall that the size of the throat in the supergravity solution is $g_s N \alpha'^{1/2}$, so this effect is bigger when g_s increases (and then the supergravity description is reliable, while for small g_s the stringy description is more precise).

That is, the object we have described as a $(p+1)$ plane on which open strings are allowed to end, has the correct properties to lead to a Dp-brane supergravity solution. The coupling to the closed string modes can be obtained from the disk diagram with a closed string insertion, see figure A.4. In particular, it allows to obtain the tension and the charge under the RR $(p+1)$-form; they are of the order of $1/g_s$, since the Euler characteristic of the disk is $\xi = 1$. It is also possible to verify they satisfy the BPS condition; indeed, we will find below that they are supersymmetric states.

One could raise a number of objections against coupling this kind of open string sectors to a sector of closed strings.

i) The open string sector is not Poincare invariant. This is not a problem, since it is describing the fluctuations of the theory around a soliton state which breaks part of the Poincare invariance of the vacuum.

ii) The 2d worldsheet bosons associated to directions transverse to the D-brane, $X^i(\sigma,t)$ (and also the 2d fermions) obey Dirichlet boundary conditions

$$\partial_t X^i(\sigma,t)|_{\sigma=0,\ell} = 0 \tag{A.1}$$

Are these boundary conditions consistent? Do we recover the same local 2d dynamics as for closed strings? In fact, we can check that Dirichlet boundary

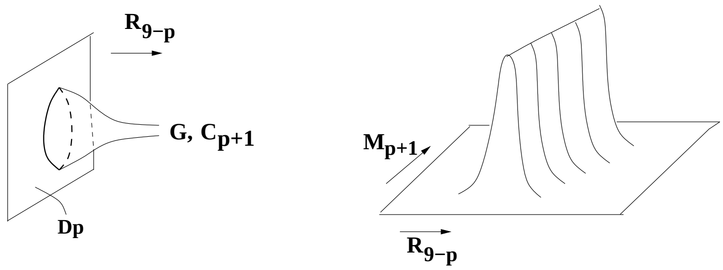


Figure A.4: D-branes interact with closed string modes, and in particular couple to the bulk graviton and $(p+1)$-form fields, i.e. they have tension (of order $1/g_s$ in string units) and carry charge. Their backreaction on the background curves and deforms it into the p-brane solution seen in the supergravity regime.

conditions do the job . Recall that the variation of the Polyakov action is

$$\begin{aligned} \delta S_P &= -\frac{1}{2\pi\alpha'}\int_\Sigma d^2\xi\, g^{ab}\partial_a X^\mu \partial_b \delta X^\mu = \\ &= -\frac{1}{2\pi\alpha'}\int_{-\infty}^{\infty} de\,(g^{ab}\delta X^\mu \partial_b X_\mu)|_{\sigma=0}^{\sigma=\ell} + \frac{1}{2\pi\alpha'}\int_\Sigma d^2\xi\,\delta X_\mu\, g^{ab}\partial_a\partial_b X^\mu \end{aligned} \quad \text{(A.2)}$$

For Dirichlet boundary conditions, the corresponding endpoint is not allowed to move, so the allowed variations must satisfy $\delta X^i = 0$. Hence the boundary term for any coordinate drops, for Neumann or Dirichlet boundary conditions.

iii) In th lecture on open strings we saw that open strings allowed to end anywhere on spacetime cannot be consistently added to type IIB theory, due to RR tadpole cancellation conditions. In fact, this kind of configurations can be understood as type IIB theory in the presence of D9-branes, which are charged under C_{10} and lead to an inconsistency in the equations of motion. For configurations with lower-dimensional Dp-branes, $p < 9$, the corresponding RR form C_{p+1} does have a kinetic term and the equation of motion can be solved. RR charges are not dangerous if there are non-compact dimensions transverse to the D-brane. Intuitively, the fluxlines created by the D-brane charge can escape to infinity along the non-compact dimension. If there are no transverse directions, or they are compact, the flux cannot escape and one should require charge cancellation as a consistency condition.

A.3 World-volume spectra for type II D-branes

The fluctuations of the theory around the soliton background are described by open strings ending on the D-brane $(p+1)$-dimensional world-volume. These modes describe the dynamics of the Dp-brane. For instance, zero mass oscillation modes of the open strings correspond to zero energy motions of the Dp-brane.

In this section we compute the spectrum of open strings ending on the Dp-brane. They give rise to fields propagating on the volume of the Dp-brane, and describe its dynamics. For concreteness we center on type IIB D-branes, which have even world-volume dimension.

A.3.1 A single Dp-brane

Consider a configuration given by a single Dp-brane with worldvolume spanning the directions X^μ, $\mu = 0, \ldots, p$ and transverse to the directions X^i, $i = p+1, \ldots, 9$. Consider an open string with both endpoints on the Dp-brane. Its worldsheet 2d theory is described by 2d bosons $X^\mu(\sigma, t)$, $\mu = 2, \ldots, p$ (in the light cone gauge) and $X^i(\sigma, t)$, $i = p+1, \ldots, 9$, and their 2d fermion partners. see fig A.5.For directions along the brane volume, we have Neumann boundary conditions, while for directions transverse to it we have Dirichlet boundary conditions

$$\partial_\sigma X^\mu(\sigma, t)|_{\sigma=0,\ell} = 0 \quad ; \quad \partial_t X^i(\sigma, t)|_{\sigma=0,\ell} = 0 \quad ; \tag{A.3}$$

Using the mode expansions, of the form

$$X(\sigma, t) \;\; = \;\; x + w\sigma + \frac{p}{p^+}t + i\sqrt{\frac{\alpha'}{2}}\sum_\nu \frac{\alpha^i_\nu}{\nu}\, e^{-\pi i\,\nu\,(\sigma+t)/\ell} + i\sqrt{\frac{\alpha'}{2}}\sum_{\tilde\nu} \frac{\tilde\alpha^i_{\tilde\nu}}{\tilde\nu}\, e^{-\pi i\,\tilde\nu\,(\sigma+t)/\ell}$$

For X^μ we obtain

$$x^\mu, p^\mu \text{ allowed} \quad ; \quad w^\mu = 0 \quad ; \quad \nu = n \in \mathbf{Z} \quad ; \quad \alpha^\mu_{\mathbf{n}} = \tilde\alpha^\mu_{\mathbf{n}} \tag{A.4}$$

For X^i we obtain

$$x^i \text{ allowed} \quad ; \quad p^i = 0 \quad ; \quad w^i = 0 \quad ; \quad \nu = n \in \mathbf{Z} \quad ; \quad \alpha^{\mathbf{i}}_{\mathbf{n}} = -\tilde\alpha^{\mathbf{i}}_{\mathbf{n}} \tag{A.5}$$

For the NN directions we have the expansion familiar from the lesson on open strings. For the DD directions we have the expansion

$$X^i(\sigma, t) = x^i + \sqrt{2\alpha'}\sum_{n\neq 0}\frac{\alpha^i_n}{n}\,\sin(\pi n\sigma/\ell)\, e^{-\pi i n t/\ell} \tag{A.6}$$

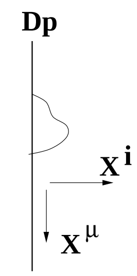


Figure A.5: .

and similarly for fermions.

In total, we obtain integer modded bosonic oscillators α_n^μ, α_n^i, and fermionic oscillators $\psi_{n+\rho}^\mu$, $\psi_{n+\rho}^i$, with $\rho = 1/2, 0$ for NS or R fermions. Note also that these states have momentum only in directions along the volume of the D-brane, and not in those transverse to it. This implies that the correspondinng particles propagate only in the $(p+1)$-dimensional D-brane world-volume.

The spectrum is very similar to that of an open superstring sector, with the same states reinterpreted with respect to a lower-dimensional Lorentz group. In particular, at the massless level the states are

Sector	State	$SO(8)$ weight	$SO(p-1)$	$(p+1)$-dim field
NS	$\psi_{-1/2}^\mu \lvert 0\rangle$	$(0,\ldots,0,\underline{\pm,\ldots,0})$	Vector	Gauge boson A_μ
	$\psi_{-1/2}^i \lvert 0\rangle$	$(\underline{\pm,\ldots,0},0,\ldots,0)$	Scalar	$9-p$ real scalars ϕ^i
R	$A_a^+ \lvert 0\rangle$	$\frac{1}{2}(\pm,\pm,\pm,\pm)$	spinor	$2^{(9-p)/2}$ ch. fermion λ_α
	$A_{a_1}^+ A_{a_2}^+ A_{a_3}^+ \lvert 0\rangle$	# - = odd		

This corresponds to a $U(1)$ vector supermultiplet with respect to 16 supersymmetries in $(p+1)$dimension. This is also often described as the dimensional reduction of the $\mathcal{N}=1$ 10d vector multiplet. A prototypical example is provided by the spectrum on a D3-brane, which corresponds to a $U(1)$ vector multiplet of $\mathcal{N}=4$ susy in 4d, given by one gauge boson, six real scalars and four Majorana fermions.

In fact, supersymmetry extends to the complete open string spectrum, implying the property that the D-brane is a 1/2 BPS state. Indeed, it is possible to verify that the boundary conditions imposed for a D-brane on the open string sector relate the spacetime supercharges arising from the

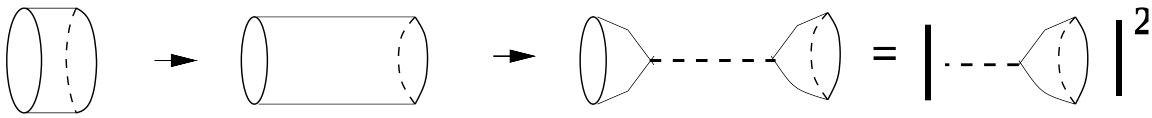


Figure A.6: The D-brane charge and tension arise from a disk diagram with insertions, which can be obtained from factorization of the annulus diagram.

left and right-movers in 2d, so that the configuration is still invariant under 10d supersymmetry transformations with parameters ϵ_L, ϵ_R (which are 10d spinors) satisfying

$$\epsilon_L = \Gamma^0 \ldots \Gamma^p \epsilon_R \tag{A.7}$$

Using the above microscopic description, and knowing how to quantize open string sectors, it is possible to compute explicitly the tension and charge of a Dp-brane. The standard technique is to evaluate the annulus amplitude, namely the one-loop vacuum amplitude for open strings with both ends on a D-brane, and go to the factorization limit where the amplitude splits into the square of the disk. The disk provides the the coupling between the D-brane and the NSNS fields, like the graviton, and the RR fields (i.e. the D-brane tension and RR charge). The computation is pictorially sketched in figure A.6 and gives the result (see section 13.3 in [71]

$$T_p^2 = \frac{\pi}{\kappa^2}(4\pi^2\alpha')^{3-p} \quad ; \quad Q_p = T_p/g_s \tag{A.8}$$

A.3.2 Effective action

The $(9-p)$ real scalars in the volume of the Dp-brane are the goldstone bosons associated to translational symmetries of the vacuum, broken by the presence of the soliton [1]. This implies that the vevs of these scalars provide the location of the Dp-brane in transverse space $\mathbf{R^{9-p}}$. It also implies that non-trivial profiles for these scalar fields (that is, configurations with x^μ-dependent backgrounds for these scalars) correspond to fluctuations of the embedding of the D-brane worldvolume on spacetime, see fig A.7. Namely

[1] Similarly, the fermions can be regarded as the goldstone fermions associated to supersymmetries of the vacuum, broken by the presence of the D-brane.

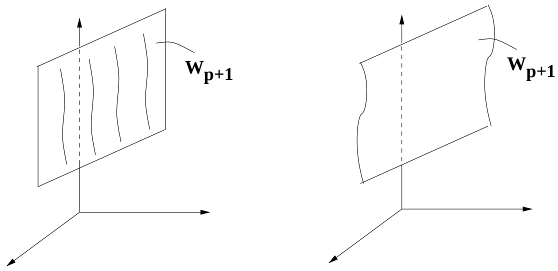


Figure A.7: A nontrivial configuration for one of the worldvolume translational zero modes corresponds to a non-trivial embedding of the soliton worldvolume in spacetime.

$\phi^i(x^\mu)$ describes the embedding of the D-brane volume in spacetime. Therefore the effective action for the massless open string modes on the D-brane worldvolume corresponds to an effective action for the D-brane, controlling its dynamics.

There are two strategies to obtain this effective action, which are conceptually analogous to the computation of effective actions for closed string sectors. The first is to compute scattering amplitudes in string theory and to cook up an action that reproduces them. The second is to couple a general background of the massless fields to the 2d worldsheet theory, and to demand conformal invariance (both locally on the 2d worldsheet and on the boundary conditions for general backgrounds); the conformla invariance constraints can then be interpreted as equations of motion for the spacetime fields, arising from some effective action, see [75].

The resulting effective action has several pieces. One of them is the Dirac-Born-Infeld action, which has the form

$$S_{Dp} = -T_p \int_{W_{p+1}} d^{p+1}x^\mu \, \left(-\det(G + B + 2\pi\alpha' F)\right)^{1/2} \tag{A.9}$$

where $G_{\mu\nu} = \partial_\mu \phi^i \partial_\nu \phi^j \, G_{ij}$ is the metric induced on the D-brane worldvolume [2], and similarly $B_{\mu\nu}$ is the induces 2-form. These terms introduce the dependence of the action on the embedding fields $\phi^i(x^\mu)$. Finally $F_{\mu\nu}$ is the field strength of the worldvolume gauge field.

[2] We have implicitly fixed the worldvolume reparametrization invariance to fix a 'static gauge'. The scalars associated to these gauge degrees of freedom do not appear in the light-cone spectrum.

The Dirac-Born-Infeld action carries the information about the coupling of the D-brane to the NSNS field. The Dirac-Born-Infeld action is α' exact in terms not involving derivatives of the field strength. Neglecting the dependence on the field strength, it reduces to the D-brane tension times the D-brane volume $\int(\det G)^{1/2})$. At low energies, i.e. neglecting the α' corrections, it reduces to a kinetic term for the scalars plus the $(p+1)$-dimensional Maxwell action for the worldvolume $U(1)$, with gauge coupling given by $g_{U(1)}^2 = g_s$. Of course the above action should include superpartner fermions, etc, but we skip their discussion.

A second piece of the effective action is the Wess-Zumino terms, of the form

$$S_{WZ} = -Q_p \int_{W_{p+1}} \mathcal{C} \wedge \mathrm{ch}(F)\, \hat{A}(R) \tag{A.10}$$

where $\mathcal{C} = C_{p+1} + C_{p-1} + C_{p-3} + \ldots$ is a formal sum of the RR forms of the theory, and $\mathrm{ch}(F)$ is the Chern character of the worldvolume gauge bundle on the D-brane volume

$$\mathrm{ch}(F) = \exp(\frac{F}{2\pi}) = 1 + \frac{1}{2\pi}\mathrm{tr}\, F + \frac{1}{8\pi^2}\mathrm{tr}\, F^2 + \ldots \tag{A.11}$$

and $\hat{A}(R)$ is the A-roof genus, characterizing the tangent bundle of the D-brane world-volume $\hat{A}(R) = 1 - \mathrm{tr}\, R^2/(2\pi^2)$. Integration is implictly defined to pick up the degree $(p+1)$ pieces in the formal expansion in wedge products. Hence we get terms like

$$\begin{aligned} S_{WZ} = \ & \textstyle\int_{W_{p+1}} -Q_p \Big(\int_{W_{p+1}} C_{p+1} + \frac{1}{2\pi} \int_{W_{p+1}} C_{p-1} \wedge \mathrm{tr}\, F + \\ & \textstyle + \frac{1}{8\pi^2} \int_{W_{p+1}} C_{p-3} \wedge (\mathrm{tr}\, F^2 - \mathrm{tr}\, R^2) + \ldots \Big) \end{aligned} \tag{A.12}$$

A very important property of this term is that it is topological, independent of the metric or on the particular field representatives in a given topological sector. This is related to the fact that these terms carry the information about the RR charges of the D-brane configuration.

A.3.3 Stack of coincident Dp-branes

As a consequence of the BPS property, the interaction between several parallel Dp-branes exactly vanishes. This can be understood from a cancellation

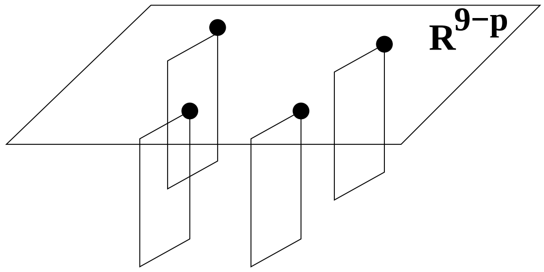


Figure A.8: .

of the attractive interaction due to exchange of NSNS fields (like the graviton) and the repulsive interaction due to exchange of RR fields. It can also be understood from the fact that the supersymmetry transformations unbroken by a D-brane depend only on the directions it spans, so several parallel D-branes preserve the same supersymmetries (A.7).

We would like to consider the spectrum of open strings in a configuration of n parallel Dp-branes, labelled $a = 1, \ldots, n$, spanning the directions x^μ, $\mu = 0, \ldots, p$, and sitting at the locations $x^i = x_a^i$ in the $(9-p)$ transverse directions. See figure A.8.

There are in this situation n^2 open string sectors, labelled ab, corresponding to open strings starting at the a^{th} D-brane and ending at the b^{th} D-brane. It is important to recall that we are working with oriented open strings (whose closed string sector is type II theory, which is oriented). For each of these n^2 sectors, the boundary conditions are NN for the 2d bosons $X^\mu(\sigma, t)$ (and fermions partners) and DD for the $X^i(\sigma, t)$ (and fermion partners). Namely, for an ab string we have

$$\begin{aligned} \partial_\sigma X^\mu(\sigma, t)|_{\sigma=0,\ell} &= 0 \\ X^i(\sigma, t)|_{\sigma=0} = x_a^i \quad ; \quad X^i(\sigma, t)|_{\sigma=\ell} &= x_b^i \end{aligned} \tag{A.13}$$

The mode expansion reads

$$X^i(\sigma, t) = x_a^i + \frac{x_b^i - x_a^i}{\ell}\sigma + i\sqrt{\frac{\alpha'}{2}} \sum_\nu \frac{\alpha_n^i}{n}\, e^{-\pi i\, n\,(\sigma+t)/\ell} + i\sqrt{\frac{\alpha'}{2}} \sum_{\tilde{n}} \frac{\tilde{\alpha}_{\tilde{n}}^i}{\tilde{n}}\, e^{-\pi i\, \tilde{n}\,(\sigma+t)/\ell}$$

The moddings etc works as in the case of just one Dp-brane. The spacetime mass formula is similar to the usual one for open strings, with and additional contribution arising from the winding term; we have

$$M^2 = \left(\sum_{i=p+1}^{9} \frac{x_b^i - x_a^i}{2\pi\alpha'} \right)^2 + \frac{1}{\alpha'}(N_B + N_F + E_0) \qquad \text{(A.14)}$$

with $E_0 = -1/2, 0$ in the NS, R sectors.

This leads to the same kind of massless states as above, for each of the n^2 ab sectors. Namely, we obtain a total of n^2 gauge bosons, $(9-p)$ times n^2 real scalars and $2^{(9-p)/2}$ times n^2 chiral fermions in $p+1$ dimensions. It is not difficult to realize that the aa strings lead to massless states, no matter what the x_a^i are, and produce a gauge group $U(1)^n$, each $U(1)$ propagating on the volume of each D-brane. On the other hand the ab states are generically massive, with mass squared proportional to $\sum_i (x_a^i - x_b^i)^2$, and have charges $(+1,-1)$ under $U(1)_a \times U(1)_b$.

When some, say k, of the location of the D-branes in transverse space $\mathbf{R^{9-p}}$ coincide, the corresponding ab states become massless. In this situation, with additional massless vector bosons, we expect the world-volume gauge group to enhance beyond $U(1)^k$. The charges of the ab gauge bosons under the aa gauge symmetries correspond to the non-zero roots of the gauge group, which is easily checked to be $U(k)$. Hence, for k coincident Dp-branes the massless open string sector yields a $U(k)$ vector multiplet with respect to the 16 unbroken supersymmetries. In other words, in the configuration of k coincident D-branes the corresponding states are described by a $k \times k$ matrix, which represents their wavefunction with respect to Chan-Paton factors. That is, Chan-Paton factors receive a geometric interpretation as encoding on which branes the string is starting and ending.

Changing continuously the locations of the D-branes away from each other corresponds to turning on a vev for the diagonal components of the scalar fields on the D-branes. This produces a Higgs effect breaking the enhanced $U(k)$ gauge symmetry, generically to the Cartan subalgebra $U(1)^k$. This is in agreement with the interpretation of these scalars as coordinates of the D-branes in transverse space. In this respect, it is amusing (and possibly a very profound property of the nature of spacetime in string theory) that these coordinates become matrices (and therefore non-commutative) at distances of the order of the string scale (where the scalars in ab open string sectors become light).

We conclude by mentioning that the effective action for world-volume massless fields in coincident D-branes should be a non-abelian generalization of the above. This is not exactly known for the Dirac-Born-Infeld piece, due to ambiguities in the precise gauge trace structure prescription. In any event, at low energies the action reduces to non-abelian Yang-Mills interactions with coupling $g_{YM}^2 = g_s$.

A.3.4 Comments

We conclude the discussion of type II D-branes with some comments:

• Although we have centered on type IIB D-branes, the same kind of results hold for type IIA D-branes, namely the worldvolume massless fields gather in vector multiplets with respect to the 16 unbroken susys, and their dynamics is described by the Dirac-Born-Infeld plus Wess-Zumino action.

• Spacetime supersymmetric D-branes exist only for p odd in type IIB and p even in type IIA. For the reverse dimensions, no GSO projection can be introduced in the open string sector (in a way consistent with open-closed duality and the GSO in the closed sector). However, there exist non-supersymmetric D-branes with p odd in type IIA and p even in type IIB. They are non-supersymmetric, contains worldvolume tachyons, and are unstable against decay. We may study them in the lecture on stable non-BPS states in string theory.

• Recall that type IIA and IIB theories are T-dual once we compactify on $\mathbf{S^1}$. The action of T-duality of D-brane states is easy to obtain, since T-duality acts on open string boundary conditions by exchanging Dirichlet and Neumann boundary conditions (see lecture on T-duality for type I). This implies the mapping

IIB on $\mathbf{S^1}$ of radius R	IIA on $\mathbf{S^1}$ of radius $1/R$
wrapped D$(2k+1)$	unwrapped D$(2k)$
unwrapped D$(2k+1)$	wrapped D$(2k)$

D-brane states moreover form a multiplet under the perturbative T-duality groups in compactifications on $\mathbf{T^d}$. For instance, consider type II compactified on $\mathbf{T^6}$, which has a T-duality group $SO(6,6;\mathbf{Z})$, and consider 4d particle-like D-brane states. Type IIB theory contains 4d particle-like states arising from D1-branes wrapped in one of the internal $\mathbf{T^6}$ directions (6 states), from D3-branes wrapped in three internal directions (20 states) and from D5-branes wrapped in five internal directions (6 states). In total, we have 32 states, transforming in the spinor representation 32 of $SO(6,6;\mathbf{Z})$.

(Similarly, for type IIA we obtain 32 states from 1 D0-brane state, 15 D2-brane states, 15 D4-brane states adn 1 D6-brane state). These states, together with the perturbative states (momentum, winding) and other non-D-brane non-perturbative states (NS5-brane states, KK monopoles) fill out multiplets of the U-duality group $E_7(\mathbf{Z})$ as described in previous lecture.

• We would like to make a small remark on some D-branes which can be defined using the microscopic stringy description, and which were not encountered in the supergravity discussion.

- The type IIB D7-brane and the type IIA D8-branes change the asymptotic metric of spacetime, which is not flat, hence are not nicely described as asymptotically flat supergravity branes. The D7-brane is magnetically charged under the type IIB RR scalar a, which suffers a shift (monodromy) $a \to a+1$ in going around a D7-brane. This is consistent because the scalar is periodic, or equivalentely, because this transformation is an exact symmetry of IIB theory (in fact, in a subgroup of $SL(2, \mathbf{Z})$.

- Type IIA D8-brane is formally magnetically charged with respect to a (-1)-form. This simply means that it acts as a domain wall for a RR 0-form (the 'field strength') of type IIA theory, which is the cosmological constant, or mass parameter of massive IIA theory (Romans theory [76]).

- The type IIB D9-brane cannot be thought as a BPS non-perturbative state of type IIB theory, since it is charged under the RR 10-form and generates a tadpole rendering the theory inconsistent. Supersymmetric D9-branes only exist in the presence of O9-plane in type I theory, and in this situation they are present already in the vacuum, they are not an excited state of the theory. In the lecture on non-BPS states we will discuss excited states of type IIB theory with D9 - anti-D9 -brane pairs. These are excited states, but are not supersymmetric.

- Finally, the type IIB D(-1)-brane, which can be defined in the theory with spacetime euclidean signatures, is a sort of instanton, localized boht in space and in time. It is formally electrically charge under the type IIB RR scalar a, hence the instanton action is weighted by e^{ia}, so a acts as a theta parameter for type IIB theory.

A.4 D-branes in type I theory

A.4.1 Type I in terms of D-branes

Type I contains a sector of open strings already in its vacuum, with endpoints allowed to be anywhere in 10d spacetime. So in a sense it contains a set of (spacetime filling) D9-branes in the vacuum, on which these open strings end. These D-branes should thus not be regarded as excitated states above the vacuum, but part of it, since the theory is inconsistent without them. Nevertheless, these vacuum D9-branes are mathematically identical to the D-branes studied above, so it is useful to use the same language to describe them.

Indeed, both kinds of branes are in a sense related, as we described at the end of the lesson on T-duality for type I string theory. Recall that type I theory has one O9-plane (set of points fixed under the orientifold action Ω), and 32 D9-branes. Compactifying on $\mathbf{S^1}$ and performing a T-duality along it we obtain type I' theory, which is type IIA theory modded out by ΩR, with $R : x^9 \to -x^9$. It contains two O8 planes sitting at $x^9 = 0, \pi R$, and 32 D8-branes located at points in $\mathbf{S^1}$, which are part of the vacuum. However, taking the limit of infinite radius, keeping the D8-branes at a finite distance, the O8-planes go off to infinity and we are left with type IIA theory in flat 10d, with D8-branes. In this setup the D8-branes should be regarded as excitations over the type IIA vacuum.

The BPS D-branes of type I theory are the D5-brane and the D1-brane. In this section we obtain their world-volume modes, by quantizing the open string sectors of the configuration. Notice that other D-branes of type IIB theory, like the D3- or the D7-brane are projected out by the Ω projection and do not exist as BPS D-branes in type I theory.

A.4.2 Type I D5-brane

A useful reference for this section should be [104].

In principle, the computation of the world-volume spectrum for type I D5-branes is similar to that of type IIB D5-branes, with two new ingredients

i) In addition to the sector of open strings with both endpoints on the D5-branes, there is a sector of open strings with one end on the D5-branes and the other end on the vacuum D9-branes.

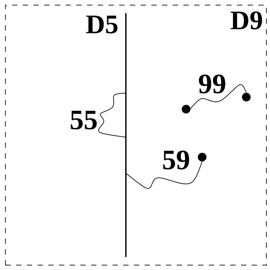


Figure A.9: Open string sectors in type I theory in the presence of a D5-brane.

ii) We need to impose the orientifold projection on the open string spectrum, since we are working with unoriented open strings

To deal with i), let us start first forgetting the Ω projection, and consider a system of k coincident D5-branes and N D9-branes in the oriented theory. The geometry of the directions spanned/transverse to the branes is depicted by lines/crosses as follows

	0	1	2	3	4	5	6	7	8	9
D9	-	-	-	-	-	-	-	-	-	-
D5	-	-	-	-	-	-	×	×	×	×

The geometry is shown in figure A.9. As we will see, the configuration preserves 8 supersymmetries, i.e. the equivalent of $\mathcal{N}=1$ 6d supersymmetry. This is the familiar criterion that some susy is preserved when the number of DN directions is a multiple of four.

In the sector of 55 strings, we obtain a $U(k)$ vector multiplet of 6d $\mathcal{N}=(1,1)$ supersymmetry, containing $U(k)$ gauge bosons, four real adjoint scalars, and two 6d Weyl fermions. In terms of 6d $\mathcal{N}=1$ supersymmetry, they correspond to a $U(k)$ vector multiplet (gauge boson plus one Weyl fermion) and an adjoint hypermultiplet (Weyl fermion plus four scalars).

In the sector of strings starting at the D5-branes and ending on the D9-branes (59 sector), the open strings have NN boundary conditions on the direction 2345 and DN conditions on 6789. The NN directions work as usual. For $X^i(\sigma,t)$, $i=6789$, we have

$$\partial_\sigma X^i(\sigma,t)|_{\sigma=0}=0 \quad ; \quad \partial_t X^i(\sigma,t)|_{\sigma=\ell}=0 \qquad \text{(A.15)}$$

Using the mode expansions for the DN directions, we obtain that the center of mass x^i is fixed at the location of the D5-brane; that momentum and winding are not allowed $p^i = 0$, $w^i = 0$; and that oscillator modding is shifted by 1/2 with respect to their usual values, namely 2d bosons have modes $\alpha^i_{n+1/2}$ and 2d fermions have modes $\psi^i_{n+\rho+1/2}$, with $\rho = 1/2, 0$ for NS, R.

The mass formula for 59 states is

$$\alpha' M^2 = N_B + N_F \tag{A.16}$$

since $E_0 = 0$ both in the NS and R sectors. In the NS sector, there are four fermion zero modes, along 6789, hence the massless groundstate is degenerate. Splitting the zero modes in creation and annihilation operators, and constructing the representaion of the zero mode Clifford algebra as usual, the GSO projection selects the massless groundstates

$$\begin{array}{cc} \text{State} & SO(4)_{6789} \\ |0\rangle & (-\frac{1}{2}, -\frac{1}{2}) \\ A^+_{a_1} A^+_{a_2}|0\rangle & (\frac{1}{2}, \frac{1}{2}) \end{array} \tag{A.17}$$

(A.18)

where the $SO(4)$ is the unbroken rotation group in 6789. These states are scalars under the 6d little group $SO(4)$. In the R sector, we have four fermion zero modes along 2345, hence the massless groundstate is degenerate. Splitting the zero modes in creation and annihilation operators, and constructing the representaion of the zero mode Clifford algebra as usual, the GSO projection selects the massless groundstates

$$\begin{array}{cc} \text{State} & SO(4)_{2345} \\ |0\rangle & (-\frac{1}{2}, -\frac{1}{2}) \\ A^+_{a_1} A^+_{a_2}|0\rangle & (\frac{1}{2}, \frac{1}{2}) \end{array} \tag{A.19}$$

(A.20)

these states are spinors under the 6d $SO(4)$ Lorentz little group. Gathering states from the 59 and 95 sectors (the latter are similar), we obtain one hypermultiplet of 6d $\mathcal{N} = 1$ supersymmetry. Noticing that the states carry D5- and D9- Chan-Paton labels, encoding on which D5- and on which D9-brane their endpoints lie, we realize the 6d $\mathcal{N} = 1$ hypermultiplet transforms in the bi-fundamental representation (N, k) under the D9- and D5-brane world-volume gauge groups.

Let us now address ii) and impose the orientifold projection. To make a long story short, let us simply say that we need to specify the action of Ω on the D9- and D5-brane Chan-Paton indices, via $N \times N$ and $k \times k$ matrices $\gamma_{\Omega,5}$, $\gamma_{\Omega,9}$, and that consistency requires $N = 32$ and [78]

$$\gamma_{\Omega,9} = \mathbf{1}_{32} quad; \quad \gamma_{\Omega,5} = \begin{pmatrix} 0 & \mathbf{1}_{k/2} \\ -\mathbf{1}_{k/2} & 0 \end{pmatrix} \tag{A.21}$$

Note that consistency requires k to be even.

The projections go as follows. In the 99 sector, all fields suffer a projection

$$\lambda = -\gamma_{\Omega,9} \lambda^T \gamma_{\Omega,9}^{-1} \tag{A.22}$$

and the surviving spectrum is the 10d $\mathcal{N} = 1$ $SO(32)$ vector multiplet.

In the 55 sector, the Ω action on oscillators along DD and NN directions differ by a sign. This follows from the definition of the action of Ω as $X^{\Omega}(\sigma, t) = X(-\sigma, t)$, and the mode expansions

$$\begin{aligned} X^{\mu}(\sigma, t) &= \ldots + \sum_{n \neq 0} \frac{\alpha_n^{\mu}}{n} \cos(\pi n \sigma/\ell)\, e^{-\pi n t/\ell} \\ X^{i}(\sigma, t) &= \ldots + \sum_{n \neq 0} \frac{\alpha_n^{i}}{n} \sin(\pi n \sigma/\ell)\, e^{-\pi n t/\ell} \end{aligned} \tag{A.23}$$

The Ω projections are different for the 6d $\mathcal{N} = 1$ hyper and vector multiplets. The surviving states must satisfy the conditions

$$\begin{aligned} \lambda = -\gamma_{\Omega,5} \lambda^T \gamma_{\Omega,5}^{-1} & \qquad \text{vect.mult.} \\ \lambda = \gamma_{\Omega,5} \lambda^T \gamma_{\Omega,5}^{-1} & \qquad \text{hypermult.} \end{aligned} \tag{A.24}$$

leading to a 6d $\mathcal{N} = 1$ $USp(k)$ vector multiplet, and one hypermultiplet in the two-index antisymmetric representation (which is reducible into a singlet and a representation of dimension $k(k-1)/2 - 1$).

Finally, the 59 sector is mapped to the 95 sector by Ω, so it is enough to keep the degrees of freedom in the 59 sector and not perform any projection. This leads to one half hypermultiplet of 6d $\mathcal{N} = 1$ susy in the representation $(k, 32)$ under $USp(k)_{55}$ and $SO(32)_{99}$. A half-hypermultiplet contains two real scalars and one Weyl fermion satisfying a reality condition, and only exists for multiplets in pseudo-real representation of the gauge group.

Some comments are in order

• The complete spectrum on the D5-brane worldvolume is

$USp(k)$ $\mathcal{N}=1$ vector multiplet
$\mathcal{N}=1$ hypermultiplet in $\begin{smallmatrix}\square\\\square\end{smallmatrix} + \frac{1}{2}(\square; 32)$

This 6d theory is chiral and miraculously free of anomlies. Again, another strong check of the self-consistency of string theory.

• As discussed in [104], a D5-brane provides the limit of zero size instantons in the D9-brane world-volume gauge theory. In fact, using the WZ couplings in the D9-brane theory, and instanton is charged under the RR 6-form C_6, exactly as a D5-brane.

$$\frac{1}{8\pi^2}\int_{10d} C_6 \wedge \mathrm{tr}\, F^2 \rightarrow k\int_{6d} C_6 \tag{A.25}$$

where $k=\frac{1}{8\pi^2}\int \mathrm{tr}\, F^2$ is the instanton number. Instanton have a bosonic zero mode which parametrizes their size. In the limit of zero size, the instanton is pointlike in four dimensions and is exactly described by a D5-brane.

A.4.3 Type I D1-brane

One can perform a similar computation of the world-volume massless spectrum for D1-branes. We consider a configuration of N D9-branes and k coincident D1-branes, the geometry is described by

	0	1	2	3	4	5	6	7	8	9
D9	-	-	-	-	-	-	-	-	-	-
D1	-	-	×	×	×	×	×	×	×	×

The configuration preserves 8 supersymmetries, more specifically $\mathcal{N}=(0,8)$ susy in the 2d volume of the D1-brane.

Before the orientifold projection, the 99 massless sector leads to the 10d $\mathcal{N}=1$ $U(N)$ vector multiplet; the 11 massless sector lead to the 2d $\mathcal{N}=(8,8)$ $U(k)$ vector multiplet. In the 19+91 sector, we have DN boundary conditions along the 8 light-cone directions; the moddings of oscillators are as for the DN directions discussed above, and the mass formula for 19 states is

$$\alpha' M^2 = N_B + N_F + E_0 \tag{A.26}$$

with $E_0 = 1/2, 0$ for the NS, R sectors. IN the NS sector, all states are massive. Massless states only arise from the R sector groundstate, which is unique since there are no fermion zero modes. The 19 and 91 groundstates behave as a 2d spinor, and transform in the representation (k, N) under the D1- and D9-brane gauge groups.

Let us now impose the orientifold projection. In this case, consistency requires

$$\gamma_{\Omega,9} = \mathbf{1}_{32} quad; \quad \gamma_{\Omega,1} = \mathbf{1}_k \tag{A.27}$$

The projections go as follows. In the 99 sector, all fields suffer a projection

$$\lambda = -\gamma_{\Omega,9} \lambda^T \gamma_{\Omega,9}^{-1} \tag{A.28}$$

and the surviving spectrum is the 10d $\mathcal{N} = 1$ $SO(32)$ vector multiplet.

In the 11 sector, the Ω action on oscillators along DD and NN directions differ by a sign. The Ω projections are different for the 2d $\mathcal{N} = (0, 8)$ vector multiplet (2d gauge bosons plus 8 left-moving 2d chiral fermions) and the 2d $\mathcal{N} = (0, 8)$ chiral multiplet (8 real scalars plus 8 2d chiral right-moving fermions). In fact we have

$$\begin{aligned} \lambda = -\gamma_{\Omega,1} \lambda^T \gamma_{\Omega,1}^{-1} \qquad & \text{vect.mult.} \\ \lambda = \gamma_{\Omega,1} \lambda^T \gamma_{\Omega,1}^{-1} \qquad & \text{ch.mult.} \end{aligned} \tag{A.29}$$

leading to a 2d $\mathcal{N} = (0, 8)$ $SO(k)$ vector multiplet and a 2d $\mathcal{N} = (0, 8)$ chiral multiplet in the two-index symmetric representation (which is reducible into a singlet and a representation of dimension $k(k + 1)/2 - 1$).

Finally, the 19 sector is mapped to the 91 sector by Ω, so it is enough to keep the degrees of freedom in the 19 sector and not perform any projection. This leads to one 2d chiral (left-moving) spinor, with just one component, in the representation $(k, 32)$ under $SO(k)_{11}$ and $SO(32)_{99}$. This is sometimes called a Fermi multiplet of 2d $\mathcal{N} = (0, 8)$ susy.

Some comments are in order

• The complete spectrum on the D1-brane worldvolume is

$SO(k)$ $\mathcal{N} = (0, 8)$ vector multiplet (gauge boson plus 8 left fermions)
$\mathcal{N} = (0, 8)$ chiral multiplet (8 scalars plus 8 right fermions)
$\mathcal{N} = (0, 8)$ Fermi multiplet (8 left fermions) in $(\square; 32)$

This 2d theory is chiral and miraculously free of anomlies. Yet another strong check of the self-consistency of string theory.

• As discussed in later lectures, this content will provide support for the interesting duality conjecture for the strong coupling regime of type I theory.

A.5 Final comments

We have shown that a detailed treatment of Dp-branes is possible from our microscopic description. It allows to rederive the results from the supergravity analysis of solitons, and to obtain new results, like the detailed world-volume theories, the appearance of enhanced gauge symmetries, etc.

Many other interesting phenomena appear in configurations with D-branes. For instance the existence of bound states of D-branes of different dimensions, configurations where D-branes end on D-branes, the D-brane dielectric effect, etc. D-branes properties is one of the hot topics in todays string theory. In the following lectures we will become familiar with some of them.

Appendix B

String theories at strong coupling and string duality

In this lecture we mainly follow section 14 of [71].

B.1 Introduction

The perturbative picture of the different superstring theories is shown in figure B.1. There are five different theories, some ofwhich are related by perturbative string dualities (T-duality) upon compactification.

Our purpose in this lecture is to study the strong coupling limit of these theories. We will find out that this limit is surprisingly quite simple, and is usually described in terms of a weakly coupled dual theory. In this description further, non-perturbative, dualities relate all the different string theories. This implies that the different perturbative string theories all arise in different limits of a unique underlying theory, as some moduli are tuned. The situation is shown in figure B.2. This is analogous to how 10d type IIA and IIB are recovered starting from a unique theory (type II on $\mathbf{S^1}$) in the two limits of large radius and small radius (large T-dual radius).

The main tool used in the exploration of the strong coupling regime is to follow the properties of BPS states as the coupling becomes strong. This can be done because such properties are protected by the supersymmetry of these states. Some of these states, which are non-perturbative and very heavy in the weakly coupled regime, become light in the strong coupling regime, and correspond to the states that dominate this regime, and provide

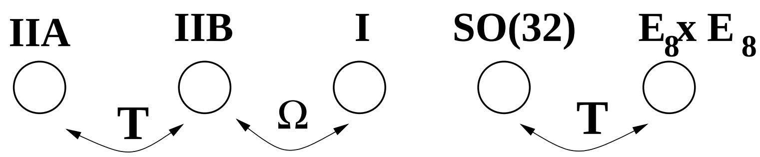


Figure B.1: The different 10d supersymmetric superstring theories in perturbation theory.

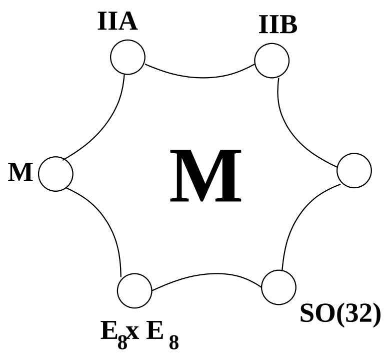


Figure B.2: Map of the moduli space of the underlying theory and its different known limits.

the elementary, perturbative, degrees of freedom of the dual theory, which is weakly coupled in that regime.

An intuitive argument indicating which BPS states dominate the dynamics at strong coupling is to associate an scale to them. For instance, the tension of a Dp-brane $T_p \simeq \alpha'^{-(p+1)/2}/g_s$ defines a mass scale $M \simeq \alpha'^{-1/2} g_s^{-1/(p+1)}$. This implies that in the strong coupling limit the lightest mass scale corresponds to the lowest p Dp-brane, suggesting these are the states dominating the low-enery dynamics in that regime.

B.2 The type IIB $SL(2, \mathbf{Z})$ self-duality

The basic reference is [104], see also [86].

B.2.1 Type IIB S-duality

At strong coupling, the lightest mass scale is set by the D1-brane states, $M \simeq \alpha'^{-1/2} g_s^{-1/2}$. The fact that these objects dominate the dynamics at strong coupling suggests that the strong coupling limit is described by a string theory. It is reasonable to imagine that it moreover corresponds to a weakly interacting string theory, hence it should correspond to one of the string theories we have studied. And the only string theories with the correct amount of supersymmetry are the type IIA and type IIB theories.

The natural proposal is that in fact the strong coupling limit of type IIB theory is described by a weakly coupled dual type IIB theory. In fact, in the low-enery limit the theory is described by type IIB supergravity, which is known to have a symmetry relating weak and strong coupling. The action of this symmetry, known as S-duality, relates the massless fields of the theory at coupling g_s (denoted as unprimed) with those of the theory at coupling $g_s' = 1/g_s$ (denoted as primed), as follows

$$a' = a \quad ; \quad \phi' = -\phi \quad ; \quad B_2' = \tilde{B}_2 \quad ; \quad \tilde{B}_2' = b_2$$
$$C_4' = C_4 \quad ; \quad G' = G \tag{B.1}$$

where G is the metric in the Einstein frame. The reason why we can trust the form of the type IIB supergravity action is that its form is fixed by supersymmetry (up to higher derivative terms, which are not relevant at low energies).

The proposal is that this symmetry of the supergravity limit is an exact symmetry of the full string theory! As a consequence, the theory at $g_s \to \infty$ is described by a perturbative type IIB theory, the transformed under S-duality, which is weakly coupled $g_s' \to 0$.

B.2.2 Additional support

We would like to mention additional evidence supporting this proposal.

• The D1-branes in the original theory are the fundamental strings of the dual one. Therefore the D1-brane 2d world-volume theory should be of the same kind as that of a fundamental type IIB string. In fact, D-brane worldvolume spectra were computed in previous lecture. For a D1-brane we have a 2d $U(1)$ gauge boson (which is non-dynamical in 2d), 8 2d real scalars $X^i(\sigma^1,\sigma^2)$ in the 8_V of the transverse $SO(8)$ Lorentz group, and 8 2d fermions $\Theta^\alpha(\sigma^1,\sigma^2)$, transforming in the 8_C of $SO(8)$. This is precisely the 2d field content of a type IIB fundamental string in the Green-Schwarz formalism (see comment in page 19 in lecture on type II superstrings).

- The BPS states of both theories agree. For instance

IIB at g_s		IIB at $g_s' = 1/g_s$
F1	$\longleftrightarrow$	D1
D1	$\longleftrightarrow$	F1
D3	$\longleftrightarrow$	D3
D5	$\longleftrightarrow$	NS5
NS5	$\longleftrightarrow$	D5

The tensions and charges of the objects match. Also, they have equivalent world-volume field theories, as we have seen for the D1/F1 and as follows from the discussion of world-volume modes for the D5/NS5-branes in the lecture on non-perturbative states in string theory.

B.2.3 $SL(2,\mathbf{Z})$ duality

In fact, type IIB supergravity has a larger symmetry group, $SL(2,\mathbf{R})$, the group of unit determinant 2×2 real matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Introducing the type IIB complex coupling $\tau = a + ie^{-\phi}$, one such transformation relates the theory at coupling τ to the theory at coupling τ', by the following action on the massless fields

$$\begin{aligned} \tau' &= \frac{a\tau+b}{c\tau+d} \\ \begin{pmatrix} B_2' \\ \tilde{B}_2' \end{pmatrix} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} B_2 \\ \tilde{B}_2 \end{pmatrix} \\ G' &= G \quad , \quad C_4' = C_4 \end{aligned} \tag{B.2}$$

As we argued two lectures ago, not the full $SL(2,\mathbf{R})$ can be an exact symmetry of the quantum theory, since it does not respect the discrete lattice of charges of brane states in the theory. However, an $SL(2,\mathbf{Z})$ subgroup of it does respect it, and we propose that it is an exact symmetry of the full quantum theory. In fact, this group is generated by the above S-duality transformation $\tau \to -1/\tau$, and a transformation $\tau \to \tau+1$, which simply shift $a \to a+1$ leaving all other fields invariant. The latter is knonw to be a symmetry to all orders in perturbation theory (it is a 0-form gauge field, so gauge invariance implies that it has no non-derivative couplings), so it is natural to propose that it is a symmetry at the non-perturbative level.

This proposal has several interesting implications.

• For instance, it implies the existence of an infinite set of points in the τ moduli space which are related by $SL(2,\mathbf{Z})$ to weak coupling; that is, whose dynamics is equivalent to a perturbative IIB string theory once described in suitable $SL(2,\mathbf{Z})$ dual variables.

• It implies that the spectrum of BPS states in type IIB string theory must arrange in $SL(2,\mathbf{Z})$ multiplets. In particular, it must contain an $SL(2,\mathbf{Z})$ orbit of string-like objects, denoted (p,q)-strings. The $(1,0)$- and $(0,1)$-strings correspond to the F1- and D1-strings. Indeed, at a point τ dual to weak coupling by an $SL(2,\mathbf{Z})$ duality, the object becoming the perturbative one is the (p,q)-string related to the F1-string by the same $SL(2,\mathbf{Z})$-transformation. Simlarly we have (p,q) 5-branes; in these cases the p,q labels transform under $SL(2,\mathbf{Z})$ as a doublet, which means that a (p,q) object can be regarded as a bound state of p $(1,0)$ objects and q $(0,1)$ objects. There are also (p,q) 7-branes, but they have a more involved $SL(2,\mathbf{Z})$ transformation rule and cannot be properly regarded as bound states of the 'elementary' solitons. The existence of these (p,q)-branes as supergravity solitons is guaranteed from the fact that $SL(2,\mathbf{Z})$ is a subgroup of the supergravity symmetry group.

Toroidal compactification has been already discussed in the lecture on non-perturbative objects in string theory. So we refer the reader to the corresponding section.

B.3 Type IIA and M-theory on $\mathbf{S^1}$

The original paper discussing this is [104]

B.3.1 Strong coupling proposal

The type IIA theory strong coupling dynamics at low energies is dominated by the D0-branes, with a mass scale of $M \simeq \alpha'^{-1/2} g_s^{-1}$. There is no BPS string becoming light in the strong coupling regime, and this suggests that the strong coupling limit is *not* described by a string theory. Instead what one finds at strong coupling is that states with n D0-branes form an infinite tower of states, with massess $M_n \simeq \frac{n}{g_s} M_s$, which is becoming extremely light.

This suggests that the strong coupling limit is a decompactification limit of some 11d theory.

Indeed, there exists and 11d supergravity theory with the correct amount of supersymmetry (32 supercharges), and which upon Kaluza-Klein compactification on a circle of radius R leads to an effective theory (neglecting KK replicas of massless modes) given by 10d type IIA supergravity, with

$$g_s = (M_{11}R)^{3/2} \tag{B.3}$$

where M_{11} is the 11d Planck scale, and R is measured in the 11d metric.

More explictly, 11d supergravity is described by a metric G, a 3-form $C_3$and and 11d gravitino. The matching of massless 11d fields and massless type IIA 10d fields is

$$\begin{array}{lcl} G_{MN} & \longrightarrow & G_{\mu\nu} \\ & & G_{\mu,10} \rightarrow A_\mu \\ & & G_{10,10} \rightarrow \phi \\ C_{MNP} & \longrightarrow & C_{\mu\nu\rho} \\ & & C_{\mu\nu,10} \rightarrow B_{\mu\nu} \\ \Psi_{M,\alpha} & \longrightarrow & \psi_{\mu\alpha}, \psi_{\mu\dot{\alpha}}, \psi_{10,\alpha}, \psi_{10,\dot{\alpha}} \end{array} \tag{B.4}$$

This suggests that type IIA at strong coupling is a new 11d quantum theory, whose low energy limit is 11d supergravity. The microscopic nature of this 11d theory is completely unknown (let us emphasize again that it it not a string theory), and it is simply called M-theory. The facts we know about M-theory are

• At low energies it reduces to 11d supergravity

• It contains 1/2 BPS states corresponding to a 2-brane and a 5-brane (denoted M2- and M5-brane). These can be constructed as BPS solutions of the 11d supergravity equations of motion, and argued to exist in the full microscopic theory (whatever it is) due to their BPS property.

• M-theory compactified on a circle of radius R is completely equivalent to full-fledged type IIA string theory at coupling $g_s = (M_{11}R)^{3/2}$.

It is interesting to point out that M-theory in 11d does not have any scalar field, and consequently does not have any dimensionless coupling constant. This means that there is no parameter which can be taken small to obtain a perturbative description, so the theory is intrinsecally non-perturbative.

Once compactified on a circle, however, there is one dimensionless quantity $M_{11}R$, which can be taken small to lead to a perturbative theory: this is precisely perturbative type IIA string theory.

B.3.2 Further comments

Let us provide some additional support for the proposal. For instance there is a precise matching of BPS states in both theories, as follows [1]

IIA at g_s	$\longleftrightarrow$	M-theory on $\mathbf{S^1}$ of radius R
D0-branes	$\longleftrightarrow$	KK momenta of 11d supergravity multiplet
F1	$\longleftrightarrow$	wrapped M2
D2	$\longleftrightarrow$	unwrapped M2
D4	$\longleftrightarrow$	wrapped M5
NS5	$\longleftrightarrow$	unwrapped M5
D6	$\longleftrightarrow$	Kaluza-Klein monopole

The tensions of these objects agree completely, and it is possible to show that they have equivalent world-volume field theories. In particular one can show that the worldvolume theory of an M2-brane wrapped on $\mathbf{S^1}$ reduces to the world-sheet theory of a fundamental type IIA string.

We would like to mention that the M-theory proposal implies very interesting properties for the D0-branes, since they are, from an 11d viewpoint, simply 11d gravitons (and partners) with non-zero momentum along the circle. For instance, a 11d graviton with n units of momentum is not the same state as n 11d gravitons with 1 unit of momentum each, although the have the same mass and charge. This implies that there should exist a bound states of n D0-branes (with zero binding energy) in type IIA theory. Moreover, scattering of this kind of states should reproduce the supergravity interactions in 11d!

This line of though has led to a proposal to define microscopically M-theory, known as the M(atrix) theory proposal [79]. It is based on describing the complete dynamics of 11d M-theory from the world-line gauge theory on stacks of D0-branes. This is a 1d quantum mechanics of $U(n)$ gauge fields, and its partners under the 16 unbroken supersymmetries (9 scalars

[1] The D8-brane is however a bit problematic, since it is a source of the type IIA mass parameter, and there is no 11d version of supergravity which reduces to massive IIA theory. This is in a sense an open issue.

and fermions). In this description, spacetime arises as the moduli space of scalars in the 1d theory; 11d gravitons are bound states in this quantum mechanics system; scattering of supergravity modes in 11d is recovered by interactions of wavepackets of bound states.

M(atrix) theory has led to very interesting results in 11d and in compactifications preserving enough supersymmetry (toroidal compactifications, etc). However, difficulties have typically arisen in trying to study more involved situations with less supersymmetry.

B.4 M-theory on $\mathbf{T^2}$ vs type IIB on $\mathbf{S^1}$

The original discussion is in [80].

There must be a direct link between M-theory compactified on $\mathbf{T^2}$ and type IIB compactified on $\mathbf{S^1}$. This can be seen by regarding M-theory on $\mathbf{T^2}$ as type IIA on $\mathbf{S^1}$ and performing a T-duality to type IIB on (a T-dual) $\mathbf{S^1}$.

11d			M
			$\downarrow \mathbf{S^1}$
10d	IIB		IIA/M
	$\downarrow \mathbf{S^1}$		$\downarrow \mathbf{S^1}$
9d	IIB	$\stackrel{T}{\longleftrightarrow}$	IIA/M

We can perform the matching of both theories even when the circles are not small, and propose they are equivalent, with the following relations.

Moduli: The τ complex coupling of type IIB theory matches with the complex structure parameter of the M-theory $\mathbf{T^2}$, $\tau = \frac{R_1}{R_2} e^{i\theta}$ (see figure B.3). The radius R of the IIB $\mathbf{S^1}$ is related to the area of the M-theory $\mathbf{T^2}$ $A = R_1 R_2$ by $M_{11}^3\, A = 1/R$.

Duality groups: The $SL(2,\mathbf{Z})$ duality group of type IIB theory (already present in 10d) matches the $SL(2,\mathbf{Z})$ invariance group of the $\mathbf{T^2}$ geometry, corresponding to large diffeomorphisms of $\mathbf{T^2}$. This is hence a nice geometric interpretation for the IIB self-duality group.

BPS states: Let us give some examples on the matching of BPS states

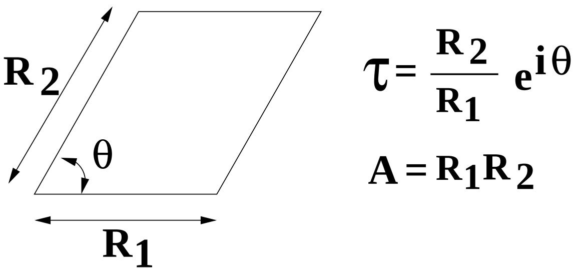


Figure B.3: Complex structure parameter and area of a two-torus.

IIB on $\mathbf{S^1}$	$\longleftrightarrow$	M-theory on $\mathbf{T^2}$
unwrapped (p, q) string	$\longleftrightarrow$	M2 wrapped on (p, q) cycle in $\mathbf{T^2}$
wrapped (p, q) string	$\longleftrightarrow$	KK momentum of 11d sugra multiplet along (p, q) direction
momentum in $\mathbf{S^1}$	$\longleftrightarrow$	M2 wrapped on $\mathbf{T^2}$
wrapped (p, q) 5-brane	$\longleftrightarrow$	M5-brane wrapped on (p, q) cycle
unwrapped (p, q) 5-brane	$\longleftrightarrow$	KK-monopole with isometry along (p, q)
unwrapped D3	$\longleftrightarrow$	M5 wrapped on $\mathbf{T^2}$
wrapped D3	$\longleftrightarrow$	unwrapped M2

The tensions of all objects agree, and they have equivalent world-volume theories.

Hence type IIB on $\mathbf{S^1}$ with radius R and coupling τ is equivalent to M-theory on a $\mathbf{T^2}$ with complex structure τ and area $A \simeq 1/R$. In particular notice that the decompactified 10d type IIB string theory can be obtained by taking M-theory on a $\mathbf{T^2}$ in the limit of vanishing area. In this limit, a tower of light states arises from M2-branes wrapped on $\mathbf{T^2}$, these are interpreted as the KK modes on the $\mathbf{S^1}$ of the dual IIB theory, which is in the decompactification limit.

B.5 Type I / $SO(32)$ heterotic duality

See [104] and [81].

B.5.1 Strong coupling of Type I theory

In 10d type I theory at strong coupling, the lightest mass scale is set by the D1-branes, with $M \simeq \alpha'^{-1/2} g_s^{-1/2}$. This suggests that the strong coupling behaviour is controlled by a string, and that the strong coupling limit may correspond to a dual string theory.

From the amount of supersymmetry, the dual string theory could be a dual type I theory, or a dual $SO(32)$ or $E_8 \times E_8$ heterotic theory. However, the D1-string of the original theory is BPS, so the dual string theory should have an F1 BPS state. This is not present in type I theory, so the strong coupling dynamics cannot correspond to a dual type I theory. Out of the two heterotic theories, the fact that the $SO(32)$ heterotic has the same gauge group as type I theory suggests that it is the correct candidate to describe the strong coupling limit of type I.

In fact, restricting to low energies, the low energy supergravity action for type I and $SO(32)$ heterotic theories is the same, up to redefinitions of the fields, as follows.

$$\begin{aligned} \phi_{\rm typeI} &= -\phi_{\rm het.} \quad \to \quad (g_s)_{\rm het} = 1/(g_s)_{\rm typeI} \\ G_{\rm typeI} &= e^{-\phi_{\rm het.}} G_{\rm het.} \quad , \quad (A_{SO(32)})_{\rm typeI} = (A_{SO(32)})_{\rm het.} \quad , \quad (H_{3,RR})_{\rm typeI} = (H_3)_{\rm het.} \end{aligned} \tag{B.5}$$

This suggest that the type I theory at coupling g_s is exactly equivalent to the $SO(32)$ heterotic at coupling $1/g_s$. And in particular that the strong coupling limit of type I theory is described by a weakly coupled $SO(32)$ heterotic string theory, and viceversa.

B.5.2 Further comments

B.5.3 Additional support

We would like to mention additional evidence supporting this proposal.

• The D1-branes in the original type Itheory are the fundamental strings of the dual heterotic theory. Therefore the type I D1-brane 2d world-volume theory should be of the same kind as that of a fundamental $SO(32)$ heterotic string. In fact, D-brane worldvolume spectra were computed in previous lecture. For a D1-brane, in the 11 sector we have an $O(1) = \mathbf{Z_2}$ gauge symmetry, 8 2d real scalars $X^i(\sigma^1, \sigma^2)$ in the 8_V of the transverse $SO(8)$ Lorentz group, and 8 2d rightmoving chiral fermions $\Theta^\alpha(\sigma^1 - \sigma^2)$, transforming in the 8_C of $SO(8)$. In addition in the 19 and 91 sectors we have 32 2d left-moving chiral fermions $\lambda^I(\sigma^1 + \sigma^2)$, singlets under the Lorentz $SO(8)$ and transforming in the fundamental of the $SO(32)$ spacetime group. This is precisely the 2d field content of a heterotic fundamental string in the Green-Schwarz formalism. The fact that it is the $SO(32)$ follows from the fact that the fermions λ^I are odd under the $\mathbf{Z_2}$ gauge symmetry, and so in building gauge invariant states of the 2d theory they suffer a GSO projection acting in the same way on the 32 2d internal fermions.

• The BPS states of both theories agree. For instance

type I at g_s		$SO(32)$ Heterotic at $g_s' = 1/g_s$
D1	$\longleftrightarrow$	F1
D5	$\longleftrightarrow$	NS5

We would like to conclude with a comment. The $SO(32)$ heterotic theory contains massive states in the spinor representation of $SO(32)$, of dimension 2^{15}. They correspond to states with internal 16d momentum

$$P = \frac{1}{2}(\pm, \ldots, \pm) \quad \text{with even number of minu signs.} \tag{B.6}$$

These states are non-BPS, but are stable due to charge conservation (there are no states lighter than them with the same charge). A prediction of heterotic/typeI duality is that states with those quantum numbers exist in type I theory. These do not appear in perturbative type I theory, or in the non-perturbative BPS states. In the lecture on stable non-BPS D-branes we will discuss the nature of these objects.

B.6 M-theory on $\mathbf{S^1}/\mathbf{Z_2}$ / $E_8 \times E_8$ heterotic

The strong coupling limit of the $E_8 \times E_8$ heterotic is difficult to analyze directly, as we will understand later on. It is somewhat easier (although highly non-trivial) to derive it starting from the discussion of compactifications of M-theory in the unique compact 1d space which is not the circle: the interval. The discussion follows the original paper [82] (wee [83] for more advanced discussions).

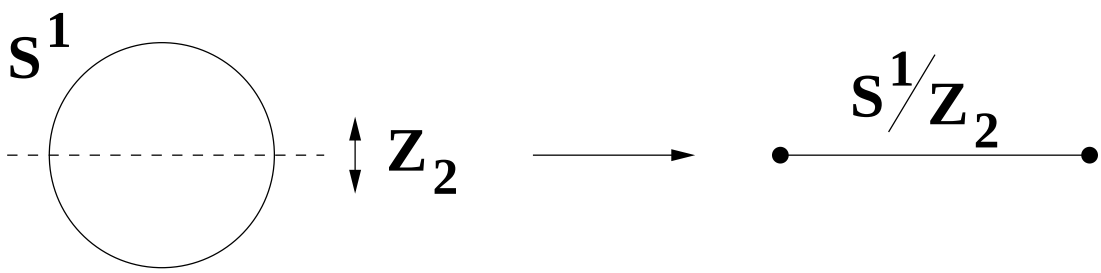


Figure B.4: The quotient of a circle by a reflection under a diameter is an interval $I = \mathbf{S^1}/\mathbf{Z_2}$.

B.6.1 Horava-Witten theory

Consider the compactification of M-theory on $\mathbf{S^1}$, modded out by a $\mathbf{Z_2}$ action, with generator acting by

$$\begin{aligned} \theta : \quad & x^{10} \to -x^{10} \\ & C_3 \to -C_3 \end{aligned} \tag{B.7}$$

which is a symmetry of the theory (at least at the supergravity level, so we are assuming implicitly this to be a symmetry of microscopic M-theory). The action on C_3 is required so that the term in the 11d supergravity action $\int_{11d} C_3 \wedge G_4 \wedge G_4$ (with $G_4 = dC_3$) is invariant.

The quotient space is an interval (see figure B.4), so that spacetime has two 10d boundaries sitting at $x^{10} = 0, \pi R$.

It is important to understand the we do not have a microscopic description of M-theory, and such a description would be required to construct an orbifold of M-theory from first principles. This is because at the fixed points of the orbifold (the boundaries of spacetime) there may be additional states which are not obtained simply from the effective field analysis. They would be the analogues of twisted sectors in string theory constructions. We will not be able to obtaine these states from first principles, but happily the consistency condition of cancellation of anomalies will be enough to show that the existence of these states, and their precise spectrum.

Let us start constructing the orbifold. We expect that the 10d theory will contain a sector given by the $\mathbf{Z_2}$ invariant states in the compactification of M-theory on $\mathbf{S^1}$ (this is the analogue of the untwisted sector in string theory orbifolds). Ignoring KK replicas, we have

11 field		10 field	$\mathbf{Z_2}$ parity	Surviving field
G_{MN}	$\longrightarrow$	$G_{\mu\nu}$	$+$	$G_{\mu\nu}$
		$G_{\mu,10} \to A_\mu$	$-$	—
		$G_{10,10} \to \phi$	$+$	ϕ
C_{MNP}	$\longrightarrow$	$C_{\mu\nu\rho}$	$-$	—
		$C_{\mu\nu,10} \to B_{\mu\nu}$	$+$	$B_{\mu\nu}$
$\Psi_{M,\alpha}$	$\longrightarrow$	$\psi_{\mu\alpha}, \psi_{\mu\dot{\alpha}}, \psi_{10,\alpha}, \psi_{10,\dot{\alpha}}$	$+, -, +, -$	$\psi_{\mu\alpha}, \psi_{10,\dot{\alpha}}$

The content of massless 10d surviving fields is exactly that of 10d $\mathcal{N}=1$ supergravity. This content is chiral, and leads to 10d anomalies, hence the theory as it stands is inconsistent.

If M-theory is consistent at the quantum level it should lead to an additional set of states. Moreover, one can check that from the 11d viewpoint the anomalies are localized on the 10d fixed locus of the orbifold. This is because in the bulk of the spacetime away from the boundaries the local dynamics is still described by 11d M-theory, which is non-chiral, while it is at the boundaries that the orbifold projection introduces chirality. The new fields cancelling the anomaly must be localized on the orbifold fixed points, as expected.

From our discussion of anomalies in heterotic theories, we know that there are two possible sets of fields that can cancel (in a very miraculous way) the anomaly of the 10d $\mathcal{N}=1$ supergravity multiplet. One of them is a 10d $\mathcal{N}=1$ $SO(32)$ vector multiplet, and the other is a 10d $\mathcal{N}=1$ $E_8 \times E_8$ vector multiplet. Clearly only the later set of fields can be split into two fixed points and cancel the two sources of anomaly, so they provide the only candidate set of multiplets that M-theory must contain in order to lead to a consisten compactification.

That is, compactification of M-theory on the interval $\mathbf{S^1}/\mathbf{Z_2}$ contains one E_8 10d $\mathcal{N}=1$ vector multiplet at each of the two 10d boundaries of spacetime, see figure B.5. This is known as Horava-Witten theory or Horava-Witten compactification of M-theory.

Notice that the final theory has the same massless spectrum as the $E_8 \times E_8$ heterotic theory. Moreover, the effective action of both theories is determined by supersymmetry, and agrees if the heterotic string coupling constant and the M-theory radius are related by $g_s = (M_{11}R)^{3/2}$. It is then natural to propose that the $E_8 \times E_8$ heterotic string theory at coupling g_s is completely equivalent to M-theory on $\mathbf{S^1}/\mathbf{Z_2}$ with radius R, related to g_s as above.

The strong coupling regime of the $E_8 \times E_8$ heterotic string theory corre-

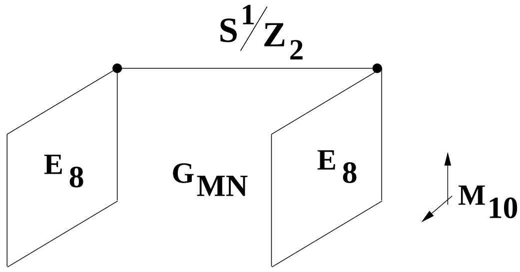


Figure B.5: The strong coupling description of $E_8 \times E_8$ heterotic involves the compatification of M-theory on a space with two 10d boundaries. Gravity propagates in 11d, while gauge interactions are localized on the 10d subpaces at the boundaries.

sponds to the large radius limit of the M-theory compactification. We can now understand why it is difficult to determine directly the strong coupling regime directly. The sign of the opening up of the extra dimension is the appearance of KK momentum modes, but these are not BPS states, due to the $\mathbf{Z_2}$ projection in M-theory language: the gauge boson that would carry the charge of these states is projected out by $\mathbf{Z_2}$; equivalentely, momentum is not a conserved charge due to violation of translational invariance in the $\mathbf{S^1}$ due to the existence of preferred points (the orbifold fixed points).

B.6.2 Additional support

We can also match BPS states in the two theories, as follows

$E_8 \times E_8$ heterotic at g_s		M-theory on $\mathbf{S^1}/\mathbf{Z_2}$ at R
F1	$\longleftrightarrow$	wrapped M2 (see fig. B.6)
NS5	$\longleftrightarrow$	unwrapped M5

Notice that other states in M-theory on $\mathbf{S_1}$, which are projected out by $\mathbf{Z_2}$ (like a warpped M5-branes, or an unwrapped M2-branes) are correctly absent in heterotic theory.

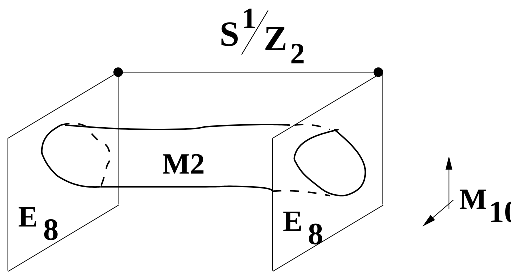


Figure B.6: The fundamental heterotic string arising in the Horava-Witten viewpoint from a M-theory M2-brane stretched along the interval. Note that it knows about the existence of the two E_8's, thus explaining why the heterotic string has (in the fermionic formulation) 2d fermions charged under the cartans of both group factors.

B.7 $SO(32)$ het/typeI on $\mathbf{S^1}$ vs M-theory on $\mathbf{S^1} \times (\mathbf{S^1}/\mathbf{Z_2})$

In this section we describe a relation between Horava-Witten theory compactified to 9d on a circle with type I theory. We will find that the type I picture in terms of D-branes, in a T-dual version, provides further insight into the appearance of the E_8 gauge multiplets on the boundaries of the interval. See section 14.5 in [71].

We consider the following chain of dualities

11d							M
							$\downarrow \mathbf{S^1}/\mathbf{Z_2}$
10d			type I	$\overset{S}{\longleftrightarrow}$	$SO(32)$ het		$E_8 \times E_8$/HW
			$\downarrow \mathbf{S^1}$		$\downarrow \mathbf{S^1}$		$\downarrow \mathbf{S^1}$
9d	type I'	$\overset{T}{\longleftrightarrow}$	type I	$\longleftrightarrow$	$SO(32)$ het	$\overset{T}{\longleftrightarrow}$	$E_8 \times E_8$/HW

Following the duality carefully allows to derive the Horava-Witten picture from type I' theory on $\mathbf{S^1}$.

T-duality relates the 9d $SO(32)$ and $E_8 \times E_8$ heterotic theories if there are Wilson lines turned on, breaking the gauge group to $SO(16)$ (see lecture on toroidal compactification of heterotic strings). We can now use the S-dual version of $SO(32)$ heterotic theory, and relate type I on $\mathbf{S^1}$ with Wilson lines breaking to $SO(16)^2$ with $E_8 \times E_8$ heterotic theory.

In fact, it is more useful to use the T-dual of type I theory, namely type I' theory, where the Wilson lines correspond to D8-brane positions (see lecture on type I toroidal compactification). We are interested in locating 16 D8-branes on top of each of the two O8-planes in the ΩR quotient of type IIA on $\mathbf{S^1}$, a configuration which leads to $SO(16)^2$ gauge group.

Thus we have a relation between IIA modded out by ΩR, $R : s^9 \to -x^9$ (with $SO(16)$ gauge multiplets on top of each of the fixed points of R) and $E_8 \times E_8$ theory (on $\mathbf{S^1}$ with Wilson lines breaking to $SO(16)^2$. We now only need to identify in type I' language what is the limit that corresponds to taking large $\mathbf{S^1}$ radius and strong coupling in the heterotic side. It can be seen to correspond also to large radius and large coupling in type I' picture.

Recall now that in the bulk of the type I' theory, away form the O8-planes, the local dynamics is that of type IIA theory. Since we are taking a strong coupling limit, a new dimension will open up (D0's are beoming light), lifting our configuration to M-theory. We recover a picture of M-theory on $\mathbf{S^1}/\mathbf{Z_2}$ (and on large circle). At the same time, we should see our $SO(16)$ gauge groups enhancing to E_8's. Indeed this is the case: near the O8-planer there are stuck D0-branes (which cannot move off into the bulk), which lead to additional light particles (in vector multiplets) transforming in the chiral spinors representation 128 of $SO(16)$ and enhancing the group to E_8 [2]

The result is exactly the Horava-Witten picture. The advantage of the present approach is that it provides a more intuitive interpretation of the E_8 gauge multiplets living on the boundaries of spacetime. The type I' picture has managed to make part of these multiplets perturbative and familiar.

Another additional advantage of the present picture is that is clarifies a little bit the role of D8-branes in the lift to M-theory, at least in this particular context. Another important feature of this picture is that it allows to understand some subtle details in the matching with heterotic string theory (namely the appearance of exceptional gauge symmetries), but these are beyond the scopes of these notes, see [81] for details.

[2]The open string sector of 08 and 80 strings leads to fermionic zero modes on the D0-brane worldline, transforming in the representation 16 of $SO(16)$. In the quantum mechanics of these particles, quantization of the fermion zero modes implies these particles transform in the spinor representation; there is a $\mathbf{Z_2}$ gauge symmetry on the D0-brane volume that forces us to project out on of the chiral representations.

B.8 Final remarks

As promised in the introduction, the study of strong coupling behaviour of string theory has enriched our picture of these theories, and shown they are all related in an intricate web of dualities, see figure B.2. The duality web get even more intricate as we compactify in more involveld geometries.

We have learnt the lesson that different string theories are simply different perturbative limits of a unique underlying theory. This underlying theory has moreover a limit described by an 11d theory, which reduces at low energies to 11d supergravity.

The theory underlying all string theories and the 11d theory is sometimes referred to as M-theory as well, in a broad sense (M-theory is often used in a restricted sense to refer to the 11d theory underlying 11d supergravity).

There are several proposals to define M-theory (in a broad sense) microscopically, but for the moment a complete definition is lacking: We do not have a complete definition of string theory beyond the perturbative corners.

Altough the discoveries in this lecture may make us feel a bit uncomfortable, we should realize that the final picture is extremely beautiful. For instance, in perturbation theory it seemed that we had five different and seemingly disconnected solutions/proposals to provide a quantum consistent description of gravity and gauge interactions. Non-perturbatively we find that in fact there is a unique answer to this problem. The issue is to extract the fundamental physical principles underlying this theory in an intrinsic way (not tied to any particular perturbative limit).

Appendix C

Non-perturbative effects in (weakly coupled) string theory

C.1 Motivation

We have seen that non-perturbative states are very important in the structure of string theory at finite string coupling. In this lecture we will discuss that non-perturbative states are also essential even in the weakly coupled regime in certain situations, in which the purely perturbative sector of the theory is incomplete and leads to divergent answers for physical quantities.

There are different situations of this kind in string theory. In this lecture we center on two particular examples: enhanced gauge symmetries in type IIA/M-theory on K3, and conifold singularities in Calabi-Yau compactifications.

C.2 Enhanced gauge symmetries in type IIA theory on K3

C.2.1 K3

K3 is the only compact topological space with four dimensions admitting a Calabi-Yau metric, i.e. of $SU(2)$ holonomy (besides the four-torus $\mathbf{T^4}$, which has trivial holonomy). We now state without proof some of its properties, see [84] for a more extensive discussion.

Its Hodge numbers are

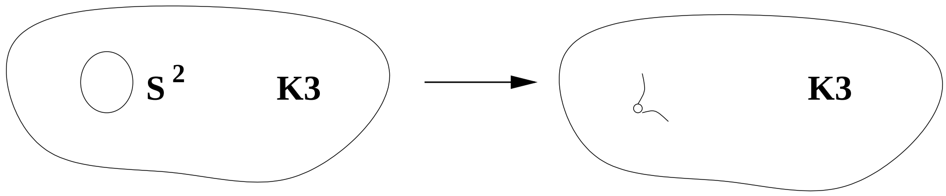


Figure C.1: In K3, singularities arise when some 2-cycles are tuned to have zero size.

$$
\begin{array}{ccccc}
 & & h_{0,0} & & \\
 & h_{1,0} & & h_{0,1} & \\
h_{2,0} & & h_{1,1} & & h_{0,2} \\
 & h_{2,1} & & h_{1,2} & \\
 & & h_{2,2} & &
\end{array}
\qquad\qquad
\begin{array}{ccccc}
 & & 1 & & \\
 & 0 & & 0 & \\
1 & & 20 & & 1 \\
 & 0 & & 0 & \\
 & & 0 & &
\end{array}
$$

The lattice of homology classes (with integer coefficients) turns out to be even and self-dual. We can split the corresponding harmonic forms in self-dual and anti-self dual forms, with respect to the 4d metric. This introduces a signature in the homology lattice, with 20 self-dual forms (given by 19 of the $(1,1)$ forms and a linear combination of the $(0,0)$ and the $(2,2)$ forms) and 4 anti-self-dual forms (one $(1,1)$ form and a combination of the $(0,0)$ and $(2,2)$). Hence, the (integer) homology of K3 has the very suggestive form of a even self-dual lattice with a lorentzian $(20,4)$ signature.

The moduli space of Calabi-Yau metrics on K3 is 58-dimensional. There are 38 parameters specifying the complex structure on K3 (i.e, telling us how to cook up comlex coordinates starting from real ones), and 20 parameters specifying the Kahler class.

At particular points (or more precisely, at some locus) in this metric moduli space, K3 develops singularities, which are always of orbifold type [1] $\mathbf{C^2/\Gamma}$, with Γ a discrete subgroup of $SU(2)$. This limits correspond to points in moduli space where some 2-cycles within K3 have been tuned to zero size, see figure C.1. The simplest such situation is $\mathbf{C^2/Z_2}$, where just one 2-cycle collapses to zero size.

Notice that tuning more parameters, one can go to a limit where the

[1] That is, the only singular local geometries that are consistent with $SU(2)$ holonomy are of orbifold type. In three complex dimensions there exist singularities consistent with $SU(3)$ holonomy, which are not of orbifold type.

whole K3 has the form of a toroidal orbifold, of the kind studied in the lecture on orbifold compactification. For instance, there exist points in the moduli space of metrics in K3 where it is of the forms $\mathbf{T^2}/\mathbf{Z_2}$, At each of the 16 fixed points of the orbifolds the local geometry is $\mathbf{C^2}/\mathbf{Z_2}$ and there is a zero size 2-cycle.

C.2.2 Type IIA on K3

We are interested in studying compactification of type IIA theory on K3. Since K3 has $SU(2)$ holonomy, each 10d gravitino leads to one 6d gravitino. The resulting 6d theory has therefore 16 unbroken supercharges and (being non-chiral) corresponds to 6d $\mathcal{N} = (1,1)$ supersymmetry. The main massless supermultiplets are

• the gravity multiplet, containing the graviton $G_{\mu\nu}$, a 2-form $B_{\mu\nu}$, a real scalar ϕ, four gauge bosons A_μ, two gravitinos $\psi_{\mu\alpha}$, $\psi_{\mu\dot{\alpha}}$, and two Weyl fermions ψ_α, $\psi_{\dot{\alpha}}$, all of opposite chiralities.

• the vector multiplet, with one gauge boson A_μ, four real scalars, and two Weyl fermions of opposite chiralities.

As usual, it will be thus enough to identify the bosonic fields in the 6d theory, since the fermions simply complete the supermultiplets.

Since K3 is curved (unless we are sitting at the point of moduli space corresponding to some global orbifold geometry) the 2d worldsheet theory is not free, and we can discuss compactification only in the supergravity approximation. This will provide the spectrum in the limit where all length scales in K3 are large (in particular all 2-cycles are large), usually refered to as large volume regime. Denoting Σ_a the 22 $(2,2)$ 2-cycles, Π, $\overline{\Pi}$ the $(2,0)$ and $(0,2)$ 2-cycles, and Π_a the 20 $(1,1)$ 2-cycles, the Kaluza-Klein reduction of the massless 10d bosonic fields gives

IIA		Gravity	Vector
G	$\rightarrow$	$G_{\mu\nu}$	38+20 scalars
B	$\rightarrow$	B_2	$\int_{\Sigma_a} B$
ϕ	$\rightarrow$	ϕ	
A_1	$\rightarrow$	A_1	
C_3	$\rightarrow$	C_3, $\int_\Pi C_3$, $\int_{\overline{\Pi}} C_3$	$\int_{\Pi_a} C_3$

We thus obtain the 6d $\mathcal{N} = (1,1)$ supergravity multiplet and 20 vector multiplets (with gauge group $U(1)^{20}$).

The structure of the moduli space is (locally) of the form

$$\frac{SO(20,4)}{SO(20)\times SO(4)} \tag{C.1}$$

In principle this can be determined from supergravity, in the large K3 volume regime. However it turns out to be completely determined by supersymmetry, so it is exactly of this form (locally), with no α' or g_s corrections. The above structure is related as we know to the moduli space of 24-dimensional $(20,4)$ lorentzian even self-dual lattices up to rotations within the 20d and 4d signature eigenspaces. In K3, it can be regarded as the moduli space of ways of splitting the 24d lattice of homology classes into sublattices of self-dual and anti-self-dual forms. More technical considerations involving mirror symmetry moreover allows to determine the global structure of moduli space of IIA on K3 [85], which turns out to be

$$\frac{SO(20,4)}{SO(20)\times SO(4)\times SO(20,4;\mathbf{Z})}. \tag{C.2}$$

C.2.3 Heterotic on $\mathbf{T^4}$ / Type IIA on K3 duality

This is a prototypical example of string duality below ten dimensions. Let us provide a list of supporting evidence for it; for details, see [86, 104].

• The spectrum of heterotic string theory on $\mathbf{T^4}$ (either for the $E_8\times E_8$ or the $SO(32)$ theories, since they are equivalent upon toroidal compactification) , at a generic point of its moduli space (see lecture on toroidal compactification of superstrings) is given by the 6d $\mathcal{N}=(1,1)$ supergravity multiplet and 20 vector multiplets (with gauge group $U(1)^{20}$). The bosonic fields arise from $G_{\mu\nu}$, $B_{\mu\nu}$, ϕ, the 24 abelian gauge bosons $G_{m\mu}$, $B_{m\mu}$, A_μ^I and the 80 scalars G_{mn}, B_{mn}, A_m^I, with $m=1,\ldots,4$, $I=1,\ldots,16$.

• The structure of the moduli space of both theories agrees, even globally. As we know, $\mathbf{T^4}$ compactifications of heterotic string theory have (C.2) as their moduli space (with the lattice corresponding to the Narain lattice of left- and right-moving momenta).

• The low-energy effective actions of bothe theories is the same, up to a redefinition of the fields. Defining the 6d dilaton by $e^{-2\phi_6}=V_{X_4}e^{-2\phi}$, with V_{X_4} the volume of the internal space, the actions agree up to the field redefinition

$$\phi_6' \quad = \phi_6 \quad ; \quad H_3 \quad = e^{-2\phi_6} *_6 H_3$$

$$G' \; = e^{-2\phi_6}\, G \quad ; \quad A_a^{I\,\prime} \; = A_a^I \tag{C.3}$$

The relations work as above in any direction of the duality. The above mapping implies that when the IIA theory has large 6d coupling, it admits a dual perturbative description in terms of weakly coupled heterotic strings, and vice versa.

- The spectrum of BPS states agrees in both theories. For instance,

Heterotic on $\mathbf{T^4}$		Type IIA on K3
F1	$\longleftrightarrow$	NS5 wrapped on K3
NS5 wrapped on $\mathbf{T^4}$	$\longleftrightarrow$	F1
momentum k_i		D2 wrapped on any
winding w_i	$\longleftrightarrow$	of the 22 2-cycles or
momentum P_I		D0, or D4 wrapped on K3

The tensions of these objects agree, and objects related as above have equivalent world-volume field theories.

The fundamental string of one theory corresponds to the wrapped five-brane of the other. Namely starting with the IIA theory and going to the limit of large 6d coupling the wrapped fivebrane becomes weakly coupled and sets the lightest scale, hence dominating the dynamics. In fact, it is possible to see that the world-volume theory on this wrapped fivebrane is that of a heterotic string (and viceversa of the IIA F1 vs the heterotic NS5).

C.2.4 Enhanced non-abelian gauge symmetry

The above duality suggests that there must exist an interesting phenomenon at particular points (loci) in the moduli space of type IIA on K3. Indeed, at particular points (or rather, subspaces) of the moduli space of heterotic theory on $\mathbf{T^4}$, some abelian gauge symmetries get enhanced to non-abelian ones. Recalling the left-moving spacetime mass formula

$$\alpha' M_L^2/2 = N_B + \frac{P_L^2}{2} - 1 \tag{C.4}$$

we see that when the parameters are tuned such that some state has $P_L^2 = 2$, we get two new massless state, corresponding to $\pm P_L$. They corresponds to a 6d vector multiplet, and carry charges ± 1 under some linear combination of the $U(1)$ gauge factors in the generic gauge group. Thus, they enhance the corresponding $U(1)$ gauge group to $SU(2)$.

This process has clear generalizations. If the parameters are tuned in such a way that additional states reach $P_L^2 = 2$, then we obtaine enhancements to larger gauge factors. In general, any non-abelian gauge symmetry with Lie algebra of type A, D or E (or products thereof) and rank ≤ 24 is possible (Note that only these algebras are possible since they are the only ones with all roots of length square equal to 2).

The states becoming massless are BPS states, so we know that there are new massless states in heterotic theory, even at strong coupling. By duality, this implies that type IIA must have enhanced non-abelian gauge symmetries at particular points in K3 moduli space, even at weak coupling. This is a very surprising conclusion: we have seem that compactification of type IIA theory on large and smooth K3 spaces leads to abelian gauge symmetries. Moreover one can use 2d conformal field theory tecniques to show (exactly in α') that any regular conformal field theory describing propagation of IIA string theory on K3 necessarily leads only to abelian gauge symmetries.

Interestingly enough, it is possible to show that there are points in moduli space of K3 where the 2d conformal field theory breaks down, i.e. the perturbative prescription to compute things in string theory gives infinite answers. Hence we suspect that it is at these points in moduli space where non-abelian gauge symmetries may arise, due to non-perturbative effects (present even at weak coupling!). These points in moduli space correspond to K3 geometries where some 2-cycle is collapsed to zero size and where the integral of B along the 2-cycle vanishes. The simplest situation corresponds to geometries with one collapsed 2-cycle C on which $\int_C B = 0$. As discussed above, this corresponds to the geometry of a local $\mathbf{C^2}/\mathbf{Z_2}$ orbifold singularity.

Now it is easy to identify how gauge symmetry enhancement occurs. The 6d theory contains a $U(1)$ gauge boson arising from $\int_C C_3$. The theory contains 6d particle states charged under it with charges ± 1, arising from D2-branes wrapped on C (with the two possible orientations). It is possible to see that these states are BPS [2], and their mass is (exactly) given by

$$M = \frac{|V_C + ib|}{g_s} \tag{C.5}$$

[2]Understanding this requires some discussion of the supersymmetry unbroken by D-branes wrapped on cycles in Calabi-Yau space. We chose to skip this discussion for our introductory overview.

where V_C denotes the volume of C and $b = \int_C B$. Hence, at the point or zero size and zero B-field the D2-brane particle is exactly massless, no matter how small the string coupling is. This effect is very surprising, since we see that the non-perturbative sector of the theory leads to significant effects (new massless particles!) even in the weak coupling regime. It is reasonable (and correct) to guess that these new particles in 6d belong to vector multiplets $\mathcal{N} = (1,1)$ supersymmetry, and therefore enhance the gauge symmetry from $U(1)$ to $SU(2)$. Clearly, these D2-brane particles are the duals to the $P_L^2 = 2$ states in heterotic theory. Note that this is in agreement with the mapping of BPS states proposed above.

Several comments are in order

• Notice that in the IIA picture we have perturbative states (the $U(1)$ gauge boson) and non-perturbative ones (the D2-brane particles) on an equal footing. Indeed, they are related by an exact gauge symmetry of the theory.

• In the above discussion we used heterotic/IIA duality to motivate the appearance of enhanced gauge symmetries in IIA compactifications on K3. However, the whole argument about the appearance of new massless charged states could have been done based simply on our knowledge of D-branes and the BPS formulae, without any use of string duality. Clearly, we have enough understanding of non-perturbative states in string theory to look for them without help from duality, as we will do in next section.

• Once the additional multiplet of non-perturbative origin is included, the physics of the configuration is completely non-singular. Equivalently, the divergent behaviour of the perturbative sector can be understood as due to incorrectly not including all the massless fields in the dynamics (as often stated, due to integrating out (= to not incluing) the non-perturbative state, incorrectly since it is a massless state that clearly must be included in discussing the low energy dynamics of the system).

• Let us emphasize again that this non-perturbative effect takes place no matter how small the string coupling is.

• The point $V_C = 0$, $b = 0$ is singular from the viewpoint of the 2d world-sheet theory, which only sees perturbative physics. This may seem puzzling at first sight: In the lesson on orbifold compactification we studied orbifold singularities with cycles collapsed to zero size, and they were perfectly well described by simple (in fact, free) 2d worldsheet theories. The key difference, realized in [88], is that the orbifold describe by a free 2d worldsheet theory corresponds to a point in moduli space where $V_C = 0$ but $b \neq 0$ (in fact $b = 1/2$ for $\mathbf{C^2}/\mathbf{Z_2}$). In this situation, the D2 particle is very massive at weak coupling, and the perturbative description is accurate and non-singular (gives finite answers for all observables in the theory).

• There is a generalized version of this that explains other non-abelian gauge symmetry enhancements. There is a classification of $\mathbf{C^2\Gamma}$ singularities with Γ a discrete subgroup of $SU(2)$. In this classification there is an infinite A series (corresponding to cyclic $\mathbf{Z_k}$ groups), and infinite D series (dihedral groups) and an E series with three groups (denoted E_6, E_7, E_8). When parameters of K3 are tuned so that it develops a $\mathbf{C^2/\Gamma}$ singularity of A, D, E type (with zero B-fields over the collapsed 2-cycles), non-perturbative states become massless and enhance the gauge symmetry to the corresponding A, D, E gauge group. This provides the IIA dual to the configurations of enhanced gauge symmetries in heterotic compactifications. Moreover, it also establishes a 'physics proof' of the so-called McKay correspondence in mathematics, which establishes a relation between the geometry of orbifold singularities $\mathbf{C^2/\Gamma}$ and Lie algebras.

C.2.5 Further comments

It is interesting to consider dual realizations of this gauge symmetry enhancement. Indeed, we will find out that it is related to a very familiar phenomenon we have already encountered.

The local geometry of $\mathbf{C^2/Z_2}$ is identical to that of a 2-center Taub-NUT geometry in the limit where the two centers coincide. In fact, it is possible to display the 2-cycle collapsing to zero size in quite an explicit way, see figure C.2a. Both spaces differ only in their asymptotic behaviour at infinity, but this is not important for the phenomenon of gauge symmetry enhancement. Therefore, we conclude that multi - Taub-NUT spaces develop enhanced gauge symmetry when two centers coincide, and the B-field is tuned to zero.

Performing now a T-duality along the isometric direction in the Taub-NUT space, the two centers of the Taub-NUT geometry turn into two parallel NS5-branes of IIB theory, sitting at points in the transverse $\mathbf{R^4}$. Their separation in $\mathbf{R^3}$ is determined by the volume and B-field of the T-dual 2-cycle. The non-perturbative D2-brane state now corresponds to a D1-brane stretched between the NS5-branes, which clearly becomes massless when the NS5-branes coincide. Performing now an S-duality on this configuration we obtain two D5-branes; the state related to the original D2-brane is now a fundametnal string stretched between the D5-branes. In this language, the exotic phenomenon of enhanced symmetry due to the D2-brane state is the

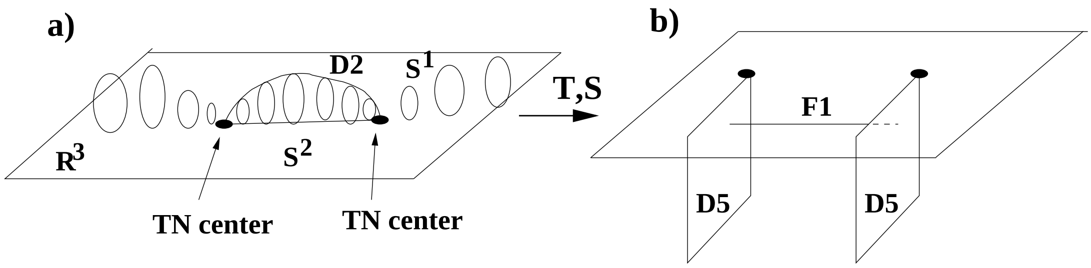


Figure C.2: The $\mathbf{S^1}$ fibration over a segment joining two centers in a multi Taub-NUT geometry defines a homologically non-trivial 2-cycle with the topology of a 2-sphere. Its volume vanishes as the two centers of the Taub-NUT are tuned to coincide.

familiar phenomenon of enhancement of 6d gauge symmetry on the volume of D5-branes when they are coincident, due to the appearance of new massless open (fundamental) strings, see fig C.2b.

We would like to conclude by briefly mentioning that compactification of type IIB theory on K3 leads to even more exotic physics [89]. Type IIB theory does not contain abelian $U(1)$ gauge symmetries associated to 2-cycles. Rather it contains abelian 2-forms, arising from the KK reduction of C_4, belonging to tensor multiplets of 6d $\mathcal{N} = (2,0)$ supersymmetry. Similarly, IIB theory does not have D2-brane states and hence does not lead to new massless particles in K3 with collapsed 2-cycles and zero B-field. Instead it leads to BPS tensionless string states, charged under the 2-form fields, arising from D3-branes wrapped on the collapsed 2-cycles. This surprising answer is completely consistent with T-duality with the type IIA answer, once we compactify both IIB and IIA on a further circle. Winding states of these IIB tensionless strings are mapped by T-duality to momentum states of the IIA massless particles.

These configurations can be used to define exotic theories in 6d if we take the limit of decoupling gravitational interactions. In particular, they can be used to define the so-called $(0,2)$ superconformal field theoty, or the so-called little string theory. Their discussion is however beyond our scope in these lectures.

C.3 Type IIB on CY_3 and conifold singularities

We now have enough understanding of BPS states in string theory to analyze non-perturbative effects in other situations, even without the help from string duality. For this section see [90].

C.3.1 Breakdown of the perturbative theory at points in moduli space

Recall that type IIB on Calabi-Yau threefolds, with Hodge numbers $(h_{1,1}, h_{2,1})$, gives rise to the $\mathcal{N} = 2$ 4d supergravity multiplet, $(h_{1,1} + 1)$ vector multiplets and $h_{2,1}$ hypermultiplets. The latter two kinds of multiplets contain scalars spanning a moduli space. We are interested in looking for regions in this moduli space where non-perturbative effects may be relevant, even at weak coupling.

There is a non-renormalization theorem for 4d $\mathcal{N} = 2$ supersymmetry that ensures that (to all orders in perturbation theory) the geometry of the moduli space of vector multiplets (the moduli space metric, which controls the kinetic terms of moduli in the effective action) does not depend on scalars in hypermultiplets, and vice versa. In type IIB, both the dilaton and the overall volume of the Calabi-Yau belong to hypermultiplets. This implies that the geometry of the vector multiplet space does not depend on the dilaton (i.e. does not suffer any quantum corrections in g_s) or on the volume scalar (i.e. does not suffer any α'/R^2 corrections). The moduli space metric determined in the classical supergravity approximation is exacty in g_s and α'.

On the other hand it is known that there are points in the moduli space of complex structures (i.e. vector multiplet moduli space) of Calabi-Yau manifolds where the effective action obtained from supergravity is singular. Since we have argued that the supergravity result is exact, there is no α' or g_s correction (to any order in perturbation theory) which removes this singularity. This means that even the α'-exact worldsheet theory (describing compactification on the Calabi-Yau space at this point in complex structure moduli space) is singular, and gives divergent answers for certain physical quantities.

This breakdown of the perturbative prescription suggests that at this points in moduli space there is some non-perturbative effect playing an essential role, even if the string coupling is weak. Our aim in this section is to discuss this effect.

C.3.2 The conifold singularity

Let us discuss, the generic, simplest, case where compactification on a CY_3 leads to a breakdown of the perturbative theory. It corresponds to sitting at a point in complex structure moduli space, such that the CY_3 has a region which locally develops a so-called conifold singularity. Namely, a piece of the CY_3 can be locally described as the complex hypersurface in $\mathbf{C^4}$ given by the equation

$$(z_1)^2 + (z_2)^2 + (z_3)^2 + (z_4)^2 = \epsilon \tag{C.6}$$

The complex structure modulus is described by the parameter ϵ, and the problematic configuration corresponds to tuning $\epsilon \to 0$.

The above geometry corresponds, as $\epsilon \to 0$ to a local singularity, which is not an orbifold, but still is quite simple and well-knonw to mathematicians (algebraic geometers). It is possible to see that the geometry (C.6) contains a 3-cycle with the topology of a 3-sphere of size controlled by $|\epsilon|$. Namely, let $\epsilon = |\epsilon| e^{i\theta}$, and define $z_i' = z_i e^{-i\theta/2}$. If we let $x_i = \mathrm{Re}\, z_i'$, $y_i = \mathrm{Im}\, z_i'$, the 3-sphere is given by

$$y_i = 0 \quad , \quad (x_1)^2 + (x_2)^2 + (x_3)^2 + (x_4)^2 = |\epsilon| \tag{C.7}$$

As $\epsilon \to 0$ the 3-cycle C collapses to zero size (see figure C.3). In the configuration with a zero size 3-cycle, the perturbative theory breaks down.

The cure of the problem is now clear. Type IIB string theory on this CY_3 contains non-pertubative particle states arising from D3-branes wrapped on the 3-cycle C. It is possible to see that this state is BPS [3] and that its mass is given by

$$M = \frac{|\epsilon|}{g_s} \tag{C.8}$$

Thus it becomes massless precisely when $\epsilon \to 0$, suggesting that this states solves the problem of the perturbative sector, as is indeed the case.

[3]The 3-cycle has the property of being Special lagrangian, which implies that D-branes wrapped on it preserve some of the supersymmetry unbroken by the CY_3.

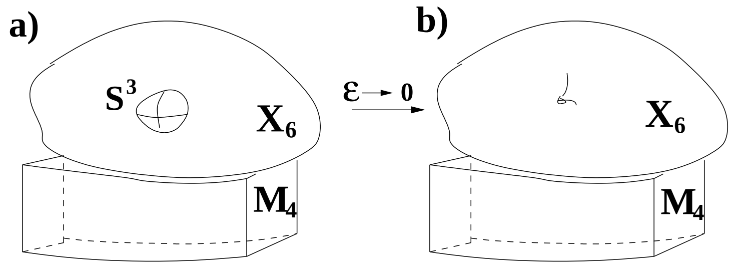


Figure C.3: Tuning a modulus in the Calabi-Yau geometry, a 3-cycle shrinks and the geometry develops a conifold singularity.

An important difference with respect to the case of IIA theory on K3, is that the massless states belong to hypermultiplets of $\mathcal{N}=2$ 4d supersymmetry. They are charged under the (perturbative) $U(1)$ gauge symmetry arising from $\int_C C_4$. Therefore the effective action for the light modes in this region in moduli space, is simply a $U(1)$ vector multiplet coupled to a charged hypermultiplet of mass equal to ϵ. In $\mathcal{N}=1$ susy language, we have a $U(1)$ vector multiplet V, a neutral chiral multiplet Φ (whose vev corresponds to ϵ) and two chiral multiplets of H, H' of charges ± 1. The action is of the form1

$$\mathcal{L} \;=\; \int d^2\theta W_\alpha W_\alpha + \int d^4\theta \,(H^\dagger e^V H - H'^\dagger e^V H') + \int d^2\theta \Phi H H' \quad \text{(C.9)}$$

This is a perfectly nice an smooth effective action. However, integrating out the massless fields H, H' leads to the singular behaviour of the perturbative sector. The pathological behaviour of the perturbative theory can be regarded as a consequence of missing important dynamical degrees of freedom in the low energy dynamics.

Again, let us emphasize that the appearance of these non-perturbative states takes place no matter how small the string coupling is.

C.3.3 Topology change

For this section see [91].

We have seen that the conifold geometry can be regarded as a limit of a smooth geometry (C.6), containing a 3-cycle, in the limit where the 3-cycle collapses to zero size. Mathematically, the conifold geometry can also be regarded as a limit of a (different) smooth geometry, containing a 2-cycle, in

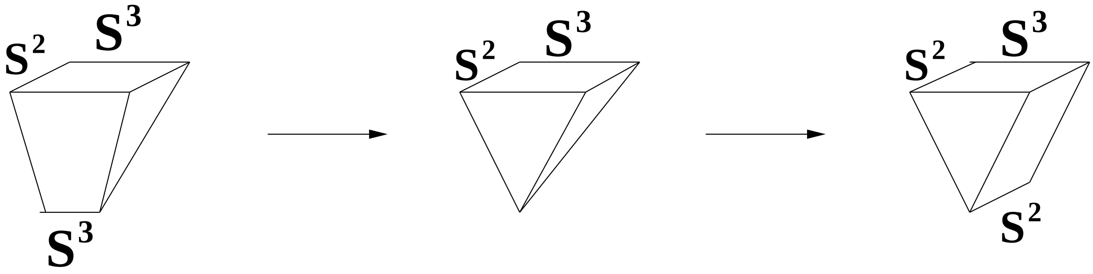


Figure C.4: Topology change in the neighbourhood of a conifold singularity. Starting with a finite size $\mathbf{S^3}$ we tune a modulus to shrink it; at this stage massless state appear; a vev for them parametrizes growing an $\mathbf{S^2}$ out of the conifold singualrity.

the limit where the 2-cycle collapses to zero size (and the B-fieldthrough it is tuned to zero).

To understand this better, consider the equation (C.6) for $\epsilon = 0$ in terms of $x_i = \mathrm{Re}\, z_i$, $y_i = \mathrm{Im}\, z_i$. We get

$$x^2 - y^2 = 0 \quad , \quad x \cdot y = 0 \tag{C.10}$$

where x, y are 4-vectors with components x_i, y_i. Equivalently, introducing a new variable r taking positive values, we have

$$x^2 = r^2 \quad ; \quad y^2 = r^2 \ , \ x \cdot y = 0 \tag{C.11}$$

The first equation implies that x describes a 3-sphere of radius r, while the last equations implies that y describes a 2-sphere of radius r. The geometry of the conifold is a cone, with base $\mathbf{S^3} \times \mathbf{S^2}$ and radial coordinate r. At $r = 0$ both the 3-sphere and the 2-sphere have zero size.

The manifold (C.6) for non-zero ϵ corresponds to a smoothing of the conifold singularity by replacing the singular tip of the cone by a finite size 3-sphere, as illustrated in C.4. This process is called deformation of the singularity. As mentioned above, there is also the possibility of smoothing the geometry by replacing the singular tip of the cone by a finite size 2-sphere, as illustrated in figure C.4. This process is called small resolution of the singularity, and mathematically the smooth space is described by the equations

$$\begin{aligned} z_+ x + w_+ y &= 0 \\ w_- x + z_+ y &= 0 \end{aligned} \tag{C.12}$$

in $\mathbf{C^4} \times \mathbf{P_1}$, where $\mathbf{C^4}$ is parametrized by $z_\pm = z_1 \pm iz_2$, $w_\pm = i(z_3 \pm iz_4)$, and $\mathbf{P_1}$ is parametrized by (x, y) (with the equivalence relation $(x, y) \simeq \lambda(x, y)$ with $\lambda \in \mathbf{C}^*$. The above equations define a smooth space, which is the same as the conifold singularity except at the tip of the cone. Namely, for each non-zero value of $z_\pm$, $w_\pm$, the above equations define a unique point, so the resolved space has a 1-1 mapping to the conifold singularity (away from the thip). When $z_\pm = w_\pm = 0$, then (x, y) are unconstrained and instead of just a singular point we obtain a whole $\mathbf{P_1}$. The resolved conifold thus corresponds to a smooth space, containing a 2-sphere, given by the $\mathbf{P_1}$. When its size goes to zero, the space becomes the conifold singularity.

Starting with a deformed conifold, we can imagine the process of shrinking the 3-cycle to zero size to reach the singular conifold geometry, and then growing a 2-cycle to obtain a resolved conifold. This process changes the topology of the space, since we have $\Delta(h_{1,1}, h_{2,1}) = (1, -1)$. This process is possible mathematically, but only passing through singular geometries. However, we have just seen that physically string theory is smooth even at the singular geometry. Therefore it is reasonable to wonder whether string theory can smoothly interpolate between the two topologically different geometries.

It can be shown that this is not really possible in the above situation, where the CY_3 has only one conifold point. The new geometry does not contain any 3-cycle, hence the low energy theory should not have any $U(1)$ gauge symmetry. This suggests that the transition to the new geometry must be triggered by a vev for the massless charged hypermultiplet. However, the field theory (C.9) does not have a flat direction where the multiplets H, H' acquire non-zero vevs. This cannot be done due to the conditions to minimize the scalar potential: these include the D-flatness constraint for the $U(1)$ gauge symmetry

$$|H|^2 - |H'|^2 = 0 \qquad \text{(C.13)}$$

and the F-flatness constraint

$$\frac{\partial W}{\partial \Phi} = H\, H' = 0 \qquad \text{(C.14)}$$

In other words, since in the Higgsing of $U(1)$ the vector multiplet must eat one hypermultiplet, we are left with not scalars whose vev parametrize the new branch.

On the other hand, this kind of topology changing transitions are possible at points in complex structure moduli space where the CY_3 develops several conifold singularities, such that the 3-cycles at the conifold points are not homologically independent. For instance, we can imagine a CY_3 with N conifold singularities, with the property that the homology classes of the corresponding 3-cycles add to zero in homology. In such situation the gauge symmetry is $U(1)^{N-1}$; equivalently there are N gauge bosons $U(1)^N$, but there is a relation between them, namely their sum is identically zero. On the other hand, we still get N independent charged hypermultiplets arising from D3-branes wrapped on the N collapsing 3-spheres. So in $\mathcal{N}=1$ multiplet language we have N pairs of chiral multiplets H_i, H_i' with charges $\pm q_a^i$ under the a^{th} $U(1)$ factor, with $a=1,\ldots,N$ and $\sum_a q_a^i=0$.

The effective theory for these field does have a flat direction where the fields H_i, H_i' acquire vevs, as can be checked from the D- and F-term constraints in this case

$$\begin{array}{c} \sum_i q_a^i(|H_i|^2-|H_i'|^2)=0 \quad , \quad a=1,\ldots,N \\ \sum_i q_a^i H_i H_i'=0 \end{array} \tag{C.15}$$

And there is a flat direction, corresponding to $\langle H_i'\rangle=v$, $\langle H_i\rangle=w$, i.e. i-independent vevs. More intuitively, we have N charged hypermultiplets to Higgs $U(1)^{N-1}$ vector multiplets. Clearly $N-1$ hypermultiplets are eaten in the Higgs mechanism, making the vector multiplets massive, while a last hypermultiplet is left. The two complex parameters v,w correspond to vevs for scalars in this overall hypermultiplet.

The geometric interpretation of this new branch is clear. Since there are no massless $U(1)$'s, all 3-spheres have disappeared from the geometry. Since there is a new massless hypermultiplet, there is a new 2-sphere. Indeed, there are N new 2-spheres at the N conifold points, which have been resolved, but the geometry forces the sizes of all these spheres to be equal [4]. String theory has managed to smoothly interpolate [5] between the two topologically different geometries, thanks to the crucial presence of massless non-perturbative states! (figure C.5).

Some comments are in order

[4]It would be a bit tricky to explain this, see [92].

[5]Notice that the topology change as we have discussed it is not really dynamical, but simpy and adiabatic change as some parameters of the model are varied. However, it is easy to imagine configurations where moduli change slowly with time, so that their vevs evolve in time, and we are really moving in moduli space as time goes by. In this setup the above topology change could occur dynamically during time evolution.

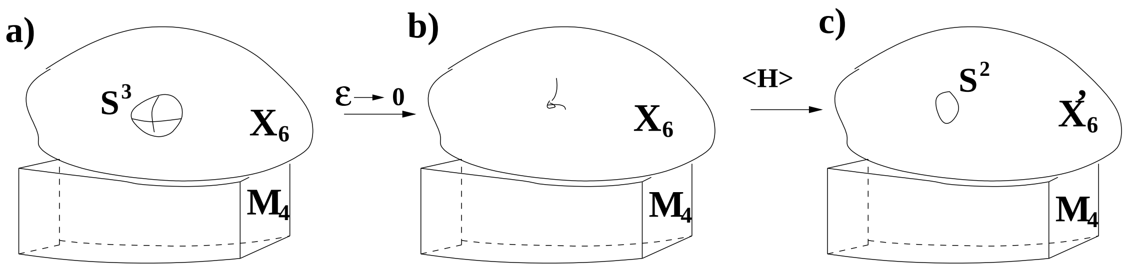


Figure C.5: Topology change in CY spaces with conifold singularities.

• Let us emphasize again that, at least in this particular setup, string theory is able to interpolate smoothly between spacetimes of different topologies. In a sense, this is a more drastic version of the statement that geometry is dynamical in theories with gravity. In string theory, even the topology of spacetime is, to some extent, dynamical and can change.

• After the transition to the small resolution branch, the original hypermultiplet which was of non-perturbative origin, becomes just a perturbative hypermultiplet arising from the KK reduction of 10d type IIB theory on a CY_2 with a 2-cycle. This is a very striking phenomenon, but certainly it is implied by our discussion of topology change.

• The topology changing transitions allow to connect the moduli spaces of different CY compactifications. Indeed it has been checked that all known Calabi-Yau manifolds are connected by this kind of transition (or generalizations of it). This is conceptually very satisfying, and suggests that the election of particular compactification is as dynamical as the choice of vevs for some fields in a(n extended) moduli space.

• Finally, we would like to point out that there exist dual versions of this phenomenon, where it looks much more familiar. For instance, there exists a dual version in terms of heterotic theory compactified on $\mathrm{K3}\times\mathbf{T^2}$, where the above process corresponds to simply deforming the internal gauge bundle of the compactification.

C.4 Final comments

There are two finals comments we would like to make

• Non-perturbative effects can be important in string theory even in the weakly coupled regime. These effects are particularly crucial in situations where the perturbative sector of the theory is singular.

• The ideas in this lecture suggest a powerful tool to determine new interesting phenomena in string theory (and check its self-consistency). Namely, cook up situations where some singular behaviour arises, and try to identify what effects solve the problem Many new phenomena of string theory have been uncovered using this idea, and many more lessons still wait to be learnt.

Appendix D

D-branes and gauge field theories

D.1 Motivation

String theory in the presence of D-branes contains sectors of gauge interactions (open string sectors). The strength of gauge and gravitational interactions in these setups is different [1], making it possible to switch off gravitational (and other closed string) interactions, while keeping the gauge sectors interacting. This can be done essentially by taking a low energy limit in the configuration. In the limit, the dynamics of the open string sector of the theory reduces to a gauge field theory.

Hence, string theory is able to reproduce the richness of gauge field theory. The idea is to use string theory to explore the dynamics of gauge field theories; for instance, study non-perturbative effects in gauge theories by exploiting what we already know about the non-perturbative dynamics in string theory. In order to do so, we must center on theories with enough supersymmetry. In this talk we center on theories with 16 supersymmetries, and four-dimensional gauge sectors (i.e. we center on configurations of parallel D3-brane in type IIB string theory).

[1] This is not true in heterotic models, for instance.

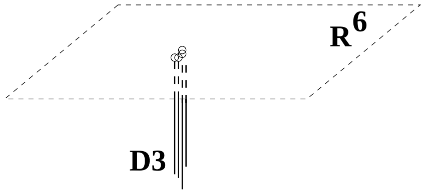


Figure D.1: Stack of coincident D3-branes in flat space..

D.2 D3-branes and 4d $\mathcal{N} = 1$ $U(N)$ super Yang-Mills

D.2.1 The configuration

Consider a stack of N coincident type IIB D3-branes in flat 10d space, see figure D.1. The open string spectrum contains massless modes corresponding to 4d $\mathcal{N} = 4$ $U(N)$ super Yang-Mills, propagating on the 4d D3-brane world-volume. This includes $U(N)$ gauge bosons, four Majorana fermions λ_r, $r = 1, \ldots, 4$ in the adjoint representation, and six real scalars ϕ_m, $m = 1, \ldots, 6$ in the adjoint representation. The configuration also contains massive open string modes, and massless and massive closed string modes.

Let us consider the limit of very low energies, or equivalentely of very large string scale (i.e. we take the limit $E/M_s \to 0$). In this limit, all massive string modes (open or closed) decouple, and moreover all interactions of massless closed string modes (which are controlled by 10d Newton's constant $\simeq M_s^{-8}$ go to zero. Interactions for massless open string modes, however, are controlled by the dilaton vev g_s and remain non-trivial. The whole dynamics of the configuration reduces to $\mathcal{N} = 4$ $U(N)$ gauge field theory. In $\mathcal{N} = 1$ supermultiplet language we have a vector multiplet V and three chiral multiplets Φ_i, $i = 1, 2, 3$, with action

$$S_{YM} = \int d^4x \left[\int d^2\theta \tau \mathrm{tr}\,(W_\alpha W_\alpha) + \int d^4\theta \sum_i \mathrm{tr}\,\Phi_i e^V \Phi_i^\dagger + \int d^2\theta \mathrm{tr}\,(\epsilon_{ijk}\Phi_i\Phi_j\Phi_k) \right]$$

where W_α is the field strength chiral multiplet, and $\tau = \theta + i/g_{YM}^2 = a + i/g_s$. This leads to the kinetic terms for gauge bosons, matter fields, and to the scalar potential proportional to square of the modulus of the commutators of scalar fields. For most of the discussion we center on $\theta = 0$, $a = 0$.

D.2.2 The dictionary

We can now establish a dictionary between properties of the gauge field theory and properties of the D3-brane configuration in string theory. A first example already described is that the complex gauge coupling constant corresponds in the underlying string theory to the complex IIB coupling constant. Also, for instance, the $SU(4) = SO(6)$ R-symmetry of $\mathcal{N} = 4$ super Yang-Mills theory, which acts on the four 4d spinor supercharges, corresponds in the underlying string theory to the $SO(6)$ rotational symmetry in the $\mathbf{R^6}$ transverse to the D3-branes.

The dictionary become particularly interesting in discussing the so-called Coulomb branch. The $\mathcal{N} = 4$ $U(N)$ gauge theory has a moduli space of vacua, parametrized by the vevs of the scalar fields. For these vevs to minimize the scalar potential, the vevs for the real scalar fields $\langle\phi_m\rangle$ should be $N \times N$ commuting matrices. Then they can be diagonalized simultaneously, with real eigenvalues $v_{m,a}$, $a = 1, \ldots, N$, namely

$$\langle\phi_m\rangle = \begin{pmatrix} v_{m,1} & & \\ & \ldots & \\ & & v_{m,N} \end{pmatrix} \tag{D.2}$$

The gauge symmetry in this vacuum is broken to $U(1)^N$, if the vevs are generic. In all cases, $\mathcal{N} = 4$ supersymmetry is unbroken in these vacua, so we have full $\mathcal{N} = 4$ vector multiplets of $U(1)^N$. In fact, each $U(1)$ gauge boson (refered to as the a^{th} $U(1)$) is associated with six massless scalars, which correspond to the moduli associated to the a^{th} set of vevs $v_{m,a}$.

States of the theory with electric charges $+1$, -1 under the a^{th} and b^{th} $U(1)$'s acquire a mass

$$M_{ab} = g_{YM}|\vec{v}_a - \vec{v}_b| \tag{D.3}$$

where $\vec{v}_a$ is a 6d vector with components $v_{m,a}$. This arises from the Higgs mechanism for gauge bosons, from the scalar potential for scalars, and from Yukawa couplings for fermions.

There are enhanced $U(n)$ gauge symmetries in the non-generic case when n of the $\vec{v}_a$ are equal. That is, the corresponding charged vector multiplets become massless.

In the underlying string picture, the moduli space of vacua corresponds to the moduli space of D3-branes. There exists a continuous set of configurations, corresponding to the choice of locations of the D3-branes in the transverse space $\mathbf{R^6}$ [2]. Labelling the D3-branes by a (Chan-Paton) index $a = 1, \ldots, N$, the configurations are described by the locations $r_{m,a}$ of the a^{th} D3-branes in the coordinate x^m in $\mathbf{R^6}$. All these configurations are $\mathcal{N} = 4$ supersymmetric. As will become clear in a moment, they correspond precisely the vacua of the $\mathcal{N} = 4$ gauge theory, via the relation $v_{i,a} = r_{i,a}/\alpha'$.

In this configuration, the gauge symmetry on the D3-branes is broken generically to $U(1)^N$, since only aa open strings are massless. If n D3-branes are located at coindicent positions in $\mathbf{R^6}$, their gauge symmetry is enhanced to $U(n)$. Furthermore, states with charges $+1$, -1 under the a^{th} and b^{th} $U(1)$'s correspond to ab open strings, see figure D.2. Their lightest modes have a mass

$$\alpha' M_{ab}^2 = \frac{|\vec{r}_a - \vec{r}_b|^2}{\alpha'} \tag{D.4}$$

in the string frame. Going to the Einstein frame, there is a rescaling of energies by by $\sqrt{g_s}$, so we get

$$M_{ab} = g_s^{1/2} \frac{|\vec{r}_a - \vec{r}_b|}{\alpha'} \tag{D.5}$$

which is in precise agreement with (D.3). That is, we can match the spectrum of electrically charges states in the gauge theory from the set of fundamental open strings stretched between the D3-branes.

As we discussed in the $\mathcal{N} = 4$ field theory appendix on the lecture of non-perturbative states, there are other BPS states in the $\mathcal{N} = 4$ gauge theory. In particular, each $SU(2)$ subgroup of the $U(N)$ is spontaneously broken to $U(1)$ in the Coulomb branch. Within each $SU(2) \to U(1)$ sector one can construct non-perturbative 't Hooft-Polyakov monopole states, carrying a magnetic charge under the corresponding $U(1)$, and mass proportional to the gauge symmetry breaking vev. More precisely, for the pair given by the a^{th} and b^{th} $U(1)$'s we have monopole states with charges $\pm(1, -1)$ under $U(a)_a \times U(1)_b$, and mass given by

$$M_{ab} = \frac{|\vec{v}_a - \vec{v}_b|}{g_{YM}} \tag{D.6}$$

[2] In order to allow for the possibility of branes separated in transverse space without decoupling them from each other in the low energy limit discussed above, the limit should also rescale the distance in transverse space $\mathbf{R^6}$. This will be discussed more carefully in section 2.

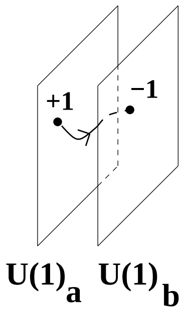


Figure D.2: In a system of parallel D3-branes, there are BPS states obtained as minimal length 1-branes (fundamental strings or otherwise) suspended between the a^{th} and the b^{th} D3-brane.

These states are BPS, etc, which guarantees that the abvove formula is exact quantum mechanically. Note that they are not charged under the diagonal $U(1)$, which is the extra $U(1)$ factor of $U(2)$ not in $SU(2)$, so we recover the $SU(2) \to U(1)$ monopole.

In the underlying string picture, there are states with these properties, corresponding to (open) D1-branes suspended between the a^{th} and b^{th} D3-branes. It is possible to see that they have the correct charges [3]. Their mass is given by its length times the D1-brane tension. In the Einstein frame,

$$M_{ab} = g_s^{1/2} \frac{|\vec{r}_a - \vec{r}_b|}{\alpha'^{1/2}} \frac{\alpha^{-1/2}}{g_s} = \frac{1}{g_s^{1/2}} \frac{|\vec{r}_a - \vec{r}_b|}{\alpha} \tag{D.7}$$

in agreement with (D.6).

Some comments are in order:

• It is possible to understand that the D1-brane states are supersymmetric, by analyzing the directions along which the D3- and D1-branes stretch. For instance, for D3-brane separated just along x^4, we have

[3]This would require describing the effect of the D1-brane pulling on the D3-brane. Such configurations are described by the so-called BIon solutions of the Dirac-Born-Infeld action on the D3-brane worldvolume: Intuitively, the D1-brane pulls the D3, so that the coordinate of the D3-brane vary as one moves away from the D1-brane endpoint. To keep energy finite, one must switch on the world-volume gauge field. The final configuration is such that there is net flux of F around a 2-sphere surrounding, at spatial infinity, the D1-brane endpoint, which thus corresponds to a magnetic monopole.

	0	1	2	3	4	5	6	7	8	9
D3	-	-	-	-	×	×	×	×	×	×
D3	-	×	×	×	-	×	×	×	×	×

The number of DN (Dirichlet Neumann) and ND directions is four, which corresponds to a supersymmetric situation (note that this would be a proof if the D1-branes were infinitely extended; since they are of finite extent, the argument is just heuristic, but gives the right answer).

• Open strings with endpoints on the D1-branes give rise to fields localized on the latter. The massless sector indeeed corresponds to the zero modes of the gauge theory monopole: bosons associated to the monopole position, and fermions due to the supersymmetries broken by the monopole state.

• In fact, the string theory configuration tells us that there are infinitely many BPS states associated to each pair of D3-branes, corresponding to (p, q) strings suspended between them. They must also exist in the gauge field theory, where they are known as dyons, which carry p and q units of electric and magnetic charge. They can be directly searched in the $\mathcal{N} = 4$ field theory, and have been constructed for particular values of (p, q). The masses of these states, (obtained in string theory language and transaled) is (for general τ)

$$M^2 = |\vec{v}_a - \vec{v}_b| \frac{1}{\Im \tau} |p + \tau q|^2 \tag{D.8}$$

• Note that in the limit of coincident D3-brane positions / coincident vevs (this is known as the origin in the Coulomb branch), the theory has massless electrically charged states, but also massless monopoles, and masless dyons. The $\mathcal{N} = 4$ theory at the origin in the Coulomb branch is highly non-trivial! It is only simple in perturbation theory, where all non-perturbative states are infinitely massive.

Finally, we would like to point out another small piece of the dictionary. The $\mathcal{N} = 4$ gauge theory has instantons, which are field configurations in the euclidean version of the theory. In the underlying string language, they are described by D(-1)-branes, which are D-branes localized in all directions in the euclidean version of the string theory. See figure D.3. This agrees with the familiar fact that a D$(p-4)$-brane on the volume of a Dp-brane behaves as an instanton (recall that an instanton carries D$(p-4)$-brane charge due to the WZ worldvolume coupling on the Dp-brane).

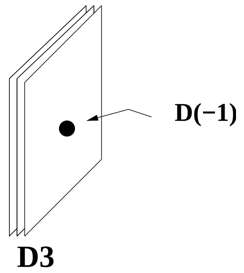


Figure D.3: A D(−1)-brane on a D3-brane corresponds to an instanton on the 4d gauge field theory on the D3-brane world-volume.

D.2.3 Montonen-Olive duality

As we discussed for $SU(2)$ in the $\mathcal{N} = 4$ field theory appendix on the lecture of non-perturbative states, there is a non-perturbative exact $SL(2, \mathbf{Z})$ symmetry of $\mathcal{N} = 4$ $U(N)$ super Yang-Mills theory, acting non-trivially on the complex gauge coupling, and exchanging the roles of perturbative and non-perturbative states.

This is easily derived from the underlying string picture. The D3-brane configuration is invariant under the exact non-perturbative $SL(2, \mathbf{Z})$ symmetry of type IIB theory, exchanging the roles of the different (p, q)-strings. This implies that the 4d $\mathcal{N} = 4$ gauge field theory inherits this as an exact symmetry, which exchanges the roles of the different electrically charged states, monopoles and dyons of the theory. Namely

Type IIB D3-branes	$\mathcal{N} = 4$ **gauge theory**
$\tau \to \frac{a\tau+b}{c\tau+d}$	$\tau \to \frac{a\tau+b}{c\tau+d}$
(p, q)-string $\to$ (p', q')-string	(p, q)-dyon $\to$ (p', q')-dyon

with $\begin{pmatrix} p' \\ q' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix}$.

As usual, this implies that e.g. the strong coupling dynamics of $\mathcal{N} = 4$ $U(N)$ super Yang-Mills is described by a dual weakly coupled $U(N)$ gauge theory where the perturbative degrees of freedom (electrically charged states in dual theory) are the original magnetic monopoles.

Montonen-Olive $SL(2, \mathbf{Z})$ duality arises from type IIB $SL(2, \mathbf{Z})$ duality in this setup.

D.2.4 Generalizations

There have been many generalizations of the possibility to study gauge theory phenomena by embedding them in the worldvolume of D-brane configurations in string theory. For a review see [93]. Some further examples and results one can show using string theory tools are

• Montonen-Olive dualities for $\mathcal{N} = 4$ gauge field theories with $SO(N)$ or $Sp(N)$ gauge groups, from configurations of D3-branes (and O3-planes) in type IIB string theory.

• For theories with 16, 8 supersymmetries in dimensions $d = 5, 6$, construction of interacting field theories which correspond to ultraviolet fixed points of the renormalization group (superconformal field theories).

• For 4d theories with 8 supersymmetries (4d $\mathcal{N} = 2$), exact computation of the low energy effective action (up to two derivatives) exactly in g_{YM} (including non-perturbative effects), in agreement with the Seiberg-Witten solution [94].

• A non-perturbative duality for 3d theories with 8 supersymmetries, known as mirror symmetry [4].

• For 4d theories with 4 supersymmetries (4d $\mathcal{N} = 1$), a non-perturbative equivalence in the infrared of theories which are different in the ultraviolet, knonw as Seiberg duality. Also, some qualitative features of $\mathcal{N} = 1$ pure Yang-Mills, like number of vacua, etc.

D.3 The Maldacena correspondence

In a sense, this is a more precise version of the relation between string theory and gauge field theory. It even allows the quantitative computation of quantities in gauge field theory from the string theory / supergravity point of view. An extensive review is [111].

D.3.1 Maldacena's argument

This section follows [96].

Consider a stack of N coincident type IIB D3-branes in flat 10d space [5] . The dynamics of the configuration is described by closed strings, and open

[4] The name is due to some relation with mirror symmetry in type II string theory.

[5] Similar arguments can be repreated for the M theory M2 and M5-branes. For other branes, non-trivial varying dilatons modify the argument substantially.

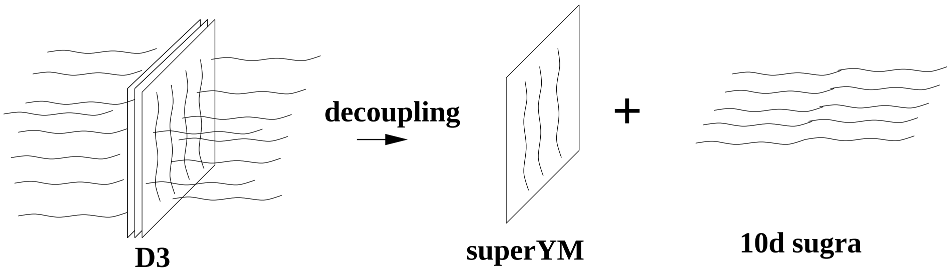


Figure D.4: Maldacena's low energy limit in the system of N coincident D3-branes, described as hypersurfaces on which open strings end. In the limit, we obtain $\mathcal{N}=4$ super Yang-Mills gauge field theory, decoupled from free 10d supergravity modes.

string ending on the D3-branes. Let us take a careful low energy limit, where we send the string scale to infinity, but keep the energies of 4d field theory excitations finite. One example of such states are open strings stretched between D3-branes separated by a distance r in transverse space (i.e. electrically charged states in the Coulomb branch), which have a mass

$$M^2 = r^2/\alpha'^2 \tag{D.9}$$

Hence we need to take the limit $\alpha' \to 0$ and $r \to 0$, keeping r/α' finite. In this limit, the theory reduces to two decoupled sectors, one of them is 4d $\mathcal{N}=4$ super Yang-Mills theory, and the other is free 10d gravitons (or better free fields corresponding to the massless closed string modes). See figure D.4.

On the other hand, the configuration has an equivalent description, as type IIB string theory in the background created by the stack of D3-branes

$$\begin{aligned} ds^2 &= Z(r)^{-1/2}\eta_{\mu\nu}dx^\mu dx^\nu + Z(r)^{1/2}dx^m dx^m \\ e^{2\phi} &= 1/g_s \\ F_5 &= (1+*)dtdx^1dx^2dx^3dZ^{-1} \end{aligned} \tag{D.10}$$

where $\mu = 0,\ldots,3$, $m = 4,\ldots,9$, and

$$r = \sum_m (x^m)^2 \quad , \quad Z(r) = 1 + \frac{R^4}{r^4} \quad , \quad R^4 = 4\pi g_s \alpha'^2 N \tag{D.11}$$

We have a gravitational background, pictorially shown in figure D.5, and a RR 5-form field strength background, such that there are N units of flux

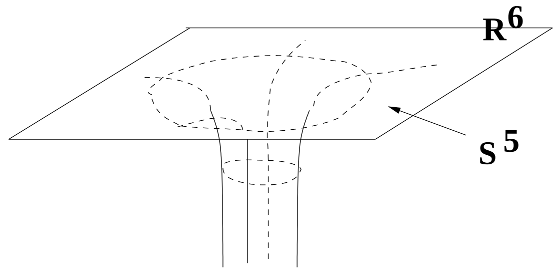


Figure D.5: Geometry of the background created by N coincident D3-branes.

piercing through a 5-sphere surrounding the origin in the transverse 6d space, $\int_{\mathbf{S}^5} F_5 = N$.

Notice that we say that this description is given by *full* string theory in this background, namely we assume that we include all stringy (i.e. α' corrections) and quantum corrections of the background.

We now would like to take the same kind of limit as above. First, it is a low energy limit; this corresponds to sending the string scale to infinity (i.e. $\alpha' \to 0$), keeping energies, as measured by an asymptotic observer in the above spacetime geometry, fixed. Second, we want to take the limit keeping energies of excitations in the near core region $r \simeq 0$ finite. Due to the non-trivial g_{tt} metric component, an excitation of proper energy $E_{(r)}$ localized at r in the radial direction, has an energy

$$E_\infty = E_{(r)} Z(r)^{-1/4} \tag{D.12}$$

as measured in the reference frame of an observer at infinity. That is, as an excitation approaches $r \simeq 0$, its energies measured in the reference frame of the asymptotic observer suffers a large redshift. For excitations near $r \simeq 0$, the above relation reads

$$E_\infty \simeq E_{(r)} \frac{r}{R} = (E_{(r)} \alpha'^{1/2}) \frac{r}{(g_s N)^{1/4} \alpha'} \tag{D.13}$$

We want to take $\alpha' \to 0$ keeping E_∞ fixed (large string scale keeping fixed energy) and $E_{(r\to 0)} \alpha'^{1/2}$ finite (finite energy for excitations in the near core region). This corresponds to taking $r \to 0$, $\alpha' \to 0$ keeping r/α' finite.

There are two decoupled sector that survive in this low energy limit. There is a sector of fields propagating in the asymptotically flat region, which

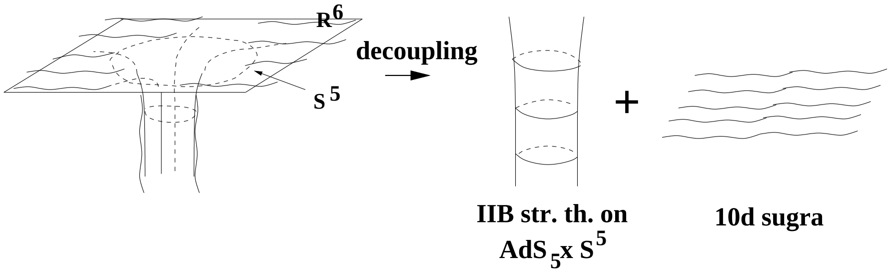


Figure D.6: Maldacena's low energy limit of the system of N coincident D3-branes, described by type IIB theory on the D3-brane background. In the limit we obtain full type IIB string theory on the near core limit $\mathrm{AdS}_5 \times \mathbf{S^5}$, decoupled from free 10d supergravity modes.

suffer no redshift; so the only fields surviving in the low energy limit are the massless 10d supergravity fields, which are free fields in this limit. A second sector corresponds to modes localized in the $r \simeq 0$ region; these fields suffer an infinite redshift, hence modes of arbitrarily large proper energy have small energy measured in the asymptotic reference frame, and survive in the low energy limit. This second sector is described by the full type IIB string theory on the background

$$\begin{gathered} ds^2 = \tfrac{r^2}{R^2}(\eta_{\mu\nu}dx^\mu dx^\nu) + \tfrac{R^2}{r^2}dr^2 \; + \; R^2 d\Omega_5^2 \\ \textstyle\int_{\mathbf{S^5}} F_5 = N \end{gathered} \tag{D.14}$$

The first two pieces of the metric describe a 5d anti de Sitter space AdS_5, of radius (or rather, length scale) R, while the last terms describes a 5-sphere. Through the latter there are N units of RR 5-form flux. See figure D.6.

Since the limit involves $r \to 0$, it is useful to rewrite the above metric in terms of the quantity $U = r/\alpha'$, which remains finite in the limit. We have

$$ds^2 = \alpha' \left[\frac{U^2}{(4\pi g_s N)^{1/2}}(\eta_{\mu\nu}dx^\mu dx^\nu) + (4\pi g_s N)^{1/2}\frac{dU^2}{U^2} \; + \; (4\pi g_s N)^{1/2} d\Omega_5^2 \right] \tag{D.15}$$

The overall factor of α' simply encodes the fact that we are zooming into the region of small r.

The Maldacena conjecture is that both descriptions, in terms of gauge field theory (plus 10d free gravitons) and in terms of string theory on the $\mathrm{AdS}_5 \times \mathbf{S^5}$ background (plus 10d free gravitons) are completely equivalent.

We thus propose the complete equivalence of 4d $\mathcal{N} = 4$ $SU(N)$ super Yang-Mills [6] gauge field theory with full fledged type IIB string theory on $\mathrm{AdS}_5 \times \mathbf{S^5}$, with radius R^2/α' given above, and N units of 5-form flux through $\mathbf{S^5}$.

Let us emphasize once again that the equivalence involves full string theory, including all stringy modes, brane states, etc. Again, this is because arbitrarily high energy modes survive in the throat region in the limit.

This correspondence is very striking. It proposes that a string theory (in a particular background) is completely equivalent to a gauge field theory. It is very striking that a theory that includes gravity, and an infinite set of fields, can be equivalent to a non-gravitational theory, which in principle looks much simpler. We will see later on how this correspondence works in more detail, although an extensive discussion is beyond the scope of this lecture. Let us also point out that this kind of relation, in the limit of large N, had been proposed by 't Hooft, see appendix.

The dictionary between the parameters of the gauge theory and the string theory are as follows

$\mathcal{N} = 4$ $SU(N)$ super Yang-Mills		Type IIB on $\mathrm{AdS}_5 \times \mathbf{S^5}$
$\tau = \theta + i\frac{1}{g_{YM}^2}$		$\tau = a + i\frac{1}{g_s}$
$N =$ number of colors		$N =$ flux
$\lambda = g_{YM}^2 N$	$\overset{R^2/\alpha' = 4\pi g_{YM}^2 N}{\longleftrightarrow}$	R^2/α'

D.3.2 Some preliminary tests of the proposal

Some additional support for the above proposal is that the two systems have the same symmetry structure.

• The $SO(6)$ isometry group of $\mathbf{S^5}$ on the string theory side exactly reproduces the $SO(6)$ R-symmetry group of the $\mathcal{N} = 4$ gauge field theory. This is analogous to the observation we made in previous section for systems of D3-branes.

• The isometry group of AdS_5 is $SO(4,2)$. This can be seen from the follwoing construction of AdS_5 space. Consider the hypersurface

$$(X^0)^2 + (X^5)^2 - (X^1)^2 - (X^2)^2 - (X^3)^2 - (X^4)^2 = R^2 \qquad \text{(D.16)}$$

[6] A subtle issue is that the Maldacena gauge/string correspondence holds for the $SU(N)$ group, rather than for $U(N)$. The difference in the large N limit is of order $1/N^2$, and hence only detectable by computing loop corrections.

in the 6d flat space with signature $(4, 2)$ and metric

$$ds^2 = -(dX^0)^2 - (dX^5)^2 + (dX^1)^2 + (dX^2)^2 + (dX^3)^2 + (dX^4)^2 \quad \text{(D.17)}$$

Clearly the above hyperboloid is a 5d space of signatures $(3, 1)$ and isometry group $SO(4, 2)$.

Performing the change of variables

$$\begin{aligned} X^0 &= \frac{1}{2u}\left[1 + u^2(R^2 + (X^1)^2 + (X^2)^2 + (X^3)^2 - t^2)\right] \\ X^i &= Rux^i \\ X^4 &= \frac{1}{2u}\left[1 - u^2(R^2 - (X^1)^2 - (X^2)^2) - (X^3)^2 + t^2\right] \\ X^5 &= Rut \end{aligned} \quad \text{(D.18)}$$

the metric on the 5d space becomes

$$ds^2 = R^2u^2(-dt^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2) + \frac{R^2}{u^2}du^2 \quad \text{(D.19)}$$

And redefining $u = U\alpha'/R^2$, we get

$$ds^2 = \alpha'\left[\frac{U^2}{R^2/\alpha'}\left[-dt^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2\right] + \frac{R^2/\alpha'}{U^2}dU^2\right] \quad \text{(D.20)}$$

which is precisely (D.15). Hence AdS_5 has an isometry group $SO(4, 2)$.

This corresponds exactly to the conformal group of the 4d gauge field theory. $\mathcal{N} = 4$ theories at the origin of the Coulomb branch are conformally invariant, even at the quantum level (that is, the beta functions which encode the running of couplings with the scale are exactly zero, so the theory is scale invariant). The $SO(4, 2)$ conformal group has a $SO(3, 1)$ Lorentz subgroup and a $SO(1, 1)$ scale transformations subgroup.

Hence the isometry group of the AdS_5 theory reproduces the conformal group of the 4d gauge field theory [7]. An important fact in this context is that the scale transformations in the gauge field theory correspond to translations in the variable U on the AdS side. Namely, this subgroup acts on the AdS_5 geometry as

$$(t, x^1, x^2, x^3, u) \to (\lambda t, \lambda x^1, \lambda x^2, \lambda x^3, \lambda^{-1}u) \quad \text{(D.21)}$$

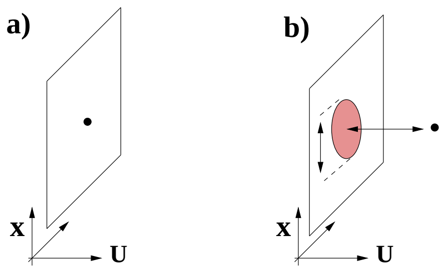


Figure D.7: The spread on the boundary of the effect of an excitation in the bulk of AdS spacetime is smaller as the excitation are localized closer to infinity in the U direction (boundary). In terms of the dual gauge field theory, the near boundary region corresponds to the ultraviolet, while the interior corresponds to the infrared.

Moreover, going to small lengths in the gauge field theory corresponds to going to infinity in U in the AdS theory, and vice versa. This is known as the UV/IR correspondence. See figure D.7.

• The supersymmetry structure is the same for both theories. The $\mathrm{AdS}_5 \times \mathbf{S^5}$ background preserves 32 supercharges. Sixteen of them were present in the full D3-brane solution, but sixteen additional one appear (accidentally) in taking the near core limit. The gauge field theory has also 32 supercharges, sixteen of them are the familiar ones of $\mathcal{N} = 4$ theories, while sixteen additional ones, known as superconformal symmetries, are generated by the previous supersymmetries and conformal transformations.

• There is non-geometric non-perturbative symmetry, which also matches in the two theories. This is the $SL(2, \mathbf{Z})$ self-duality of type IIB string theory, which corresponds to the $SL(2, \mathbf{Z})$ Montonen-Olive self-duality of $\mathcal{N} = 4$ $SU(N)$ super Yang-Mills.

It would be interesting to test the proposal beyond a mere matching of the symmetries of the system. However, we do not know how to quantize type IIB string theory on $\mathrm{AdS}_5 \times \mathbf{S^5}$. This is difficult because there is curvature in spacetimes, hence the 2d worldsheet theory is not free (and not exactly solvable, for the moment). In addition, there are RR fields in the background, and this makes the worldsheet theory even more complicated [8].

[7]Hence the familiar name of AdS/CFT correspondence.

[8]It is interesting to point out that a subsector of this theory (corresponding to a so-called

Therefore we can analyze this system only in the supergravity approximation, i.e. keeping the leading behaviour in α'/R^2. This will be a good approximation for R^2/α' large, when all length scales of the geometry are large, and when the density of RR field strength is small. In the language of the corresponding gauge field theory, this corresponds to the limit of large $\lambda = g_{YM}^2 N$, also known as 't Hooft limit (where the so-called 't Hooft coupling λ is large), see appendix. We also need to restrict to classical supergravity, hence we ignore string loop corrections, and take g_s to be small. This implies that the AdS side is tractable when $N \to \infty$m, $g_s \to 0$, and λ is finite and large. In this limit the gauge field theory is not tractable. As we will see in a moment, although g_s is small the right parameter weighting loops is λ, so at large λ the perturbative expansion breaks down.

On the other hand, the gauge field theory is tractable in the perturbative regime, namely when g_s is small and N is small. In this limit, the string theory has a strongly coupled 2d worldsheet theory, and the supergravity approximation breaks down. Hence, the above correspondence is analogous to the duality relations studied in other lectures. There is an exact equivalence of two different descriptions, but when one of them is weakly coupled and tractable, the other is not.

The usual way in which the Maldacena correspondence is exploited is to consider the classical supergravity limit to compute certain quantities, protected (or expected to be protected) by supersymmetry. These quantities can then be computed in perturbative gauge theory, extrapolated to the 't Hooft limit, and compared with the supergravity result. We will discuss some example in next section. For quantities not protected by any symmetry, the supergravity result need not agree with the perturbative gauge theory result. It can then be regarded as a prediction for the behaviour of that quantity in the 't Hooft limit.

In this setup, the mapping of systematic corrections beyond the 't Hooft limit / classical supergravity limit is as follows

gauge theory side	string theory side
λ corrections	α'/R^2 corrections
$\lambda/N = g_{YM}^2$ corrections	g_s loop corrections

Penrose limit), given by type IIB string theory on $AdS_5 \times \mathbf{S^5}$ can be quantized exactly in α' in the light-cone gauge. In this situation it is possible to find stringy effects/states and try to identify them in the gauge theory description, with great success [97].

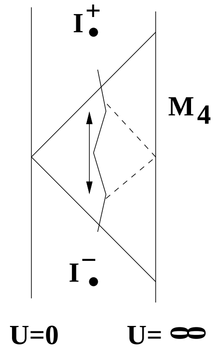


Figure D.8: Penrose diagram from AdS spacetimes (only the directions U are shown). Light-rays travel at 45 degrees. The points $I^{\pm}$ are the infinite future/past of timelike lines. A timelike observer can send a light signal to the boundary and get it back in finite proper time.

This agrees very nicely with the picture of corrections to the 't Hooft limit in gauge field theories, see appendix.

D.3.3 AdS/CFT and holography

For these section, see [104] and [99]. These authors have proposed a precise recipe to obtain correlation functions in the super Yang-Mills theory, via a computation in type IIB theory in the $\mathrm{AdS}_5 \times \mathbf{S^5}$ background. Moreover, the proposal leads to a nice interpretation of 'where' the field theory is living, in the AdS picture, For most of the discussion in this section, the essential features arise from the geometry of AdS spaces. This suggests a generalization of the correspondence to a relation between type IIB string theory on $\mathrm{AdS}_5 \times \mathbf{X_5}$ (with $\mathbf{X_5}$ a compact Einstein space) and 4d conformal gauge field theories with lower or no supersymmetry.

AdS_5 space has a conformal boundary at $U \to \infty$, which is 4d Minkowski space M_4 (plus a point). That is, there is a conformally equivalent metric ($ds^2 \to e^{2f(t,x,U)} ds^2$) such that infinity is brought to a finite distance . The Penrose diagram, encoding the causal structure of the AdS space (light-like geodesics run at 45 degrees in the diagram) is shown in figure D.8. A remarkable feature of AdS spacetime is that an timelike observer can send a light signal to the boundary of spacetime, and receive the reflected signal a finite amount of time after sending it, see figure D.8. This means that, although the boundary of AdS spacetime is at infinity, information on the boundary can interact with information in the bulk within finite time. Hence, AdS_5 behaves as a box of finite size, and this makes it important to specify boundary conditions in order to define any theory on AdS_5.

For instance, the partition function of the theory (namely, the vacuum path integral over all the spacetime fields of the theory on AdS space) is in general a functional of the boundary values ϕ_0 for all the spacetime fields ϕ of the theory

$$Z_{\rm part.funct.}[\phi_0] = \int \mathcal{D}(\text{IIB fields})\, e^{-S_{\rm spacetime\ IIB}}[\text{fields}] \tag{D.22}$$

The importance of boundary conditions, along with the fact that the boundary M_4 of AdS_5 spacetime is of the same form as the space on which the gauge field theory lives, motivates the following proposal. Quantities in the gauge field theory on M_4 provide the boundary conditions for fields (of the string theory) propagating on the AdS_5 spacetime. More precisely, the proposal is

• For each field ϕ propagating on AdS_5 there is an operator $\mathcal{O}_\phi$ in the gauge field theory. The field in AdS_5 can be *any* field associated to a 5d state of string theory in $AdS_5 \times \mathbf{S^5}$, for instance a massless 5d supergravity mode, any mode in the KK reduction of the massless 10d supergravity mode, any massive 10d string state, or even any state from the non-perturbative sectors of the type IIB string theory. The properties of $\mathcal{O}_\phi$, ϕ, like their behaviour under the symmetries of the systems, are related as we discuss a bit later.

• The value ϕ_0 of ϕ at the boundary at infinity

$$\phi_0(t,x) = \lim_{U\to\infty}\phi(t,x,U) \tag{D.23}$$

is a function on M_4. This value acts as a source for the corresponding operator $\mathcal{O}_\phi$ in the field theory in M_4, namely, its lagrangian includes a term $\Delta\mathcal{L} = \phi_0\mathcal{O}_\phi$. Equivalently, a term in the lagrangian of the gauge field theory (given by a linear combination of operators, with some coefficients) corresponds to introducing specific boundary conditions for the corresponding fields in AdS space. Hence, the field theory data can be regarded as encoded in the boundary of AdS space, and as providing boundary conditions for string theory in AdS space.

• Correlation functions of operators $\mathcal{O}_\phi$ in the gauge field theory can be computed by taking functional derivatives of a generating functional

$Z_{\rm gauge}[\phi_0]$, which is a path integral with a source ϕ_0 for the operator

$$Z_{\rm gauge}[\phi_0] = \int \mathcal{D}(\text{gauge th. fields}) e^{-S_{YM}+\phi_0\mathcal{O}_\phi} \tag{D.24}$$

For instance, the two-point correlation function

$$\langle \mathcal{O}_\phi \mathcal{O}_\phi \rangle = \frac{\delta Z[\phi_0]}{\delta\phi_0\delta\phi_0}\,|_{\phi_0=0} \tag{D.25}$$

The proposal is that the partition function $Z_{\rm part.func.}[\phi_0]$ of IIB theory on AdS_5 with boundary conditions ϕ_0 for the 5d field ϕ (this for all fields of the theory), corresponds exactly to the generating functional $Z_{\rm gauge\,th}[\phi_0]$ of the gauge field theory, with ϕ_0 as source term for the corresponding operator $\mathcal{O}_\phi$. That is

$$Z_{\rm part.funct}[\phi_0] = Z_{\rm gauge\,th}[\phi_0] \tag{D.26}$$

This is a precise correspondence that allows to encode all the dynamics of string theory on AdS in the dynamics of gauge field theory, and vice versa.

The above proposal can be used to obtain a relation between the mass m of a 5d field ϕ in string theory AdS_5 (which appears in the computaion of the partition function in the free field approximation) and the conformal dimension Δ of the corresponding operator $\mathcal{O}_\phi$ in the gauge field theory (which appears in the two-point correlation function). The relation, for a p-form field in AdS_5, reads

$$(\Delta + p)(\Delta + p - 4) = m^2 \tag{D.27}$$

One can verify this matching by considering operators whose conformal dimensions are protected by supersymmetry. For instance, chiral operators are operators which belong to chiral multiplets when the $\mathcal{N} = 4$ theory is written in terms of the $\mathcal{N} = 1$ subalgebra. For instance, chiral operators are $\mathrm{Tr}\,(\Phi_{i_1}\ldots\Phi_{i_r})$, or $\mathrm{Tr}\,(W_\alpha W_\alpha \Phi_{i_1}\ldots\Phi_{i_r})$. Chiral operators are BPS like, in the sense that they belong to shorter multiplets, and their conformal dimensions is related to their R-charge. The conformal dimensions can then be computed in the perturbative Yang-Mills theory (small g_s, small λ), and then extrapolated and compared with the masses of (BPS) states in the string theory side. These states are easy to identify and correspond to the KK reduction on $\mathbf{S^5}$ of massless 10d supergravity modes. The perfect matching between towers of KK modes in AdS_5 and infinite sets of opertors in the gauge theory is a strong check of the correspondence.

Beyond these kind of checks, which have been extended in several directions, there are other qualitatively different checks that we would like to mention

• There is a precise recipe to compute Wilson loops in the gauge field theory (expectation values of operators given by path ordered integrals of the gauge field over a circuit C in 4d) from the string theory side, as the action of a minimal area worldsheet asymptoting to the circuit C as it approaches the boundary at infinity [100].

• Some D-brane states in string theory on $AdS_5 \times \mathbf{S^5}$ have been identified to operators in the gauge theory. For instance, a D5-brane wrapped on $\mathbf{S^5}$ has been shown to correspond to a baryonic operator in the gauge field theory [101].

• Taking a particular limit of the correspondence, which amounts to centering on a subsector of states/operators with large $SO(6)$ quantum numbers, a complete matching of stringy states and operators has been carried out [97]. On the string theory side, the limit reduces to string theory on a pp-wave background, which can be quantized exactly in α'.

D.3.4 Implications

We would like to conclude by mentioning some implications of this far-reacing correspondence

• It is a holographic relation! A theory with gravity in 5d is described in terms of a non-gravitational theory with degrees of freedom in 4d. This has deep implications for instance on question like the information problem in black holes in AdS space. The correspondence with gauge theory allows (in principle, although in practice it is not known how to do it) to describe the process of creation and evaporation of a black hole purely in terms of a manifestly unitary quantum field theory. Hence, no violation of the rules of quantum mechanics is involved.

• The correspondence provides a complete non-perturbative definition of string/M theory, in a particular background. This certainly changes the way to think about string theories. It is however difficult to extract the main physical principles to allow to develop a background independent definition.

• The correspondence and its generalizations provides a new powerful tool to analyze gauge field theories in the 't Hooft limit using supergravity duals.

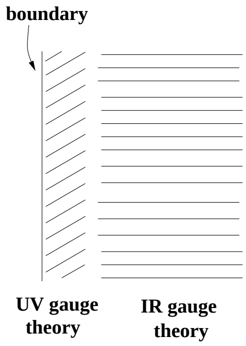


Figure D.9: Rough holographic picture for non-conformal gauge field theories.

In particular it has been possible to describe non-conformal theories in these terms, for instance by finding the field profiles that must be introduced in supergravity to describe the introduction of mass terms for some of the matter fields of the $\mathcal{N} = 4$ theories. The gauge theories are a small perturbation of $\mathcal{N} = 4$ in the ultraviolet, and flow to interacting non-conformal theories in the infrared, sometimes showing interesting behaviour like confinement, etc. In the supergravity side, the solutions are asymptotically AdS near the boundary at infinity, and are deformed in the inside. The structure in the inside region reproduces the infrared features of the gauge field theory. See figure D.9. For instance, confinement in the gauge field theory is usually associated to the presence of a black hole in the interior of the 5d space (see e.g. [102]).

The gauge/string correspondence is one of the deepest recent results in string theory and gauge theory. A lot of research is devoted to gaining a better understanding of the lessos it has for us concerning the nature of string theory, of holography, and of a new language to describe gauge field theory phenomena.

.1 Large N limit

The Maldacena correspondence fits well with 't Hooft's proposal that the large N limit of gauge field theories seems to be described by a string theory

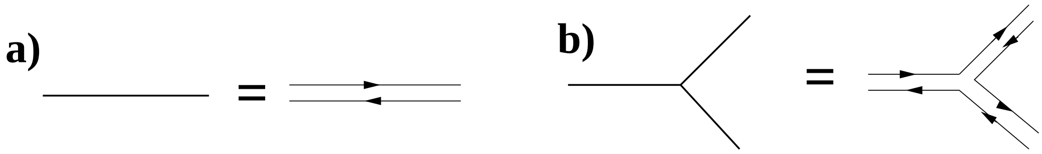


Figure 10: Propagator and 3-point interaction vertex in double line notation.

[9].

The main observation is that in a $SU(N)$ gauge theory, the effective coupling constant is not g_{YM}^2, but $\lambda = g_{YM}^2 N$. The factor of N intuitively takes into account that the number of particles running in loops increases as we increase the number of colors. Hence in the large N limit perturabation theory breaks down, no matter how small g_{YM}^2 is. However, 't Hooft realized that there are additional simplifications in this limit, of large N keeping λ finite, that suggests it might have a simple description in terms of a dual string theory. Namely in this 't Hooft limit, for any amplitude the Feynmann expansion chan be recast as a double expansion in λ and $1/N$.

To understand this, let us introduce the double line notation, where in a Feynmann diagram a field in the adjoint representation is drawn as a pair of oppositely oriented arrows (can be thought of as representing degrees of freedom in the fundamental and antifundamental representations), see figure 10. One can classify diagrams according to its number of vertices V, external lines E, and closed loops of lines F. From (D.1) each vertex is weighted by N/λ, while each propagator is weighted by λ/N, while each loop of lines gives a factor of N. Each diagram is therefore weighted by a factor

$$N^{V-E+F}\,\lambda^{E-V} \tag{28}$$

The number $\xi = V - E + F$ is known as the Euler number of the diagram, and g, defined by $\xi = 2 - 2g$, is knons as the genus of the diagram. We have the double expansion

$$\sum_{g=0}^{\infty} N^{2-2g} \sum_{i=0}^{\infty} \lambda^i\, c_{g,i} \tag{29}$$

[9] 't Hooft was interested in QCD, and hence on non-supersymmetric and confining pure gauge theories, where the string is supposed to correspond to the confined gauge field fluxlines. The AdS/CFT has shown that similar ideas actually extend (often in a subtle way) to non-confining theories as well.

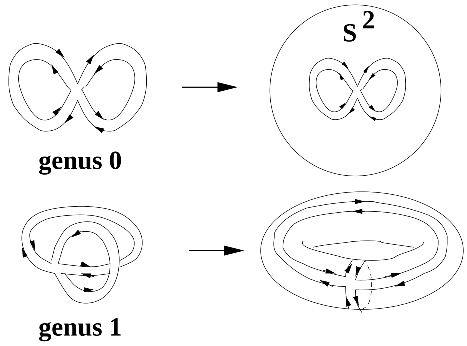


Figure 11: Two two-loop diagrams at genus 0 and 1.

The genus has a geometric intrepretation. A diagram of genus g if such that it can be drawn in double line notation without crossings on a Riemann surface of genus g, and cannot be drawn in a surface of genus $g-1$. In figure 11 we show two-loop diagrams of genus 0 and 1.

In the large N limit, keeping λ fixed, any amplitude has a genus expansion [10] in $1/N$. Hence, the large N limit is dominated by the so-called planar diagrams, which correspond to genus 0. This limit corresponds to a weakly coupled string theory, which is dominated by the genus 0 terms. The expansion in $1/N$ is supposed to reproduce the genus expansion of the dual string theory. Geometrically, this corresponds to 'filling the holes' of the gauge theory diagram in the double line notation to form the corresponding Riemann surface. This has been physically understood in a related gauge/string duality context in [103].

[10] Note that, since λ is fixed, one can recast the series as an expansion in $\lambda/N = g_{YM^2}$, which becomes the coupling constant of the string theory (in fact, it is g_s in the AdS/CFT case).

Appendix A

Brane-worlds

A.1 Introduction

We have seen that branes in string theory may lead to gauge sectors localized on their world-volumes. This can be exploited, as we did in previous lecture, to take a decoupling limit where dynamics reduces to gauge field theory, and try to use string theory tools to gain new insights into gauge field theory dynamics.

In this lecture we would like to center on a different application of branes and their gauge sectors. There exist string theory or M-theory *vacua* with gauge sectors localized on the volume of branes, or on lower-dimensional subspaces of spacetime. For intance, in Horava-Witten compactifications, or in type I' theory (or its T-dual versions). These vacua can be regarded as a new possible setup in which to construct four-dimensional models with physics similar to that of the observed world, i.e. gravitational and gauge interactions, which charged chiral fermions. In this lecture we discuss different possible constructions containing gauge sectors that come close enough to the features of the Standard Model. Their main novelty is that gravitaional interactions and gauge interactions propagate over different spaces. See figure A.1. This implies a different scaling of their interaction strength as functions of the underlying parameters/moduli of the model.

Heterotic model building

To understand better this point, recall the setup of compactifications of heterotic string theory on Calabi-Yau manifolds $\mathbf{X_6}$. The 4d gauge group is given by the commutant of H in G (namely the elements of G commuting

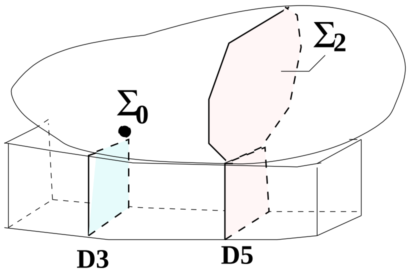


Figure A.1: In compactifications with D-branes, the gauge sectors like the Standard Model could propagate just on a lower-dimensional subpace of spacetime, e.g. the volume of a suitable set of D-branes, like any of the shaded areas.

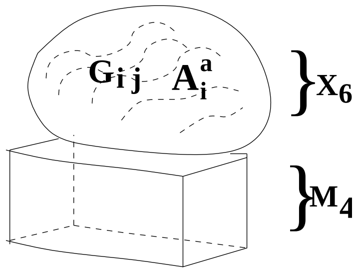


Figure A.2: Picture of heterotic string compactification.

with H), where G is the 10d $E_8 \times E_8$ or $SO(32)$. Thus, 4d gauge interactions are inherited from 10d ones, and so propagate all over 10d spacetime. Fig. A.2 shows configurations of this kind.

A very important property in this setup is the value of the string scale, which follows form analyzing the strength of gravitational and gauge interactions, as we quickly review. The 10d gravitational and gauge interactions have the structure

$$\int d^{10}x \, \frac{M_s^8}{g_s^2} \, R_{10d} \quad ; \quad \int d^{10}x \, \frac{M_s^6}{g_s^2} \, F_{10d}^2 \tag{A.1}$$

where M_s, g_s are the string scale and coupling constant, and R_{10d}, F_{10d} are the 10d Einstein and Yang-Mills terms. Powers of g_s follow from the Euler characteristic of the worldsheet which produces interactions for gravitons and gauge bosons (the sphere). Upon Kaluza-Klein compactification on $\mathbf{X_6}$, these interactions reduce to 4d and pick up a factor of the volume V_6 of $\mathbf{X_6}$

$$\int d^4x \, \frac{M_s^8 \, V_6}{g_s^2} \, R_{10d} \quad ; \quad \int d^4x \, \frac{M_s^6 \, V_6}{g_s^2} \, F_{10d}^2 \tag{A.2}$$

From this we may express the experimental 4d Planck scale and gauge coupling in terms of the microscopic parameters of the string theory configuration

$$M_P^2 \, = \, \frac{M_s^8 \, V_6}{g_s^2} \, \simeq \, 10^{19} \, \mathrm{GeV} \quad ; \quad \frac{1}{g_{YM}^2} \, = \, \frac{M_s^6 \, V_6}{g_s^2} \, \simeq \, \mathcal{O}(.1) \tag{A.3}$$

From these we obtain the relation

$$M_s \, = \, g_{YM} \, M_P \, \simeq \, 10^{18} \, \mathrm{GeV} \tag{A.4}$$

which implies that the string scale is necessarily very large in this kind of constructions. The key points in the derivation are that all interactions propagate on the same volume, and their strengths have the same dilaton dependence.

Brane-world constructions

Models where gravitational and gauge interactions propagate on different spaces are knonw as brane-worlds, since fields in the Standard Model (those that make up the observable world) are localized on some brane (or in general, some subspace of spacetime. In these constructions 4d gauge and gravitational interaction strength have a different dependence on the internal volumes.

The prototypical case [1] is provided by a compactification of type II theory (or some orientifold quotient thereof) on a 6d space $\mathbf{X_6}$, with a gauge sector localized on the volume of a stack of Dp-branes [2] wrapped on a $(p-3)$-cycle $\Pi_{(p-3)}$, with $\Pi_{(p-3)} \subset \mathbf{X_6}$. Namely, the $(p+1)$-dimensional world-volume of the Dp-brane is of the form $M_4 \times \Pi_{(p-3)}$. Before compactification, gravitational and gauge interactions are described by an effective action

$$\int d^{10}x \, \frac{M_s^8}{g_s^2} \, R_{10d} \, + \int d^{p+1}x \, \frac{M_s^{p-3}}{g_s} \, F_{(p+1)d}^2 \tag{A.5}$$

[1]The following analysis does not apply directly to Horava-Witten compactifications, see [104] for the corresponding discussion.

[2]For the moment, the D-brane configuration is simplified for convenience. Later on we will see detailed configurations leading to interesting world-volume spectra.

where the powers of g_s follow from the Euler characteristic of the world-sheet which produces interactions for gravitons (sphere) and for gauge bosons (disk).

Upon compactification, the 4d action picks up volume factors and reads

$$\int d^4x\, \frac{M_s^8 V_6}{g_s^2}\, R_{4d} + \int d^4x\, \frac{M_s^{p-3} V_{\mathbf{\Pi}}}{g_s}\, F_{4d}^2 \tag{A.6}$$

This allows to read off the 4d Planck mass and gauge coupling, which are experimentally measured.

$$\begin{aligned} M_P^2 &= \frac{M_s^8 V_{X_6}}{g_s^2} \simeq 10^{19}\,\mathrm{GeV} \\ 1/g_{YM}^2 &= \frac{M_s^{p-3} V_{\mathbf{\Pi}}}{g_s} \simeq 0.1 \end{aligned} \tag{A.7}$$

If the geometry is factorizable, we can split $V_{X_6} = V_{\mathbf{\Pi}} V_\perp$, with $V_\perp$ the transverse volume, and obtain

$$M_P^2\, g_{YM}^2 = \frac{M_s^{11-p} V_\perp}{g_s} \tag{A.8}$$

This shows that it is possible to generate a large Planck mass in 4d with a low string scale, by simply increasing the volume transverse to the brane, or tuning the string coupling. In particular, it has been proposed to lower the string scale downto the TeV scale to avoid a hierarchy with the weak scale [105, 106]. The hierarchy problem is recast in geometric terms, namely the stabilization of the compactification size in very large volumes. These are difficult to detect since they are only felt by gravitational interactions. Present bounds on the size of 'gravity-only' extra dimensions come from tabletop experiments (like the Cavendish experiment), and impose only that their length scale in not larger than 0.1 millimeter. Notice however that a low string scale is not compulsory in models with some solution to the hierarchy problem, e.g. supersymmetric models.

A.2 Model building: Non-perturbative heterotic vacua

In this and the following section, we describe the basic rules for the construction of vacua of string theory or M-theory, with localized gauge sectors with features similar to those of the Standard Model. Explicit models with spectrum extremely close to that of the Standard Model have been constructed. However, in this lecture we will be happy by simply describing the apperance of charged chiral fermions, and the underlying reason for family replication. More detailed model building issues are left for the references.

We start by considering the setup provided by compactifications of Horava-Witten theory. This can be considered as the strong coupling limit of compactifications of the $E_8 \times E_8$ heterotic string theory, and hence most of the tools are already familiar. There are however some interesting new ingredients.

Consider M-theory compactified to 4d on $\mathbf{X_6} \times \mathbf{S^1/Z_2}$. In general we will be interested in supersymmetric models, hence we choose $\mathbf{X_6}$ to be a Calabi-Yau threefold [3].

As in compactifications of the heterotic string theory, the

compactification is required to satisfy certain consistency conditions, arising from the equation of motion for some p-form fields. Namely, in heterotic theory the interactions for the NSNS 6-form B_6

$$\int_{10d} B_6 \wedge *B_6 + \int_{10d} B_6 \wedge (\mathrm{tr}\, F^2 - \mathrm{tr}\, R^2) \tag{A.9}$$

led to the equation of motion for the NSNS 2-form

$$dH_3 = \mathrm{tr}\, F^2 - \mathrm{tr}\, R^2 \tag{A.10}$$

In Horava-Witten theory, we need to consider two gauge bundles on the 10d boundaries of the interval, each with structure group a subgroup of E_8. The action for the 6-form C_6 (which is just the lift of the heterotic B_6) reads

$$\begin{aligned} S_{C_6} \;=\; & \int_{11d} *G_7 \wedge G_7 + \\ & + \int_{11d} \delta(x^{10})(\mathrm{tr}\, F_{E_8}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \wedge C_6 + \int_{11d} \delta(x^{10} - \pi R)(\mathrm{tr}\, F_{E_8'}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \wedge C_6 = \end{aligned}$$

[3]A motivation for supersymmetry in this setup is that there is only one 'gravity-only' dimension. If we build a non-supersymmetric model, and try to lower the 11d Planck scale to the TeV range to avoid a hierarchy problem, we should take this dimension very large to generate a large 4d Planck scale. In fact, so large that it would conflict with the experimental bounds. Hence, a large 11d Planck scale is convenient in this setup, and supersymmetry is the most reasonable way to stabilize the weak scale against it. It however may be somewhat lower than the 4d Planck scale.

$$= \int_{11d} dG_4 \wedge C_6 + \\ + \int_{11d} \delta(x^{10})(\mathrm{tr}\, F_{E_8}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \wedge C_6 + \int_{11d} \delta(x^{10} - \pi R)(\mathrm{tr}\, F_{E_8'}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \wedge C_6$$

where $\delta(x)$ is a bump 1-form localized in the interval. We have a similar equation of motion for the M-theory 3-form, namely

$$dG_4 = \delta(x^{10})(\mathrm{tr}\, F_{E_8}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \, + \, \delta(x^{10} - \pi R)(\mathrm{tr}\, F_{E_8'}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \quad \text{(A.11)}$$

Taking this relation in cohomology, we obtain

$$[\mathrm{tr}\, F_{E_8}^2] + [\mathrm{tr}\, F_{E_8'}^2] - [\mathrm{tr}\, R^2] = 0 \quad \text{namely} \quad c_2(E) = c_2(R) \qquad \text{(A.12)}$$

We would like to point out that the class of models is in fact richer. We can consider compactifications to 4d, where the background configuration also includes sets of k_a M5-branes [4] sitting at a point x_a^{10}in the interval, and with two of their world-volume dimensions wrapped on a 2-cycle $\Pi_a \subset \mathbf{X_6}$. Since the M5-branes are magnetically charged under the M-theory 3-form, the action for the 11d dual 6-form C_6 is

$$S_{C_6} = \int_{11d} *G_7 \wedge G_7 + \sum_a k_a \int_{M_4 \times \Pi_a} C_6 + \\ + \int_{11d} \delta(x^{10})(\mathrm{tr}\, F_{E_8}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \wedge C_6 + \int_{11d} \delta(x^{10} - \pi R)(\mathrm{tr}\, F_{E_8'}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \wedge C_6 = \\ = \int_{11d} dG_4 \wedge C_6 + \sum_a k_a \int_{11d} \delta(x^{10} - x_a^{10})\delta(\Pi_a) \wedge C_6 + \\ + \int_{11d} \delta(x^{10})(\mathrm{tr}\, F_{E_8}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \wedge C_6 + \int_{11d} \delta(x^{10} - \pi R)(\mathrm{tr}\, F_{E_8'}^2 - \frac{1}{2}\mathrm{tr}\, R^2) \wedge C_6$$

where $\delta(\Pi_a)$ is a bump 4-form with support on the 2-cycle Π_a. The equation of motion for C_6, taken in cohomology gives the consistency condition for this kind of compactification, which reads

$$[\mathrm{tr}\, F_{E_8}^2] + [\mathrm{tr}\, F_{E_8'}^2] + \sum_a k_a[\delta(\Pi_a)] - [\mathrm{tr}\, R^2] = 0 \qquad \text{(A.13)}$$

[4]Notice that taking the limit of small interval size shows that this possibility is also available in heterotic theory. Hence, there exist compactifications of heterotic on Calabi-Yau threefolds, with NS5-branes. Due to the presence of the latter, these vacua are non-perturbative, even if the string coupling is small.

where $[\Pi_a]$ is the 4-cohomology class dual to the 2-homology class of the 2-cycle $[\Pi_a]$. Namely, M5-branes contribute to the condition of cancellation of 6-form charge, via the homology class of the 2-cycle they wrap.

Compactifications with M5-branes have been studied in [107]. Since the M5-brane classes help in satisfying the consistency condition, it follows that there is additional freedom in the gauge bundles, and hence in the low-energy spectra of the theory. They lead to additional phenomena, for instance there may be transitions where some M5-brane moves towards the boundary in the interval and is diluted as an instanton class in the boundary gauge field. We will not go into these discussions.

Once the topology of the gauge bundles over the boundaries, namely their structure groups H, H', and characteristic classes, and the M5-brane configuration, are specified, the computation of the 4d massless spectrum is similar to that in heterotic theory.

• We obtain the 4d $\mathcal{N} = 1$ supergravity multiplet, the dilaton chiral multiplet, and $(h_{1,1}) + h_{2,1}$ chiral mulitplets arising from geometric moduli.

• We obtain vector multiplets for the gauge group given by the commutant of H, H' in E_8. Notice that the choice $H = SU(3)$, $H' = 1$ still leads to $E_6 \times E_8$, but does not correspond to embedding the spin connection into the gauge degrees of freedom, since the latter would involve both E_8 factos in a symmetric way.

• Charged chiral multiplets arise from the KK reduction of the 10d gaugino, and their multiplicity is given by the index of the Dirac operator coupled to the gauge bundle (in a representation corresponding to the their 4d gauge representation).

• There may be additional multiplets arising from the KK reduction of the M5-brane world-volume theory on the 2-cycle Π_a. These can be trickier to discuss, so we skip their details.

Taken overall, many of the features of these models are similar to compactifications of heterotic string theory. However, the existence of the 'gravity only' dimension allows to lower the fundamental scale somewhat below the 4d Planck scale.

A.3 Model building: D-brane-worlds

Another class of models with localized gauge sectors can be obtained by considering compactifications with D-branes. An additional advantage of

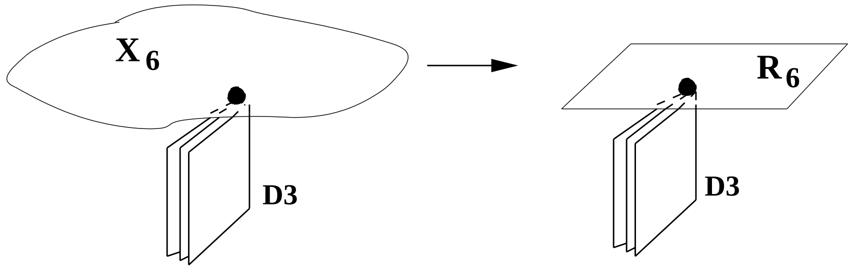


Figure A.3: Isolated D-branes at a smooth point in transverse space feel a locally trivial geometry and lead to non-chiral open string spectra.

these setups is that, for simple enough D-brane configurations (i.e. in the absence of curvatures) the quantization of open string sectors can be carried out exactly (in the sense of the α expansion).

A first issue that we should address is how to obtain D-brane sectors containing chiral fermions in the corresponding open string spectrum. In fact, the simplest D-brane configurations, like D-branes in flat space (or in toroidal compactifications), with trivial world-volume gauge bundle (zero field strength for world-volume gauge fields, preserve too much supersymmetry to allow for chirality (that is, they have at least 4d $\mathcal{N} = 2$ supersymmetry) [5].

In fact, we can heuristically argue that *isolated*D-branes sitting at a *smooth* point in transverse space lead to non-chiral open string spectra. Considering for instance the case of D3-branes, sitting at a smooth point P in Transverse 6d space $\mathbf{X_6}$, see figure A.3. Since chiral matter is necessarily massless, if present it should arise from open strings located at P and stretching between the D3-branes. Hence, only the local behaviour of $\mathbf{X_6}$ around P is important. If P is smooth this local behaviour is that of $\mathbf{R^6}$, hence the massless open string sector if simply that on D3-branes in flat space, which is non-chiral.

There are two ways which have been used in the construction of D-brane Configurations with chiral open string sectors; they arise from relaxing each

[5]One way to generate chiral fermions is in fact to consider introducing a non-trivial bundle for the D-brane world-volume gauge field, with support on the internal cycle Π_{p-3}wrapped by the Dp-brane. This kind of model is, in some respects (like in the computation of the spectrum, etc) similar to heterotic models, and we do not discuss it here.

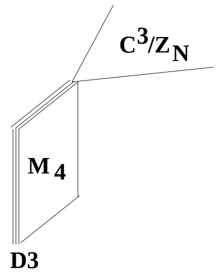


Figure A.4: Stack of D3-branes at an orbifold singularity

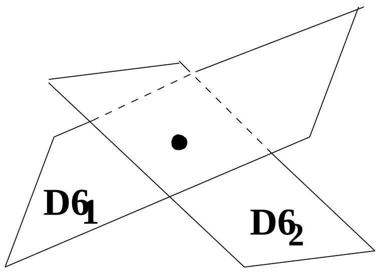


Figure A.5: Two intersecting D6-branes in flat space.

of the above conditions in italic writing:

• Relaxing the smoothness condition, we may consider D-branes sitting at singular points in transverse space. The prototypical example is provided by a stack of D3-branes located at an orbifold singularity, $\mathbf{C^3}/\mathbf{Z_N}$. See figure A.4.

• Relaxing the condition of isolatedness, we may consider configurations of D-branes intersecting over subspaces of their world-volume. The prototypical case is provided by D6-branes intersecting over 4d subpaces of their world-volumes. See figure A.5

In the following we discuss the appearance of chiral fermions, and the spectrum in these two kinds of D-brane configurations.

A.3.1 D-branes at singularities

For concreteness, let us center of a stack of n D3-branes sitting at the Origin of a $\mathbf{C^3}/\mathbf{Z_N}$ orbifold singularity. These models were first Considered in [108].

The $\mathbf{Z_N}$ generator θ acts on the three complex coordinates of $\mathbf{C^3}$ as follows

$$(z_1, z_2, z_3) \rightarrow (e^{2\pi I\, a_1/N} z_1, e^{2\pi i\, a_2/N} z_2, e^{2\pi i\, a_3/N} z_3) \tag{A.14}$$

where the $a_i \in \mathbf{Z}$ in order to have and order N action [6]. We will center on orbifolds that preserve some supersymmetry, hence their holonomy must be in $SU(3)$ and thus we require $a_1 \pm a_2 \pm a_3 = 0$ mod N, for some choice of signs.

The closed string spectrum in the configuration can be obtained using the techniques explained in the corresponding lecture. Moreover, this sector will be uncharged under the gauge group on the D-brane world-volume, so it is not too interest for our discussion and we skip it.

Concerning the open string sector, the main observation is that there are no twisted sectors. This follows because the definition of twisted sectors in closed strings made use of the periodicity in the worldsheet direction σ, and this is not allowed in open strings. Hence, the spectrum of open strings on a set of D3-branes at a $\mathbf{C^3}/\mathbf{Z_N}$ orbifold singularity is simply obtained by considering the open string spectrum on D3-branes in flat space $\mathbf{C^3}$, and keeping the $\mathbf{Z_N}$-invariant ones. Each open string state on D3-branes in flat space is given by a set of oscillators acting on the vacuum, and an $n \times n$ Chan-Paton matrix λ encoding the $U(n)$ gauge degrees of freedom. The action of θ on one such open string state is determined by the action on the corresponding set of oscillators and the action on the Chan-Paton matrix. For concreteness,let us center on massless states. The eigenvalues of the different sets of oscillators for these states are

Sector	State	θ eigenvalue
NS	$(0,0,0,\pm)$ 1	
	$(\underline{+,0,0},0)$	$e^{2\pi i\, a_i/N}$
	$(\underline{-,0,0},0)$	$e^{-2\pi i\, a_i/N}$
R	$\pm\frac{1}{2}(+,+,+,-)$	1
	$\frac{1}{2}(\underline{-,+,+},+)$	$e^{2\pi i\, a_i/N}$
	$\frac{1}{2}(\underline{+,-,-},-)$	$e^{-2\pi i\, a_i/N}$

The eigenvalues can be described as $e^{2\pi i\, r\cdot v}$, where r is The $SO(8)$ weight and $v = (a_1, a_2, a_3, 0)/N$. The above action can easily be understood by

[6]One also needs $N \sum_i a_i$ = even (so that the quotient is a spin manifold, i.e. allows spinors to be defined).

decomponsing the $SO(8)$ representation with respect to the $SU(3)$ subgroup in which the $\mathbf{Z_N}$ is embedded. In fact we have $8_V = 3 + \overline{3} + 1 + 1$, and $8_C = 3 + \overline{3} + 1 + 1$, and noticing that (A.14) defines the action on the representation 3. Notice that the fact that bosons and fermions have the same eigenvalues reflects the fact that the orbifold preserves $\mathcal{N} = 1$ supersymmetry on the D-brane world-volume theory. In fact we see that the different states group into a vector multiplet V, with eigenvalue 1, and three chiral multiplets, Φ_i with eigenvalue $e^{2\pi i\, a_i/N}$.

On the other hand, the action of θ on the Chan-Paton degrees of freedom corresponds to a $U(n)$ gauge transformation. This is defined by a unitary order N matrix $\gamma_{\theta,3}$, which without loss of generality we can diagonalize and write in the general form

$$\gamma_{\theta,3} = \text{diag}\,(1_{n_0}, e^{2\pi i/N} 1_{n_1}, \ldots, e^{2\pi i(N-1)/N} 1_{n_{N-1}}) \tag{A.15}$$

with $\sum_{a=0}^{N-1} n_a = n$. The action on the Chan-Paton wavefunction (which transforms in the adjoint representation) is

$$\lambda \to \gamma_{\theta,3} \lambda \gamma_{\theta,3}^{-1} \tag{A.16}$$

We now have to keep states invariant under the combined action of θ on the oscillator and Chan-Paton piece. For states in the $\mathcal{N} = 1$ vector multiplet, the action on the oscillators is trivial, hence the surviving states correspond to Chan-Paton matrices satisfying the condition

$$\lambda = \gamma_{\theta,3} \lambda \gamma_{\theta,3}^{-1} \tag{A.17}$$

The surviving states correspond to a block diagonal matrix. The gauge group is easily seen to be

$$U(n_0) \times \ldots \times U(n_{N-1}) \tag{A.18}$$

For the i^{th} chiral multiplet Φ_i, the oscillator part picks up a factor of $e^{2\pi i\, a_i/N}$. So surviving states have Chan-Paton wavefunction must satisfy

$$\lambda = e^{2\pi i\, a_i/N} \gamma_{\theta,3} \lambda \gamma_{\theta,3}^{-1} \tag{A.19}$$

The surviving multiplets correspond to matrices with entries in a diagonal shifted by a_i blocks. It is easy to see that the surviving multiplets transform in the representation

$$\sum_{i=1}^{3} \sum_{a=0}^{N-1} (\Box_a, \overline{\Box}_{a+a_i}) \tag{A.20}$$

We clearly see that in general the spectrum is chiral, so we have achieved the construction of D-brane configurations with non-abelian gauge symmetries and charged chiral fermions. Moreover, we see that in general the different fermions have different quantum numbers. The only way to obtain a replication of the fermion spectrum (i.e. a structure of families, like in the Standard Model), we need some of the a_i to be equal (modulo N). The most interesting example is obtained for the $\mathbf{C^3}/\mathbf{Z_3}$ singularity, with $v = (1, 1, -2)/3$. The spectrum on the D3-brane world-volume is given by

$$\begin{array}{rl} \mathcal{N} = 1 \text{ Vect.Mult.} & U(n_0) \times U(n_1) \times U(n_2) \\ \mathcal{N} = 1 \text{ Ch.Mult.} & 3\,[\,(n_0, \overline{n}_1, 1) + (1, n_1, \overline{n}_2) + (\overline{n}_0, 1, n_2)] \end{array} \qquad \text{(A.21)}$$

we see there is a triplication of the chiral fermion spectrum. Hence in this setup the number of families is given by the number of complex planes with equal eigenvalue.

We would like to point out that, as usual in models with open strings, there exist some consistency conditions, known as cancellation of RR tadpoles. Namely, there exist disk diagrams, see figure A.6, which lead to the coupling of D-branes at singularities to RR fields in the θ^k twisted sector. When the θ^k twist has the origin as the only fixed point, the corresponding RR fields do not propagate over any dimension transverse to the D-brane. This implies that they have compact support, and Gauss law will impose the corresponding charges must vanish, namely that the corresponding disk diagrams cancel. The coefficient of the disk diagram is easy to obtain: from the figure, we see that any worldsheet degree of freedom must suffer the action of θ^k as it goes around the closed string insertion. In particular it means that the Chan-Paton degrees of freedom suffer the action of $\gamma^k_{\theta^k,3=(\gamma_{\theta,3}}$ as they go around the boundary. Hence the disk amplitude is proportional to $\mathrm{tr}\,\gamma_{\theta^k,3}$, and the RR tadpole condition reads

$$\mathrm{Tr}\,\gamma_{\theta^k,3} = 0 \quad , \text{for } ka_i \neq 0 \bmod N \qquad \text{(A.22)}$$

For instance, for the above $\mathbf{Z_3}$ model these constraint require $n_0 = n_1 = n_2$. In general, the above constrains ensure that the 4d chiral gauge field theory on the volume of the D3-branes is free of anomalies.

Clearly the above model is not realistic. However, more involved models of this kind, with additional branes (like D7-branes, also passing through the singularity), can lead to models much closer to the Standard Model, see [109]. Their study is however beyond the topic of this lecture.

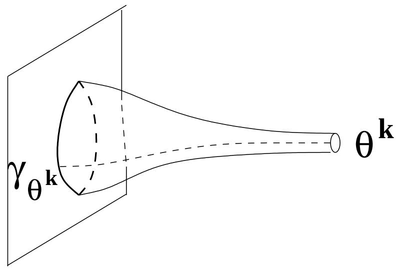


Figure A.6: D3-branes at singularities are charged under RR forms in the θ^k twisted sector, via a disk diagram. Worldsheet degrees of freedom suffer the action of θ^k as they go around the cut, shown as a dashed line. The amplitude is proportional to $\mathrm{tr}\,\gamma_{\theta^k}$.

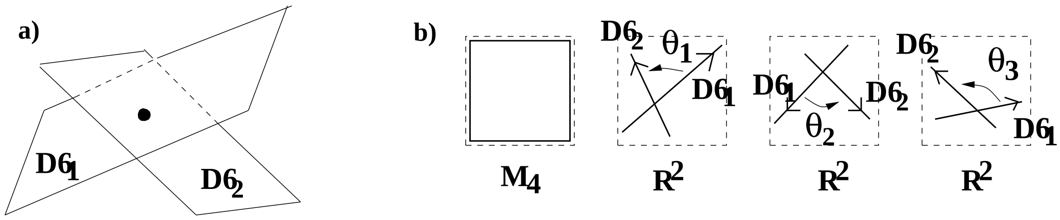


Figure A.7: Two picture of D6-branes intersecting over a 4d subspace of their volumes.

A.3.2 Intersecting D-branes

In this section we consider a different class of D-brane configurations leading to chiral 4d fermions. Consider two stacks of D6-branes (denoted $D6_1$- and $D6_2$-branes) in flat 10 space, intersecting over a 4d subspace of their worldvolumes, see figure A.7a. A slightly more explicit picture of the configuration is shown in figure A.7b. The local geometry is determined by the three angles $\pi\theta_i$ that relate the two D6-branes in the 6d space transverse to the 4d intersection. For the following analysis, see [110].

Two such sets of D6-branes, intersecting at general angles, break all the supersymmetries of the theory. The supersymmetries preserved by one of the stacks are broken by the other, and vice versa. Consider the $D6_1$-branes to span the direction 0123456. The supersymmetry transformations unbroken by these D6-branes are of the form $\epsilon_L Q_L + \epsilon_R Q_R$ with

$$\epsilon_L = \Gamma^0 \ldots \Gamma^6 \epsilon_R \tag{A.23}$$

where the subindices L, R denote the supersymmetries arising from the left or right movers. Denoting by R the $SO(6)$ rotation rotating the D6_1-branes to the D6_2-branes, the supersymmetries unbroken by the latter are

$$\epsilon_L = R^{-1}\,\Gamma^0 \ldots \Gamma^6\, R\,\epsilon_R \tag{A.24}$$

where here R denotes the action of the rotation in the spinor representation.

In general, there are no spinors surviving both conditions. However, for Special choices of the angles θ_i, i.e. of the rotation R, there may exist solutions to the above two conditions. In fact, it is easy to realize that if R is a rotation in an $SU(3)$ subgroup of $SO(6)$, there is one component of the spinor which is invariant under R, and both condition become identical. Therefore, intersections of D6-branes related by angles θ_i satisfying

$$\theta_1 \pm \theta_2 \pm \theta_3 = 0 \tag{A.25}$$

for some choice of signs, preserve 4 supercharges (1/4 of the supersymmetries preserved by the first stack of branes). This is the equivalent of 4d $\mathcal{N} = 1$, hence we may expect these configurations to lead to chiral 4d fermions. We will check below that this is indeed the case. Notice also that if the rotation is in a subgroup of $SU(2)$ (e.g. $\theta_1 \pm \theta_2 = 0$, $\theta_3 = 0$), the system preserves more spinors, in fact 8 supersymmetries, the equivalent of 4d $\mathcal{N} = 2$ supersymmetry.

Let us compute the spectrum of open strings in the above configuration of two Intersecting stacks of D6-branes, at generic angles θ_i. Consider open strings stretching among the N_1 D6_1-branes. This sector does not notice the presence of the second stack, so gives the same answers as for isolated D6-branes. We obtain $U(N_1)$ gauge bosons and their superpartners with respect to the 16 unbroken supersymmetries, propagating over the 7d volume of these D6-branes. For the sector of open strings stretching among the N_2 D6_2-branes, we similarly obtain $U(N_2)$ gauge bosons and their partners (under the 16 susys unbroken by the second D6-branes; notice these are not the same susy as above), propagating over the 7d volume of these D6-branes.

The novelty arises in the sector of open strings stretching between D6_1- and D6_2-branes. This sector feels both branes, and hence notices the amount of supersymmetry preserved by the two-stack system. We thus expect the spectrum in this sector to be non-supersymmetric for generic angles θ_i, and to gather into supermultiplets only for a constrained set of angles. Let us carry out the quantization of the sector of $6_1 6_2$ open strings. The only difference with respect to other open string sectors is in the boundary conditions. Consider two coordinates X_1, X_2 in a two-plane in which the D6-branes are rotated by an angle θ. The boundary conditions for the corresponding worldsheet fields for an open string are

$$\begin{aligned} \partial_\sigma X_1|_{\sigma=0} &= 0 \\ \partial_t X_2|_{\sigma=0} &= 0 \\ \cos\pi\theta\,\partial_\sigma X_1 \;+\; \sin\pi\theta\,\partial_\sigma X_2|_{\sigma=\ell} &= 0 \\ -\sin\pi\theta\,\partial_t X_1 \;+\; \cos\pi\theta\,\partial_t X_2|_{\sigma=\ell} &= 0 \end{aligned} \tag{A.26}$$

In complex coordinates $Z = X_i + iX_2$, we have

$$\begin{aligned} \partial_\sigma(\mathrm{Re}Z)|_{\sigma=0} &= 0 \\ \partial_t(\mathrm{Im}Z)|_{\sigma=0} &= 0 \\ \partial_\sigma(\mathrm{Re}e^{i\theta}Z)|_{\sigma=\ell} &= 0 \\ \partial_t(\mathrm{Im}e^{i\theta}Z)|_{\sigma=\ell} &= 0 \end{aligned} \tag{A.27}$$

Imposing these boundary conditions on the open string oscillator expansion leads to the constraints that: the center of mass position of the open string is located at the intersection point; momentum and winding are necessarily zero; oscillators have moddings shifted by $\pm\theta$. Applying this rule to the three complex coordinates corresponding to intersecting D6-branes, we obtain oscillators $\alpha^i_{n+\theta_i}$, $\alpha^{\bar{\imath}}_{n-\theta_i}$ for the complexified 2d bosons, and $\Psi^i_{n+\nu+\theta_i}$, $\Psi^{\bar{\imath}}_{n+\nu-\theta_i}$ for the 2d fermions, with $n \in \mathbf{Z}$ and $\nu = 1/2, 0$ for the NS and R sectors. The computation of the spectrum is formally similar to the computation of the spectrum on the left movers in an orbifold. In particular the fractional modding of oscillators introduces a modified vacuum energy. The final result for the spectrum, centering on light states, is as follows (we assume $\theta_i \in (-1/2, 1/2)$)

Sector	State	$\alpha' M^2$	4d Lorentz
NS	$\Psi^1_{-1/2+\theta_1}\|0\rangle$	$\frac{1}{2}(-\theta_1+\theta_2+\theta_3)$	Scalar
	$\Psi^2_{-1/2+\theta_2}\|0\rangle$	$\frac{1}{2}(\theta_1-\theta_2+\theta_3)$	Scalar
	$\Psi^3_{-1/2+\theta_3}\|0\rangle$	$\frac{1}{2}(\theta_1+\theta_2-\theta_3)$	Scalar
	$\Psi^1_{-1/2+\theta_1}\Psi^2_{-1/2+\theta_2}\Psi^3_{-1/2+\theta_3}\|0\rangle$	$\frac{1}{-}12(\theta_1+\theta_2+\theta_3)$	Scalar
R $\|0\rangle_R$	Weyl spinor		

All these fields propagate on the 4d intersection of the two D6-branes, and transform in the bifundamental representation (N_1, N_2) of the gauge group $U(N_1) \times U(N_2)$. The $6_2 6_1$ open string sector is quantized analogously, and in fact provides the antiparticles for the above fields. We see that generically bosons and fermions are unpaired, and only when the angles define a rotation in $SU(3)$ one of the bosons becomes massless and pairs up with the 4d fermion in the R sector, to give a 4d chiral multiplet. Notice that in the non-supersymmetric case, the scalars in the NS sector may have positive or negative mass square. If all scalars have positive mass square, the configuration of intersecting branes is stable. On the other hand, the existence of some tachyonic scalar signals an instability against a process in which the intersecting D6-branes recombine into a single smooth one. We will not say much more about this interesting process.

The important point in the above construction is that it provides a new setup with D-branes containing non-abelian gauge symmetries and charged chiral fermions. We now briefly describe how to exploit it in the construction of 4d models. For a review, see [111].

Although intersecting D6-branes provide 4d chiral fermions already in flat 10d space, gauge interactions remain 7d and gravity interactions remain 10d unless we consider compactification of spacetime. Hence, the general kind of configurations we are to consider (see figure A.8) is type IIA string theory on a spacetime of the form $M_4 \times \mathbf{X_6}$ with compact $\mathbf{X_6}$, and with stacks of N_a $\mathrm{D6}_a$-branes with volumes of the form $M_4 \times \Pi_a$, with $\Pi_a \subset \mathbf{X_6}$ a 3-cycle. It is important to realize that generically 3-cycles in a 6d compact space intersect at points, so the corresponding wrapped D6-branes will intersect at M_4 subspaces of their volumes. Hence, compactification reduces the 10d and 7d gravitational and gauge interactions to 4d, and intersections lead to charged 4d chiral fermions. Also, generically two 3-cycles in a 6d space intersect several times, therefore leading to a replicated sector of opens strings at intersections. This is a natural mechanism to explain/reproduce the appearance of replicated families of chiral fermions in Nature!

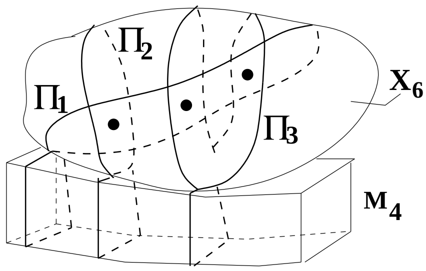


Figure A.8: Compactification with intersecting D6-branes wrapped on 3-cycles.

Denoting the 3-homology classes of the wrapped 3-cycles by $[\Pi_a]$, the intersection number is computed $I_{ab} = [\Pi_a] \cdot [\Pi_b]$, computed as described in the lecture on topology. The 4d spectrum on the resulting configuration is easy to obtain. From the sector of open strings strecthing among the $D6_a$-branes, we obtain the KK reduction on Π_a of the 7d $U(N_a)$ gauge bosons and partners. In general we obtain 4d $U(N_a)$ gauge bosons [7]. From the sector of open string stretching between the a^{th} and b^{th} stacks of D6-branes, we obtain a chiral 4d fermion in the bifundamental for each intersection of the corresponding 3-cycles. There are in general additional light scalars, which may become massless if the intersection is locally supersymmetric (ie the intersection angles define a rotation in $SU(3)$). Taken overall, the (chiral part of the) 4d spectrum is

$$\begin{array}{rl} \text{Gauge} & \prod_a U(N_a) \\ \text{Left.Ch.Fm.} & \sum_{a<b} I_{ab}(\Box_a, \overline{\Box}_b) \end{array} \qquad (A.28)$$

We note that a negative intersection number indicates the fermions have the opposite chirality.

These models have to satisfy some consistency conditions, namely cancellation of RR tadpoles. The D6-branes act as sources for the RR 7-forms via the disk coupling $\int_{W_7} C_7$. The consistency condition amounts to requiring the total RR charge of D-branes to vanish, as implied by Gauss law in a compact space (since RR field fluxlines cannot escape). The condition of RR tadpole cancellation can be expressed as the requirement of consistency

[7]Plus some partners if the 3-cycle Π_a is special lagrangian, i.e. the wrapped D-brane preserves some supersymmetry. We will not enter into this discussion.

of the equations of motion for RR fields. In our situation, the terms of the spacetime action depending on the RR 7-form C_7 are

$$\begin{aligned} S_{C7} &= \int_{M_4\times\mathbf{X_6}} H_8\wedge *H_8 + \sum_a N_a \int_{M_4\times\Pi_a} C_7 = \\ &= \int_{M_4\times\mathbf{X_6}} C_7\wedge dH_2 + \sum_a N_a \int_{M_4\times\mathbf{X_6}} C_7\wedge\delta(\Pi_a) \end{aligned} \quad \text{(A.29)}$$

where H_8 is the 8-form field strength, H_2 its Hodge dual, and $\delta(\Pi_a)$ is a bump 3-form localized on Π_a in $\mathbf{X_6}$. The equations of motion read

$$dH_2 = \sum_a N_a\,\delta(\Pi_a) \quad \text{(A.30)}$$

The integrability condition is obtained by taking this equation in homology, yielding

$$[\Pi_{\rm tot}] = \sum_a N_a\,[\Pi_a] = 0 \quad \text{(A.31)}$$

As usual, cancellation of RR tadpoles in the underlying string theory configuration implies cancellation of four-dimensional chiral anomalies in the effective field theory in our configurations.

Let us provide one simple example, obtained by taking $\mathbf{X_6} = \mathbf{T^6}$, and a simple set of 3-cycles. We consider $\mathbf{X_6}$ to be a six-torus factorized as $\mathbf{T^6} = \mathbf{T^2}\times\mathbf{T^2}\times\mathbf{T^2}$. Also for simplicity we take the 3-cycles Π_a to be given by a factorized product of 1-cycles in each of the 2-tori. For a 3-cycle Π_a, the 1-cycle in the i^{th} 2-torus will be labeled by the numbers (n_a^i, m_a^i) it wraps along the horizontal and vertical directions, see figure A.9 for examples.

The intersection number is given by the product of the number of intersections in each 2-torus, and reads

$$I_{ab} = (n_a^1 m_b^1 - m_a^1 n_b^1)\times(n_a^2 m_b^2 - m_a^2 n_b^2)\times(n_a^3 m_b^3 - m_a^3 n_b^3) \quad \text{(A.32)}$$

To give one interesting example, consider a configuration of D6-branes on $\mathbf{T^6}$ defined by the following wrapping numbers

$N_1 = 3$	$(1,2)$	$(1,-1)$	(1,-2)
$N_2 = 2$	$(1,1)$	$(1,-2)$	(-1,5)
$N_3 = 1$	$(1,1)$	$(1,0)$	(-1,5)
$N_4 = 1$	$(1,2)$	$(-1,1)$	(1,1)
$N_5 = 1$	$(1,2)$	$(-1,1)$	(2,-7)
$N_6 = 1$	$(1,1)$	$(3,-4)$	(1,-5)

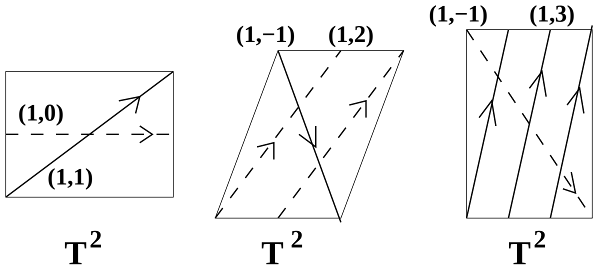


Figure A.9: Examples of intersecting 3-cycles in $\mathbf{T}^6$.

The intersection numbers are

$$\begin{array}{ccccc} I_{12}=3 & I_{13}=-3 & I_{14}=0 & I_{15}=0 & I_{16}=-3 \\ I_{23}=0 & I_{24}=6 & I_{25}=3 & I_{26}=0 & I_{34}=-6 \\ I_{35}=-3 & I_{36}=0 & I_{45}=0 & I_{46}=6 & I_{56}=3 \end{array}$$

A $U(1)$ linear combination, playing the role of hypercharge, remains massless

$$Q_Y = -\frac{1}{3}Q_1 - \frac{1}{2}Q_2 - Q_3 - Q_5 \tag{A.33}$$

The chiral fermion spectrum, with charges with respect to the Standard Model - like gauge group, is

$$\begin{gathered} SU(3)\times SU(2)\times U(1)_Y \times \dots \\ 3(3,2)_{1/6} + 3(\overline{3},1)_{-2/3} + 3(\overline{3},1)_{1/3} + 6(1,2)_{-1/2} + \\ +3(1,2)_{1/2} + 6(1,1)_1 + 3(1,1)_{-1} + 9(1,1)_0 \end{gathered} \tag{A.34}$$

Notice however, that the model contains additional $U(1)$ factors and other gauge factors, as well as matter beyond the context of the Standard Model.

In any event this general setup therefore allows the construction of a large class of models with 4d gravitational and non-abelian gauge anomalies, and charged chiral fermions. We leave their more detailed discussion for the interested reader (see [111] for a review) and simply point out that, although most models constructed in this setup are non-supersymmetric, there exist several explicit supersymmetric examples in the literature.

A.4 Final comments

The main message of this lecture is that there exist constructions in string and M-theory which have the potential of leading to low-energy physics very close to that observed in Nature. Perturbative heterotic string are simply one such setup, but there are others, like compactifications of Horava-Witten theory, or models with D-branes. There is life beyond perturbative heterotic theory!

The novelty about these new setups is that they have localized gauge sectors, and hence allow for fundamental scales not directly tied up to the 4d Planck scale, and can even be significantly lower then the latter. In models with a too low fundamental scale, there may be dangerous processes, like too fast proton decay. In many of the D-brane models above, there exist some symmetries which forbit this violation of baryon number.

The models are also interesting in that they provide an essentially new way to obtain gauge symmetries and chiral fermions in string theory. In particular this can be exploited to imagine new sources for the hierarchy of Yukawa couplings and fermion masses in the standard model.

Besides these novelties and successes, it is however important not to loose perspective and recognize that the models still leave many unanswered questions.

- If supersymmetry is present, how to break supersymmetry? If not, how to stabilize moduli at values that may correspond to (seemingly unnatural) large volumes?

- The moduli problem: Or how to get rid of the large number of massless scalars which exist in many compactifications in string theory (and whose vevs encode the parameters of the underlying geometry and gauge bundle (like sizes of the internal manifold, etc)).

- The vaccum degeneracy problem: Or the enormous amount of consistent vacua which can be constructed, out of which only one (if any at all) is realized in the real world. Is this model preferred by some energetic, cosmological, anthropic criterion? Or is it all just a matter of chance?

- The cosmological constant problem, which in general is too large once we break supersymmetry. Does string theory say anything new about this old problem?

As one can notice, the list is 'isomorphic' to the one we had in perturbative heterotic models. This means that certainly these are difficult problems which permeate any model buiding setup in string theory. Clearly we need better theoretical understanding of new aspects theory. This is not impossible, however, as for instance there are recent proposals to stabilize most compactification moduli by studying compactifications with non-trivial field strength fluxes for p-form fields [113]. Thus the above problems, which are central questions in string phenomenology, will hopefully solved perhaps by next-generation students like you!

Appendix B

Non-BPS D-branes in string theory

B.1 Motivation

In this lecture we present a new viewpoint on D-branes, arising from the Study of configurations of D-branes and anti-D-branes in string theory. The construction will imply some interesting insights into the meaning of tachyonic modes in string theory. Also, this viewpoint will lead to the construction of new stable non-BPS D-branes in string theory, which will allow to carry out a check of duality beyond supersymmetry. Some useful references for this talk are [114, 115].

B.2 Brane-antibrane pairs and tachyon condensation

B.2.1 Anti-D-branes

In analogy with particles and antiparticles in quantum field theory, every object in string theory has the corresponding antiobject, with equal tension but opposite charges. In particular, for every Dp-brane there exists a corresponding anti-Dp-brane state, denoted $\overline{Dp}$-brane, such that when they are put together they can annihilate each other into the vacuum.

$\overline{Dp}$-branes and Dp-branes have the same tension but opposite charges under the RR $(p+1)$-form. Note that this implies that $\overline{Dp}$-branes are also

BPS states, which preserve half of the supersymmetry of the vacuum, but they preserve the supersymmetries broken by the Dp-branes, and vice versa. Namely, the supersymmetry generators $\epsilon_L Q_L + \epsilon_R Q_R$ unbroken by the presence of these objects in type II theory, are of the form

$$\begin{aligned} Dp \quad & \longrightarrow \epsilon_L = \Gamma^0 \ldots \Gamma p \epsilon_R \\ \overline{Dp} \quad & \longrightarrow \epsilon_L = -\Gamma^0 \ldots \Gamma^p \epsilon_R \end{aligned} \tag{B.1}$$

$\overline{Dp}$-branes are described, just as Dp-branes, as $(p+1)$-dimensional subspaces on which open strings are allowed to end. It is thus natural to consider what features distinguish Dp-branes and $\overline{Dp}$-branes , from the viewpoint of the 2d worldsheet. Equivalently, considering a configuration including both kinds of objects, what distinguishes open strings with both ends on the same kind of object, and open strings starting on branes and ending on antibranes (or viceversa). This is addressed in the following section.

B.2.2 Dp-$\overline{Dp}$-brane pair

Consider a configuration with a single Dp- and a single $\overline{Dp}$-brane in type II theory, with coincident worldvolumes along the directions $01\ldots p$. A prominent feature of this configuration is that it is non-supersymmetric. Namely there is no supercharge which is preserved by both the Dp- and the $\overline{Dp}$-brane. Another way to obtain the result is to notice that the state as a whole is not BPS: denoting T_p, Q_p the tension and charge of a Dp-brane, the state as a whole has tension $2T_p$ and charge 0. The tension of a BPS state in the topological sector of zero charge should be zero, hence the brane-antibrane pair is a non-BPS excited state. Notice that there is a clear BPS state in the zero charge sector of the theory, namely the type II

vacuum. Therefore we expect the non-BPS state given by the brane-antibrane pair to be unstable against decay to the vacuum, since both states have the same charges, and the vacuum is energetically favoured.

Let us compute the spectrum of open strings in the presence of the brane-antibrane pair. Clearly open strings with both ends on the Dp-brane (Dp-Dp strings) are not sensitive to the presence of the $\overline{Dp}$-brane, hence are quantized as usual. They lead to a $(p+1)$-dimensional $U(1)$ gauge boson and their superpartners with respect to the 16 supersymmetries unbroken by the Dp-brane. Similarly, $\overline{Dp}$-$\overline{Dp}$ open strings lead to a $(p+1)$-dimensional $U(1)$ gauge boson and its superpartners with respect to the 16 supersymmetries

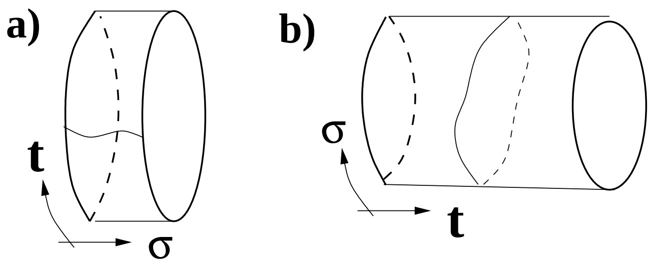


Figure B.1: The annulus diagram regarded in the open and in the closed string channel.

unbroken by the $\overline{Dp}$-branes (and which are the opposite of the above ones). Finally, we need to consider Dp-$\overline{Dp}$ and $\overline{Dp}$-Dp open strings. The boundary conditions are exactly the same as for the above sectors, namely Neumann for the directions $0, \ldots, p$ and Dirichlet for the directions $p+1, \ldots, 9$. Hence, the Hilbert space of open string states, before any GSO projection, is the usual one. The lightest modes are

Sector	State	αM^2	Field
NS	$\|0\rangle$	-1	Scalars
	$\psi^{\mu}_{-1/2}\|0\rangle$	0	Gauge bosons + Scalars
R	8_C	0	Fermions
	8_S	0	Fermions

We now show that open-closed duality forces to choose the GSO projection in the Dp-$\overline{Dp}$ sector opposite to the usual one (namely, that in Dp-Dp or $\overline{Dp}$-$\overline{Dp}$ sector). To see that, consider the annulus diagram, with two boundaries on Dp-branes, see figure B.1a. Computing this amplitude in the open string channel, as a loop of Dp-Dp strings, we get

$$
\begin{aligned}
Z(T)_{pp} \;&\simeq\; \frac{1}{2}\left(\mathrm{tr}_{NS} q^{N_F+E_0^F} + \mathrm{tr}_{NS}(q^{N_F+E_0^F}(-)^F\right) - \frac{1}{2}\left(\mathrm{tr}_{R} q^{N_F+E_0^F} + \mathrm{tr}_{R}(q^{N_F+E_0^F}(-)^F\right) = \\
&= \frac{1}{2}\eta^{-4}\left(\vartheta\begin{bmatrix}0\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^4\right) - \frac{1}{2}\eta^{-4}\left(\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^4 - \vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^4\right) \qquad \text{(B.2)}
\end{aligned}
$$

This quantity can be rewritten as an amplitude of a closed string propagating for a time $T' = 1/(2T)$, by performing a modular transformation, leading to

$$Z(2T')_{pp} = \eta^{-4}\left(\vartheta\left[\begin{array}{c}0\\0\end{array}\right]^4 - \vartheta\left[\begin{array}{c}1/2\\0\end{array}\right]^4\right) - \eta^{-4}\left(\vartheta\left[\begin{array}{c}0\\1/2\end{array}\right]^4 - \vartheta\left[\begin{array}{c}1/2\\1/2\end{array}\right]^4\right) \quad \text{(B.3)}$$

The amplitude in the closed channel describes the interaction between Dp-branes, via exchange of NSNS and RR fields, see figure B.1b).

The amplitude for a Dp-$\overline{Dp}$-brane annulus, in the closed channel, differs from (B.3) in the sign of the terms corresponding to the exchange of RR fields. This is because of the opposite sign of the RR charge of the $\overline{Dp}$-brane with respect to the Dp-brane charge. Hence we obtain

$$Z(2T')_{p\overline{p}} = \eta^{-4}\left(\vartheta\left[\begin{array}{c}0\\0\end{array}\right]^4 + \vartheta\left[\begin{array}{c}1/2\\0\end{array}\right]^4\right) - \eta^{-4}\left(\vartheta\left[\begin{array}{c}0\\1/2\end{array}\right]^4 + \vartheta\left[\begin{array}{c}1/2\\1/2\end{array}\right]^4\right)$$

Going back to the open string channel, the annulus amplitude for Dp-$\overline{Dp}$ open strings going in a loop corresponds to

$$Z(T)_{pp} \simeq \frac{1}{2}\Big(\mathrm{tr}_{NS}\, q^{N_F+E_0^F} - \mathrm{tr}_{NS}(q^{N_F+E_0^F}(-)^F\Big) - \frac{1}{2}\Big(\mathrm{tr}_{R}\, q^{N_F+E_0^F} - \mathrm{tr}_{R}(q^{N_F+E_0^F}(-)$$

Hence we see that the signs imposing the GSO projection are flipped. Therefore, for Dp-$\overline{Dp}$ and $\overline{Dp}$-Dp open strings, the lightest modes are

Sector	State	αM^2	Field
NS	$\|0\rangle$	-1	Scalars
R	8_S	0	Fermions

These fields carry charges $\pm(+1,-1)$ under the $U(1)\times U(1)$ on the D- and anti-D-branes. The spectrum in these sectors is very different from the Dp-Dp and $\overline{Dp}$-$\overline{Dp}$ sectors. In particular, there is no enhanced gauge symmetry when branes and antibranes coincide. Note also that these sectors lead to a complex tachyon of the world-volume. We will discuss it in detail later on. Before that, let us simply mention that for branes and antibranes separated by a distance L in transverse space, the lightest mode in the NS sector has mass

$$M^2 = -\frac{1}{\alpha} + \frac{L^2}{(2\pi\alpha)^2} \quad \text{(B.4)}$$

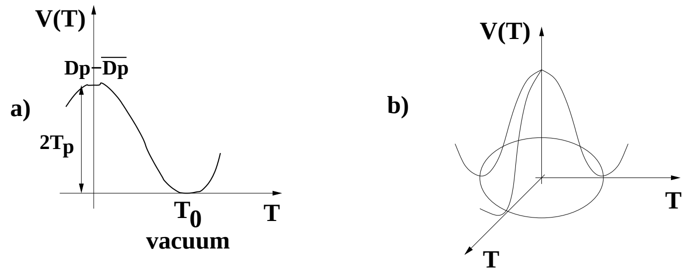


Figure B.2: Two pictures of the tachyon potential for the brane-antibrane system.

The tachyonic mode develops for distances smaller than a critical distance of the order of the string scale, $L_c \leq 4\pi^2\alpha^{1/2}$. However, branes and antibranes initially at a distance larger than L_c feel a mutual attraction (since they have equal tension and opposite charge, hence attract both gravitationally and by RR Coulomb interactions), and tend dynamically to approach and decrease this distance until they reach the tachyonic regime.

B.2.3 Tachyon condensation

The meaning of this tachyon is that the configuration is unstable against annihilation of the brane-antibrane pair to the vacuum of type II theory. That is, the brane-antibrane pair corresponds to a configuration of the system which is sitting at the top of a potential. The negative mass square of the tachyon field simply means that the second derivative of the potential as a function of this field is negative at the top of the potential, see figure B.2. One should therefore let the tachyon roll down to the minimum of the potential, if it exists, to obtain a stable configuration. This process, by which the tachyon field acquires a vacuum expectation value T_0, minimizing the potential, is known as tachyon condensation.

A remarkable feature of this process is that there is a clear spacetime picture of its endpoint. The tachyon simply represents the instability of the brane-antibrane system against annihilation to the vacuum. This spacetime picture of the process of tachyon condensation implies that we know exactly the final state of this process: it is the vacuum of type II theory. So, although the initial state is non-supersymmetric, we can make exact statement about its fate after tachyon condensation..

Note that from the viewpoint of the world-volume theory, this process is similar in some respects to a Higgs mechanism. This is not completely precise, though. It is true that in the process a charged field (the tachyon) gets a vev, and breaks a gauge symmetry (the antisymmetric linear combination of the $U(1)$s on the D- and the anti-D-branes). However, the final state is the vacuum, where no open string states, and hence also the diagonal $U(1)$ linear combination, under which the tachyon is uncharged should also disappear. More strikingly, in the final state all open string modes of the initial state must be absent. Hence in the process of tachyon condensation an infinite number of fields disappear from the theory. This kind of processes have been successfully described only within the approach of string field theory.

Let us emphasize how remarkable it is the fact that we exactly know the final state of tachyon condensation. It leads to a number of exact statements about the properties of a non-supersymmetric brane-antibrane pair when the world-volume tachyon has a constant vev T_0. All of them are encoded in the statement that a brane-antibrane pair with a tachyon vev T_0 is indistinguishable from the vacuum. This is very surprising, for instance, the final state has an enhanced supersymmetry, it has zero energy, etc. The set of predictions (as well as several others to be studied later) following from this spacetime picture of tachyon condensation are known as Sens conjectures.

Also, very remarkably, we have succeeded in understanding the meaning of open string tachyons. In fact, we can extend this understanding to othe open string tachyons in string theory. For instance, tachyons in the open string sector of open bosonic string theory are now understood as an instability of bosonic D-branes to decay into the vacuum. This is consistent, since bosonic D-branes do not carry any conserved charge. Hence, we are recovering the result, briefly mentioned in the lecture on open strings, that open bosonic string theory is unstable into decay to purely closed bosonic string theory, with no open string sector at all.

B.3 D-branes from brane-antibrane pairs

In this section we discuss other processes of tachyon condensation in brane-antibrane systems, where the final state is not the vacuum, but a lower-dimensional D-brane.

B.3.1 Branes within branes

For this section, see [116]. Recall that a Dp-brane in charged not only under the RR $(p+1)$-form, but also Under other lower-degree RR forms, if the world-volume gauge bundle is non-trivial. For instance, consider the Wess-Zumino couplings for a D3-brane

$$S_{WZ} = \int_{D3} C_4 + \int_{D3} C_2 \wedge \mathrm{tr}\, F + \int_{D3} C_0 \wedge \mathrm{tr}\, F^2 \qquad \text{(B.5)}$$

Consider a world-volume gauge bundle with non-zero first Chern class, i.e. $\mathrm{tr}\, F$ is non-trivial on the D3-brane world-volume. This intuitively corresponds to turning on a magnetic field along two of the directions, say 23, in the D3-brane volume, with total integral e.g. $\int_{23} F = 1$. The above couplings imply that the D3-brane is charged under the RR 2-form C_2, or that we are dealing with a bound state of a D3-brane and a D1-brane (with volume along 01). In a sense, the system can be thought of as a D3-brane with a D1-brane diluted in its volume [1].

Similarly, a non-trivial $\mathrm{tr}\, F$ on a general Dp-brane induces D$(p-2)$-brane charge, a non-trivial second Chern class (or instanton number) $\mathrm{tr}\, F^2$ induces non-trivial D$(p-4)$-brane charge, etc.

B.3.2 D-branes from brane-antibrane pairs

Consider a $\overline{D3}$-brane with trivial world-volume gauge bundle, and a D3-brane with one unit of induced D1-brane charge, see figure B.3. The complete system has zero D3-brane charge, one unit of D1-brane charge, and non-zero 3-brane tension (slightly larger than but around $2T_3$).

Clearly the state is non-supersymmetric. One way to understand this is to note that there exists a state with the same charges and much less energy, namely a BPS D1-brane. Hence we expect, from the spacetime viewpoint, that the initial system is unstable to decay into a D1-brane state. Notice that decay to the vacuum is not consistent with charge conservation. Heuristically, the decay to the D1-brane state can be understood by considering the magnetic field to be localized in a more or less compact core in the directions 23, and translationally invariant along 01. Asymptotically away from the

[1] Indeed this is quite precise. Starting with a configuration of coincident D3- and D1-branes there is a dynamical process diluting the D1-brane as world-volume gauge field strength on the D3-brane.

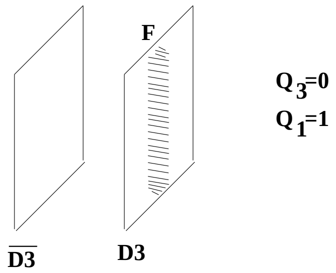


Figure B.3: Brane-antibrane system with induced lower-dimensional brane charges.

core, we just have a D3-$\overline{D3}$-brane pair, with no magnetic field density, so the system will suffer tachyon condensation annihilating them in the asymptotic region. Near the core, the magnetic field changes things, and annihilation leads an object compactly supported in 23, namely the D1-brane.

From the viewpoint of the 4d world-volume, the above system is described as follows. In a D3-$\overline{D3}$-system, we have a gauge group $U(1)^2$, and a complex scalar T with charge $(+1,-1)$, with a Mexican hat potential shown in figure B.2b. Note that gauge invariance implies that the potential is function of the modulus of T, $V(|T|)$. The diagonal $U(1)$ subgroup decouples and will be irrelevant for the following discussion. This field theory has soliton solutions, which correspond to topologically non-trivial world-volume field configurations. Finite energy solitons must have a tachyon field asymptoting to the value $|T| = T_0$. Considering configurations which are translationally invariant in 01, $T = T(x^2, x^3)$, the tachyon field taken at the $\mathbf{S^1}$ at infinity in 23 defines a map from the spacetime $\mathbf{S^1}$ to the set of minima of the potential, which is also an $\mathbf{S^1}$. Topologically inequivalent solitons correspond to topologically inequivalent tachyon field configurations, which correspond to topologically inequivalent maps $\mathbf{S^1} \to \mathbf{S^1}$. The latter are classified by the homotopy group $\Pi_1(\mathbf{S^1}) = \mathbf{Z}$, i.e. there are an infinite number of inequivalent solitons, characterized by an integer, known as winding number of the above map. A simple example is provided by the winding one configuration. Defining $z = x^2 + ix^3$, the tachyon profile for the corresponding soliton is

$$T(z) = T_0 \frac{z}{|z|} \tag{B.6}$$

Representing the complex value of the tachyon by an arrow, the field config-

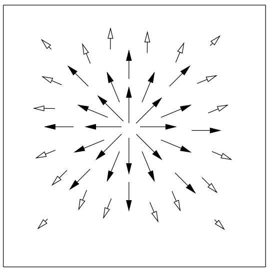


Figure B.4: Picture of the hedgehog configuration for the tachyon field in the vortex solution.

uration is of the hedgehog form shown in figure B.4. In order to have a finite energy configuration, we also need to turn on a non-trivial gauge field, so that the covariant derivatives approach zero fast enough as $|z| \to \infty$. This gauge field is such that there is a non-trivial first Chern class over 23, $\int_{23} F = 1$. The whole field configuration is known as vortex, and is the world-volume description of the tachyon condensed configuration. Indeed, asymptotically the system approaches the configuration of a D3-$\overline{D3}$-brane with a tachyon vev of T_0, hence describing asymptotic annihilation. Near the core, the tachyon value is approximately zero, and no annihilation is implied. In fact, near the core we have a D3-$\overline{D3}$ system with non-condensed tachyon; hence, open strings are allowed to end in the near core region of the above system. The system described an object localized in 23, charged under C_2 and on which open strings can end. This is clearly a D1-brane, which we have construced as a bound state of a higher-dimensional brane-antibrane pair.

The above construction suggests the construction of D-branes as bound states, upon tachyon condenstation, of higher dimensional brane-antibrane pairs. This is a surprising new viewpoint, where D-branes are regarded as solitons on the world-volume of brane-antibrane pairs.

We can use a similar strategy to construct other D-brane states, in particular unstable Non-BPS Dp-branes in type II theory (with p even for IIB and odd for IIA). For instance, consider a D3-$\overline{D3}$ pair, with a tachyon profile corresponding to a kink in one dimension, say 3, see figure B.5. The field configuration is localized in a compact region in x^3, and has trivial field strengths.

This world-volume configuration is not topologically stable, the kind can be continuously unwound into a trivial configuration. This implies that the

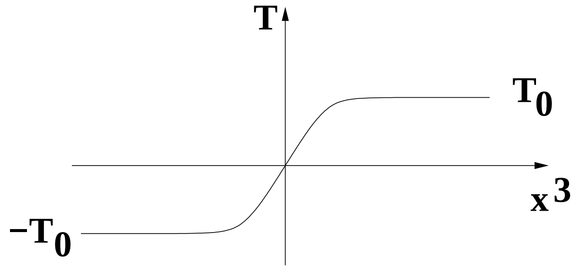


Figure B.5: Picture of the tachyon profile in the kink configuration..

resulting D2-brane, denoted $\widehat{D2}$-brane, is unstable (against decay to the vacuum), which is consistent since it carries no conserved charges. We would like to mention two further facts on these non-BPS branes: First, they admit a microscopic description, as subspaces on which open strings end. In this situation, the fact that these D-branes do not carry RR charges implies, by open-closed duality, that open strings stretching between non-BPS D-branes of this kind have a world-volume spectrum with no GSO projection. This spectrum is easily obtained, and in particular contains a real tachyon. Second, a further kink configuration on this world-volume tachyon corresponds to the condensation of an unstable non-BPS Dp-brane to a BPS D$(p-1)$-brane, of the usual kind (these relations are known as descent relations).

B.4 D-branes and K-theory

Let us generalize the idea that D-branes are constructed as bound states of higher-dimensional brane-antibrane pairs, upon tachyon condensation. The latter statement means that, at the topological level, states which differ by processes of creation/annihilation of brane-antibrane pairs must be considered equivalent.

Let us apply these ideas to type IIB theory on a spacetime X, and try to classify all D-brane charges. Namely, consider a type IIB configuration with n D9-$\overline{D9}$-brane pairs. Note that this is consistent, since the tadpole for the RR 10-form C_{10} generated by the D9- and the $\overline{D9}$-branes cancel each other [2]. In general, the D9-branes carry a world-volume $U(n)$ gauge bundle E, and

[2]Recall that the only inconsistency in coupling 10d Poincare invariant open string sector to oriented type IIB theory arose from RR tadpole cancellation. Note also that

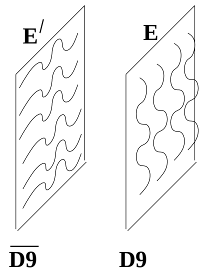


Figure B.6: Brane-antibrane pairs with general world-volume gauge bundles.

the $\overline{D9}$-branes carry another $U(n)$ gauge bundle E. Hence, different D-brane states or D-brane charges are classified by pairs of bundles (E, E). However, configurations that differ by the nucleation of D9- and $\overline{D9}$-branes, both with world-volume gauge bundle H, must be considered topologically equivalent. Therefore the set of topologically inequivalent D-brane states is given by the set of pairs of bundles (E, E), modulo the equivalence relation

$$(E, E) \quad \simeq \quad (E \oplus H, E \oplus H) \tag{B.7}$$

The set of pairs of topological bundles up to this equivalence relation is a finitely generated group, known as (complex) K-theory group of the spacetime X, denoted $K(X)$.

Let us describe the classification of type IIB charged D-branes in flat 10d space from this viewpoint. If we are interested in p-brane states (i.e. states with Poincare invariance in $(p+1)$ dimensions)the bundles are non-trivial only over the $(9-p)$-dimensional transverse space. Also, we are interested in bundles with compact support, so that the resulting states are localized in $\mathbf{R^{9-p}}$. See figure B.6. Bundles with compact support on $\mathbf{R9-p}$ can be described as general bundles over $\mathbf{S^{9-p}}$. The corresponding K-theory groups have been computed by mathematicians and read

$$\begin{aligned} K(\mathbf{S^{9-p}}) &= \mathbf{Z} \quad \mathbf{p}\text{ odd} \\ &= 1 \qquad p\text{ even} \end{aligned} \tag{B.8}$$

Hence type IIB theory contains stable Dp-branes for p odd. These branes are stable since their charge, classified by the K-theory class, which is topological,

---

the configuration with equal number of D9 and $\overline{D9}$-branes is regarded not as a new string theory, but as an excited state of type IIB theory (connected to the vacuum via tachyon condensation).

forbids their decay to the vacuum. The fact that the K-theory classes are **Z** valued implies that their charge is additive. In fact, these are the familiar BPS we already know about, and their charge is the charge under the RR $(p+1)$-form field. Hence the classification of D-brane states using K-theory agrees with the classification using cohomology (namely computing the charge of a D-brane state as the flux of a certain form field over a cycle surrounding the D-brane) [3] . This is not so surprising, since there is a natural map from K-theory to cohomology, which to each K-theory class represented by a pair of bundles (E, E) it assigns the cohomology class

$$(E, E) \longrightarrow \mathrm{ch}(E) - \mathrm{ch}(E) \tag{B.9}$$

where $\mathrm{ch}(E)$ is the Chern character, defined by

$$\mathrm{ch}(E) = \mathrm{tr}\ e^{F/2\pi} = 1 + \frac{1}{2\pi}\mathrm{tr}\, F + \frac{1}{8\pi^2}\,\mathrm{tr}\, F^2 + \ldots \tag{B.10}$$

The Chern character is additive, $\mathrm{ch}(E\oplus) = \mathrm{ch}(E) - \mathrm{ch}(H)$, hence the above map is independenent of the representative of the K-theory class. Finally, notice that the Chern character enters in the Wess-Zumino couplings on the world-volume of a D-brane, hence it carries the information on the induced D-brane charges under the RR p-form fields.

There are situations, however, where the above mapping is not injective. Namely there may are situations where there exist non-trivial K-theory classes whose Chern character vanishes. Namely, there exist bundles which are topologically non-trivial but whose characteristic classes all vanish. This implies the existence of D-branes which are stable (since they carry a non-trivial topological quantum number) but are uncharged under the RR fields.

The simplest example of this kind is provided by type I theory. The classification of D-branes can be carried out as above. Namely, introduce n additional D9-$\overline{D9}$ over the vacuum of type I theory (note that this is consistent, since the total system contains $(n + 32)$ D9-branes and n $\overline{D9}$-branes, leading to a total RR 10-form tadpole cancelling that of the O9-planes). The D9- and $\overline{D9}$-branes carry $SO(n)$ bundles E, E. D-brane configurations are topologically classified by pairs of bundles (E, E) modulo the equivalence

[3]Notice that we are classifying topological D-brane states. In particular, the unstable D-branes of type IIB theory do not appear in this classification since they can decay to the vacuum, namely are topologically equivalent to it.

relation (B.7). The resulting set is a group known as the real K-theory group of the spacetime X $KO(X)$.

Let us classify type I D-brane charges in flat space. As before, we need to compute the groups $KO(\mathbf{S^{9-p}}$, which have been computed by mathematicians. We obtain the following sets of D-branes

$$\begin{array}{rccl} KO(\mathbf{S^1}) & = \mathbf{Z_2} & \rightarrow & \widehat{D8} \\ KO(\mathbf{S^2}) & = \mathbf{Z_2} & \rightarrow & \widehat{D7} \\ KO(\mathbf{S^4}) & = \mathbf{Z} & \rightarrow & D5 \\ KO(\mathbf{S^8}) & = \mathbf{Z} & \rightarrow & D1 \\ KO(\mathbf{S^9}) & = \mathbf{Z_2} & \rightarrow & \widehat{D0} \\ KO(\mathbf{S^{10}}) & = \mathbf{Z_2} & \rightarrow & D\widehat{(-1)} \end{array} \tag{B.11}$$

Beyond the familiar BPS D1- and D5-branes, the K-theory classification implies the existence of non-BPS D8-, D7-, D0- and D(-1)-brane charges. They are completely uncharged under the RR fields, however they carry a non-trivial $\mathbf{Z_2}$ charge and cannot decay into the vacuum.

We would like to conclude with some comments

• For type IIA theory, there also exists a K-theory classification of D-brane charges. It is based on classifying bundles over spacetime filling unstable $\widehat{D}9$-branes. The relevant K-theory groups are known as (complex) reduced K-theory groups $K^{-1}(X)$. There is a relation between these and the type IIB groups, which is consistent with T-duality. For instance in 10d space, we have $K^{-1}(\mathbf{S^n}) = \mathbf{K}(\mathbf{S^{n-1}})$. This leads to the familiar set of BPS states of type IIA theory.

• The above construction is valid for D-branes, since they naturally carry world-volume gauge bundles. It is still an open issue to extend this kind of classification scheme to other theories without D-branes, like heterotic theories or M-theory, and to other objects, like NS5-branes.

B.5 Type I non-BPS D-branes

We have seen that type I contains different non-BPS D-branes with non-trivial topological charge. Since these charges are topological, states with these charges must exist in the spectrum for all values of the moduli (although their microscopic description may change in between). This allows to test string duality for this particular class of non-BPS states, i.e. test string duality beyond supersymmetry [4]

B.5.1 Description

The $\widehat{D0}$-brane

Altough we have described it starting with D9-$\overline{D9}$-brane pairs, the simplest construction starts from a D1-$\overline{D1}$-brane pair in type I theory. The world-volume gauge group is $\mathbf{Z_2} \times \mathbf{Z_2}$. The fact that this gauge group is discrete ensures that a kink configuration for the world-volume tachyon cannot be unwound, and hence describes tachyon condensation to a stable state. This is the stable $\widehat{D0}$-brane of type I theory [5]. The fact that it carries a $\mathbf{Z_2}$ charge means that two of these states can annihilate to the vacuum. This is understood in the D1-$\overline{D1}$-brane pair because two kinks can unwind to the trivial configuration for the world-volume tachyon, describing decay to the vacuum.

There is a microscopic description for the type I $\widehat{D0}$-brane, as a 1d subspace on which open strings can end. Such open strings have no GSO projection, in agreement (via open-closed duality) with the fact that they carry no RR charges. The light spectrum on the world-volume of a stack of n $\widehat{D0}$-branes is as follows. In the 00 sector, there is no GSO projection. The states are computed as usual (with some subtlety due to the fact that

there are not enough NN directions to use the light-cone gauge), and projected into Ω-invariant states. In the NS sector, there are massless $SO(n)$ gauge bosons, and 9 scalars in the representation $\Box\!\Box$; there are also a world-volume real tachyon, transforming in the representation $\begin{smallmatrix}\Box\\\Box\end{smallmatrix}$, so it is absent for $n=1$ (in which case the system is stable). In the R sector, the groundstates give rise to fermions in the $\Box\!\Box + \begin{smallmatrix}\Box\\\Box\end{smallmatrix}$ representation. In the $09+90$ sector, the NS states are massive, while in the R sector the groundstate gives rise to massless 1d fermions in the 32 of the D9-brane $SO(32)$ group.

For $n=1$ we have a stable particle, with worldline described by the above fields. It has nine worldline bosons, so the particle propatates in 10d.

---

[4] Note that the lack of the BPS property however implies that we do not have much control over properties like the tension of the object, as the moduli change. Hence the tests are much less exhaustive than for supersymmetric states.

[5] Equivalently, one can describe the type I $\widehat{D0}$-brane as the type IIB $\widehat{D0}$-brane, modded out by Ω, which projects out the world-volume tachyon of the latter.

On the other hand, there are worldline zero modes, which imply that in the quantum theory the particle belongs to a multiplet. Quantization of fermion zero modes in the 00 sector gives a 256-fold multiplicity, implying the particle state belongs to a non-BPS multiplet. Quantization of fermion zero modes in the 09 sector imply that the particle transforms in a non-trivial representation of $SO(32)$, in particular a 2^{15}-fold dimensional chiral spinor representation (there also appears the spinor representation of opposite chirality, but it is eliminated by the world-volume $\mathbf{Z_2}$ gauge group).

The fact that type I contains states in the spinor representation of a given chirality implies that its spacetime gauge group is globally not $SO(32)$. All perturbative and non-perturbative states are consistent with a gauge group $Spin(32)/Z_2$ (where $Spin$ allows the existence of spinor representations, and $\mathbf{Z_2}$ forbids the existence of spinors of one chirality and of states in vector representation).

The $\widehat{D8}$-brane

There exists a type I $\widehat{D8}$-brane described microscopically as a 9d subspace on which open strings end, and which carries the correct K-theory charge. However, the brane contains a world-volume tachyon arising in the sector of open strings stretching from the brane to the background D9-branes. This tachyon implies the non-BPS brane is unstable to decay, but not to the vacuum (which is forbidden by charge conservation) but to a different configuration carrying the same charge. The latter configuration is a non-trivial bundle on the D9-branes where there is a $\mathbf{Z_2}$ Wilson line on one of the D9-branes [6]

The $\widehat{D(-1)}$-brane

The $\widehat{D(-1)}$-brane can be constructed starting from a type I D1-$\overline{D1}$-brane pair with a vortex configuration for the world-volume tachyon. Equivalently it can be described as a D(−1)-$\overline{D(-1)}$-brane pair of type IIB theory, modded out by Ω, which exchanges the D(−1) and the $\overline{D(-1)}$-brane, and eliminates the world-volume tachyon. The latter description provides a simple microscopic description for the type I D-instanton, but we will skip its detailed discussion.

The $\widehat{D7}$-brane

[6]This Wilson line is topological in the sense that it is an element in $Spin(32)/\mathbf{Z_2}$ but not of $SO(32)$.

This can be described as a type IIB D7-$\overline{D7}$-brane pair, modded out by Ω, which exchanges the objects in the pair, and eliminates the world-volume tachyon in the $7\overline{7}$ sector. The tachyon in the 79 and $\overline{7}9$ sectors however survive, implying that the system is unstable against decay, not to the vacuum but to a non-trivial bundle on the background D9-branes. The bundle is described by two $\mathbf{Z_2}$ Wilson lines which commute up to a sign in $SO(32)$, namely which commute in $Spin(32)/\mathbf{Z_2}$ but not in $SO(32)$.

B.5.2 Heterotic/type I duality beyond supersymmetry

Non-BPS states in type I theory, which are nevertheless stable due to charge conservation, must exist not only at weak coupling (where we have provided a microscopic description), but at all values of the coupling. This implies that they lead to results which can be extrapolated to strong coupling, and be compared with properties of the heterotic theory.

The perturbative group of type I theory is $O(32)$. However we have seen that the global structure of the group is $Spin(32)/\mathbf{Z_2}$, since the theory contains states that transform in a chiral spinor representation, and states described by gauge configurations which do not exist in $SO(32)$. Finally, the non-BPS D-instanton plays a crucial role in describing the change in gauge group. Namely, it is not invariant under large gauge transformations in $SO(32)$. The true gauge group consistent with all non-BPS states of the theory is in fact the appropriate one to agree with the heterotic theory upon type I / heterotic duality. This is a first non-trivial result of non-BPS type I theory.

An even more remarkable check is that the particles described as type I $\widehat{D0}$-branes provide states that transform in a chiral spinor representation of the spacetime gauge group. By duality, heterotic theory should contain some states with the same basic properties, namely same gauge representation, and same non-BPS supermultiplet. Indeed, the $Spin(32)/\mathbf{Z_2}$ heterotic theory does contain states with these properties, they are given by massive perturbative heterotic states, with left-handed internal 16d momentum

$$P = \frac{1}{2}(\pm, \ldots, \pm) \quad , \quad \# - = \text{even} \tag{B.12}$$

This is Sens great idea on using these states to test string duality beyond supersymmetry.

B.6 Final comments

In this lecture we have studied a beautiful set of ideas, concerning a new viewpoint on D-branes. They have been widely generalized to more involved configurations, like orbifolds and orientifolds.

The construction of D-branes as bound states of higher-dimensional brane-antibrane pairs has allowed us to make precise exact statements on tachyon condensation processes in non-supersymmetric systems. Finally, these results have provided a new tool to test and partially confirm string duality beyond supersymmetry.

There are many other applications of these ideas to other related contexts. For instance the study of condensation of tachyons of other kinds (most interestingly the study of closed string tacyons is an open issue), or the application of antibranes and non-BPS D-branes as a source of supersymmetry breaking in model building. We leave these questions for the interested reader.

Appendix A

Modular functions

There is a lot of mathematical literature on modular functions, namely functions of the parameter τ which have nice transformation properties under the $SL(2,\mathbf{Z})$ modular group. A useful reference for them is [40].

Recall that the modular group is the set of transformations

$$\tau \to \frac{a\tau+b}{c\tau+d} \quad \text{with } a,b,c,d \in \mathbf{Z} \quad \text{and } \mathbf{ad-bc=1} \tag{A.1}$$

and is generated by $\tau \to \tau+1$, $\tau \to -1/\tau$

The Dedekind eta function

Introduce $q = e^{2\pi i\tau}$.The Dedekind eta function is defined by

$$\eta(\tau) \,=\, q^{1/24} \prod_{n=1}^{\infty} (1-q^n) \tag{A.2}$$

Under modular transformations

$$\begin{aligned} \eta(\tau+1) &= e^{\pi i/12}\,\eta(\tau) \\ \eta(-1/\tau) &= (-i\tau)^{1/2}\,\eta(\tau) \end{aligned} \tag{A.3}$$

(The first is trivial to show, while the second is tricky and one should consult the literature).

The theta functions

For future use it is useful to introduce the theta function with characteristics θ, ϕ

$$\vartheta\left[\begin{array}{c}\theta\\ \phi\end{array}\right](\tau) \,=\, \eta(\tau)\, e^{2\pi i\,\theta\phi}\, q^{\frac{1}{2}\theta^2-\frac{1}{24}} \prod_{n=1}^{\infty} (1+q^{n+\theta-1/2}\, e^{2\pi i\phi})\,(1+q^{n-\theta-1/2}\, e^{-2\pi i\phi}) \tag{A.4}$$

These functions also have an expression as infinite sums

$$\vartheta\left[\begin{array}{c}\theta\\ \phi\end{array}\right](\tau) \,=\, \sum_{n\in\mathbf{Z}} q^{(n+\theta)^2/2}\, e^{2\pi i\,(n+\theta)\phi} \tag{A.5}$$

The fact that (A.4) and (A.5) are equal is related to *bosonization*, namely the fact that in two dimensions a theory of free fermions can be rewritten as a theory of free bosons (with a compact target space). The two expressions for the theta functions correspond to the partition functions of the same theory in terms of different field variables. This will be understood better when we study 2d theories with fermions in the superstring.

Some useful and often appearing values of the characteristics are 0, 1/2. For future convenience, we list the product form of the corresponding theta functions

$$\begin{aligned}
\vartheta\left[\begin{array}{c}0\\ 0\end{array}\right](\tau) &= \prod_{n=1}^{\infty}(1-q^n)\,(1+q^{n-1/2})^2\\
\vartheta\left[\begin{array}{c}0\\ 1/2\end{array}\right](\tau) &= \prod_{n=1}^{\infty}(1-q^n)\,(1-q^{n-1/2})^2\\
\vartheta\left[\begin{array}{c}1/2\\ 0\end{array}\right](\tau) &= q^{1/8}\prod_{n=1}^{\infty}(1-q^n)\,(1+q^n)\,(1+q^{n-1}) \,=\\
&= 2\,q^{1/8}\prod_{n=1}^{\infty}(1-q^n)\,(1+q^n)^2\\
\vartheta\left[\begin{array}{c}1/2\\ 1/2\end{array}\right](\tau) &= i\,q^{1/8}\prod_{n=1}^{\infty}(1-q^n)^2\,(1-q^{n-1}) \,=\, 0
\end{aligned} \tag{A.6}$$

Finally, we list some useful properties of theta functions. Under integer shifts of the characteristics

$$\vartheta\left[\begin{array}{c}\theta+m\\ \phi+n\end{array}\right](\tau) \,=\, e^{2\pi i\theta n}\,\vartheta\left[\begin{array}{c}\theta\\ \phi\end{array}\right](\tau) \tag{A.7}$$

This can be shown very easily using the infinite sum form (A.5).

Under modular transformations

$$\begin{aligned}
\vartheta\left[\begin{array}{c}\theta\\ \phi\end{array}\right](\tau+1) &= e^{-\pi i(\theta^2-\theta)}\,\vartheta\left[\begin{array}{c}\theta\\ \theta+\phi-1/2\end{array}\right](\tau)\\
\vartheta\left[\begin{array}{c}\theta\\ \phi\end{array}\right](-1/\tau) &= (-i\tau)^{1/2}\,\vartheta\left[\begin{array}{c}\phi\\ -\theta\end{array}\right](\tau)
\end{aligned} \tag{A.8}$$

The first is very easy to show, using the infinite sum form (A.5) and using the trick that $e^{\pi i n^2} = e^{\pi i n}$ (since $n^2 = n$ mod 2). The second is also easy in the infinite sum form using the Poisson resummation formula

$$\sum_{n\in\mathbf{Z}} \exp\left[-\pi A(n+\theta)^2 + 2\pi i\,(n+\theta)\phi\right] \;=\; A^{-1/2} \sum_{k\in\mathbf{Z}} \exp\left[-\pi A^{-1}(k+\phi)^2 - 2\pi i k\theta\right] \quad \text{(A.9)}$$

This general formula can be shown by repeatedly using the one-dimensional one (A.9).

Particular cases of this transformation are

$$\vartheta\begin{bmatrix}0\\0\end{bmatrix}(\tau+1) = \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}(\tau) \quad ; \quad \vartheta\begin{bmatrix}0\\0\end{bmatrix}(-1/\tau) = (-i\tau)^{1/2}\,\vartheta\begin{bmatrix}0\\0\end{bmatrix}(\tau)$$

$$\vartheta\begin{bmatrix}0\\1/2\end{bmatrix}(\tau+1) = \vartheta\begin{bmatrix}0\\0\end{bmatrix}(\tau) \quad ; \quad \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}(-1/\tau) = (-i\tau)^{1/2}\,\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}(\tau)$$

$$\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}(\tau+1) = e^{-\pi i/4}\,\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}(\tau) \quad ; \quad \vartheta\begin{bmatrix}1/2\\0\end{bmatrix}(-1/\tau) = (-i\tau)^{1/2}\,\vartheta\begin{bmatrix}0\\1/2\end{bmatrix}(\tau)$$

$$\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}(\tau+1) = e^{-\pi i/4}\,\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}(\tau) \quad ; \quad \vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}(-1/\tau) = i(-i\tau)^{1/2}\,\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}(\tau)$$

Appendix B

Rudiments of group theory

In this appendix we provide some basic techniques in group theory that we will need to be familiar with. Useful references are [120, 121] and the more formal [122, 123].

B.1 Groups and representations

B.1.1 Group

A **group** G is a set on which there exists a multiplication, satisfying

- Closure: For any $g, h \in G$, $g \cdot h \in G$
- Identity element: there exists an element $e \in G$ such that $e \cdot g = g \cdot e = g$ for any $g \in G$
- Inverse: For any $g \in G$ there exists an element g^{-1} such that $g \cdot g^{-1} = g^{-1} \cdot g = e$
- Associativity: $(g \cdot h) \cdot k = g \cdot (h \cdot k)$ for any $g, h, k \in G$

Notice that commutativity $g \cdot h = h \cdot g$ is not required to be a group. If any pair of elements commute, the group is called abelian.

B.1.2 Representation

A **representation** R of a group is a mapping that, to each element of G associates a linear operator $R(g)$ acting on a vector space V, in a way compatible with the group multiplication, namely

$$R(g)\,R(h) \;=\; R(g \cdot h) \qquad \forall g, h \in G \tag{B.1}$$

Hence a representation is a homomorphism between G and the set of linear operators on V. If it is an isomorphism (injective and onto), then the representation is called **faithful**.

The vector space V is called the **representation space**, and vectors in V are said to form the representation R of G. The group G is said to acto on V (or on vectors of V) in the representation R.

If the dimension of V is n, and we fix a basis $|e_i\rangle$, any linear operator can be regarded as an $n \times n$ matrix via

$$R(g)_{ij} \;=\; \langle e_i|\, R(g)\, |e_j\rangle \tag{B.2}$$

So a representation can be defined also as a homomorphism between G and the set of $n \times n$ matrices. We call these matrix representations of G.

Notice that the explicit matrix that represents an element $g \in G$ in a matrix representation, depends on the basis. Hence, it makes sense to define an equivalence relation of matrix representations. Two matrix representations R and R' are **equivalent** if there exist a similarity transformation S ($n \times n$ invertible matrix) such that

$$R'(g) \;=\; S\,R(g)\,S^{-1} \qquad \forall g \in G \tag{B.3}$$

Namely the matrices $R(g)$ and $R'(g)$ are related by a (g-independent) change of basis in V.

OBS: Often, one find a group acting on a physical system in a particular representation. It is however important to distinguish between the abstract group and its different representations.

B.1.3 Reducibility

A representation R is **reducible** if it has a matrix version equivalent to a representation with block diagonal matrices

$$R(g) \;=\; \begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix} \qquad \forall g \in G \tag{B.4}$$

Hence V splits into V_1 and V_2, which are acted on, but not mixed, by $R_1(g)$ and $R_2(g)$, respectively.

An **irreducible** representation (**irrep** for short) is one which is not reducible.

B.1.4 Examples

• The **trivial** representation, which exists for any group G. To every element, it associates the 1×1 matrix 1.

$$R(g) = 1 \quad \forall g \in G \tag{B.5}$$

It is clearly a homomorphism, but not an isomorphism. It is not a faithful representation

• Irreps of $\mathbf{Z_3}$. The group $\mathbf{Z_3}$ has three elements, 1, g and g^2, with the group multiplication law $g^k \cdot g^l = g^{k+l}$, $g^3 = 1$.

It has three inequivalent irresps, which are all 1-dimensional. One of them is the trivial

$$1 \to 1 \quad ; \quad g \to 1 \quad ; \quad g^2 \to 1 \tag{B.6}$$

There are two faithful representations

$$\begin{array}{lllllll} R_1 & : & 1 \to 1 & ; & g \to e^{2\pi i/3} & ; & g^2 \to e^{4\pi i/3} \\ R_2 & : & 1 \to 1 & ; & g \to e^{4\pi i/3} & ; & g^2 \to e^{2\pi i/3} \end{array} \tag{B.7}$$

In fact, it is easy to show that for an abelian group all irreducible representations are necessarily 1-dimensional.

• Group of symmetries of the square. This group is generated by two elements: α, a rotation or 90 degrees around the center of the square, and β, a flip around a vertical axis. Any other element can be obtained by taking products of these. A simple 2-dimensional faithful irrep of this group is

$$\alpha \to \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad ; \quad \beta \to \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{B.8}$$

and the corresponding product matrix for other elements.

B.1.5 Operations with representations

It is useful to define them in terms of matrix representation. Let R_1, R_2 be representations of a group G on vector spaces V_1, V_2, on which we specify a basis $|e_i\rangle$, $|f_m\rangle$, of dimensions n_1, n_2 respectively.

• **Sum of representations** We define the sum representation $R_1 \oplus R_2$, acting on $V_1 \oplus V_2$ as

$$R(g) = \begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix} \tag{B.9}$$

It has dimension $n_1 + n_2$, and is clearly reducible.

• **Tensor product of representations**. We define the product represention $R = R_1 \otimes R_2$, acting on $V_1 \times V_2$ (which has basis $|e_i\rangle \otimes |f_m\rangle$) as

$$(R(g))_{im,jn} \;=\; (R_1(g))_{ij}\,(R_2(g))_{mn} \tag{B.10}$$

It has dimension $n_1 n_2$ and is in general reducible. The decomposition of tensor product representations as sum or irreps is a canonical question in group theory, which can be systematically solved using Clebsch-Gordan techniques.

B.2 Lie groups and Lie algebras

B.2.1 Lie groups

A **Lie group** G is a group where the elements are labeled by a set of continuous real parameters, ξ^a, $a = 1, \ldots, N$, with the multiplication law depending smoothly on the latter. Namely

$$g(\xi) \cdot g(\xi') \;=\; g(f(\xi, \xi')) \tag{B.11}$$

with $f^a(\xi, \xi')$ a continuous (usually also C^∞) function of ξ, ξ'.

OBS: The Lie group is a differentiable manifold, and the ξ are coordinates. Usually we define the parameters such that $g(\xi = 0) = e$, the identity element of G. The number of parameters N is called the dimension of the group.

We will be interested in compact Lie groups (which are compact as manifolds), although there exist very important non-compact Lie groups, for instance, the Lorentz group (where the boost parameters correspond to noncompact directions).

Lie groups also have representations. As usual, to each element $g(\xi) \in G$ they associate a linear operator $R(g(\xi))$ on a vector space V, compatibly with the group law. The dimension of V is unrelated to N the dimension of the group. For short we denote $R(g(\xi))$ by $R(\xi)$.

B.2.2 Lie algebra A(G)

Formally, it is the tangent space to the manifold G at the point corresponding to the identity element, see fig B.1. Since the geometry of G is so constrained

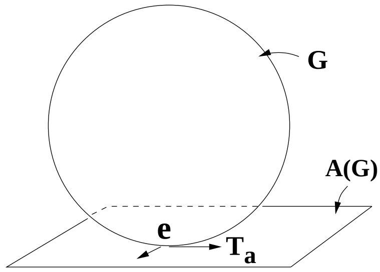


Figure B.1: The Lie algebra is in a very precise sense the tangent space to the Lie group at the point corresponding to the identity element.

by the group law, its structure is almost completely encoded just in the tangent space.

Recall the differential geometry definition of tangent space of a manifold M at a point P. It is the vector space generated by the objects ∂_a, $a = 1, \ldots, \dim M$; the latter are vectors, formally defined as mappings from the space of functions on M, $\mathcal{F}(\mathcal{M})$ to the real numbers

$$\begin{aligned} \partial_a : \quad \mathcal{F}(\mathcal{M}) \rightarrow & \quad \mathbf{R} \\ f(x) \rightarrow & \quad \partial_a f(x)|_P \end{aligned} \tag{B.12}$$

In Lie groups, the natural functions of G are matrix valued functions compatible with the group law, namely representations. Hence we define the vectors T_a as mappings from the space of representations of G, $\mathcal{R}(G)$ to the space of matrices Mat

$$\begin{aligned} T_a : \quad \mathcal{R}(\mathcal{M}) \rightarrow & \quad \mathrm{Mat} \\ R(g(\xi)) \rightarrow & \quad -i\partial_a R(g(\xi))|_{\xi=0} \end{aligned} \tag{B.13}$$

This formal definition is used to emphasize that the properties of the T_a are properties of the group and not of any particular representation. In this sense, this can be formally written as '$T_a = -i\partial_a g|_e$'. However, it is often useful to discuss properties etc in terms of representations.

For a fixed representation R, we call $-i\partial_a R(\xi)|\xi = 0$ the representation of T_a in the representation R, and call it t_a^R. It is interesting to note that changes of coordinates in G induce linear transformations on the T_a's, as follows

$$T'_a = \frac{\partial \xi^b}{\partial \xi'^a} T_b \tag{B.14}$$

We can form linear combinations and multiply the T_a's, as induced from sum and product of matrices. Roughly speaking the Lie algebra is the algebra generated by the T_a's with this sum and product. The linear combinations $\sum_a \lambda_a T_a$ are called generators of the group/algebra (often, just the T_a are called generators of the algebra).

B.2.3 Exponential map

Generators provide infinitesimal transformations

$$g(0,\dots,\delta\xi^a,\dots,0) \,=\, e \,+\, \partial_a g\,\delta\xi^a \,=\, e \,+\, i\,T_a\,\delta\xi^a \tag{B.15}$$

In fact, they are associated to whole one-parameter subgroups of G (which are said to be generated by T_a). In any representation R

$$\begin{aligned} R(0,\dots,\xi^a+\delta\xi^a,\dots,0) \,&=\, R(0,\dots,\delta\xi^a,\dots,0)\,R(0,\dots,\xi^a,\dots,0) \,= \\ &=\, (1+\partial_a R|_{\xi=0}\delta\xi^a)\,R(0,\dots,\xi^a,\dots,0) \end{aligned} \tag{B.16}$$

On the other hand

$$R(0,\dots,\xi^a+\delta\xi^a,\dots,0) \,=\, R(0,\dots,\xi^a,\dots,0) \,+\, \partial_a R|_{\xi=0}\delta\xi^a \tag{B.17}$$

So we get

$$\partial_a R(0,\dots,\xi^a,\dots,0) \,=\, i\,t^R_a\,R(0,\dots,\xi^a,\dots,0) \tag{B.18}$$

Hence

$$R(0,\dots,\xi^a,\dots,0) \,=\, e^{it^R_a\xi^a} \quad (\text{no sum}) \tag{B.19}$$

In the abstract group/algebra

$$g(0,\dots,\xi^a,\dots,0) \,=\, e^{i\,T_a\,\xi^a} \quad (\text{no sum}) \tag{B.20}$$

In fact, any element of the group $g(\xi)$ continously connected to the identity can be written as

$$g(\xi) = e^{i\,\sum_a T_a\,\xi^a} \tag{B.21}$$

for a suitable generator $\sum_a \xi^a T_a$ in the algebra, see figure B.2. So the whole group can be recovered from the structure of the algebra [1]

[1]In fact, some global information on the group may not be recovered from the algebra. There are groups which are globally different yet have the same Lie algebra. They are typically quotients of each other, so they differ in their homotopy groups. The group recovered from the algebra is the so-called universal cover group, which is the only simply connected group with that algebra. This subtle issue is what makes $SU(2)$ and $SO(3)$ have the same Lie algebra although $SU(2)$ is simply connected and $SO(3) = SU(2)/\mathbf{Z_2}$.

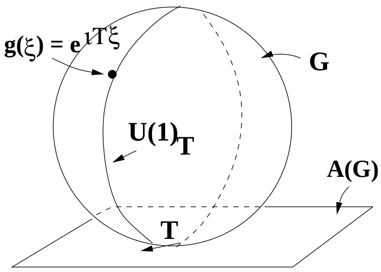


Figure B.2: Any element in the group (in the component continuously connected to the identity) can be obtained from a generator in the Lie algebra by the exponential map.

B.2.4 Commutation relations

The generators T_a satisfy simple commutation relations

$$[T_a, T_b] = i f_{abc} T_c \tag{B.22}$$

where f_{abc} are called the structure constants of the group/algebra.

i) They are determined by the group multiplication law. To see this, consider the group element $g(\lambda)$ defined by

$$g_{ab}(\lambda) \;=\; e^{i\lambda T_b}\, e^{i\lambda T_a}\, e^{-i\lambda T_b}\, e^{-i\lambda T_a} \tag{B.23}$$

Expanding around $\lambda = 0$, we have

$$g_{ab}(\lambda) \;=\; 1 \,+\, \lambda^2\, [T_a, T_b] \,+ \ldots \tag{B.24}$$

Since $g(\lambda)$ is a group element, infinitesimally close to the identity, it also has the expansion as identity plus some element in the algebra

$$g_{ab}(\lambda) \;=\; 1 \,+\, \lambda^2 \sum_c f_{abc} T_c \tag{B.25}$$

By comparing, we get the commutation relations (B.22)

ii) They determine the group multiplication law, at least for elements connected to the identity. To see that, consider two group elements $e^{i\lambda^a T_a}$ and $e^{i\sigma^a T_a}$, their product is some element $e^{i\rho^a T_a}$. The Lie algebra information is enough to find the ρ^a in terms of the λ^b, σ^c. By expansion of the relation

$$e^{i\lambda^a T_a}\, e^{i\sigma^a T_a} \;=\; e^{i\rho^a T_a} \tag{B.26}$$

we get

$$\rho^a = \lambda^a + \sigma^a - \frac{1}{2} f_{abc} \lambda^b \sigma^c + \dots \tag{B.27}$$

this verifies our claim.

The commutation relations satisfy the Jacobi identities

$$[T_a, [T_b, T_c]] + [T_c, [T_a, T_b]] + [T_b, [T_c, T_a]] = 0 \tag{B.28}$$

(as in any representation they are simply matrices which obviously satisfy this relation). This can be easily translated into a relation among the structure constants.

A representation R of the Lie algebra is a mapping that to each T_a it associates a linear operator t_a^R (acting on a space V of some dimension n, independent of the dimension N of the group), consistently with linear combinations and with the commutation relations, namely

$$[t_a^R, t_b^R] = i f_{abc} t_c^R \tag{B.29}$$

Clearly the structure constants are a property of the group/algebra and not of the representation.

Clearly, given a representation of the group we can build a representation of the algebra (by taking representations of group elements close to the identity $t_a^R = -i\partial_a R(\xi)$), and viceversa (by the exponential mapping $R(\xi) = e^{it_a^R \xi^a}$).

The structure constants depend on the choice of basis in the Lie algebra, so it is convenient to fix a canonical choice. To fix it, consider the quantity $\mathrm{tr}\,(t_a^R t_b^R)$ in any representation R; it is a real and symmetric matrix, which can be diagonalized by a change of basis in the Lie algebra. Once we are in such basis $\mathrm{tr}\,(t_a^R t_b^R) = k_R \delta_{ab}$ and we obtain the structure constants as

$$f_{abc} = -\frac{i}{k_R}\,\mathrm{tr}\,([t_a^R, t_b^R] t_c^R) \tag{B.30}$$

and are completely antisymmetric.

Since this can be played for any representation R, it shows that there exists a basis in the abstract Lie algebra where (B.22) hold with completely antisymmetric structure constants.

In the remaining of this lecture we will center on compact Lie groups, for which any representation is equivalent to a unitary representation. In such representation all generators are hermitian and the structure constants are real.

B.2.5 Some useful representations

There is a very useful representation which is canonically built in the structure of the Lie algebra. It is the **adjoint** representation, which is N-dimensional (same dimension as the group). Consider an N-dimensional vector space, with a set of basis vectors labeled by the generators of the algebra $|T_a\rangle$, $a = 1, \ldots, N$. And represent T_a by the linear operator $t_a^{\rm Adj}$ defined by

$$t_a^{\rm Adj}|T_b\rangle \;=\; |[T_a, T_b]\rangle \;=\; i\, f_{abc}\, |T_c\rangle \tag{B.31}$$

Namely we have the matrix elements $(t_a^{\rm Adj})_{bc} = -i f_{abc}$.

Given any representation R, with generators represented by t_a^R, we can build another representation R^*, called the **conjugate representation**, with generators represented by $-(t_a^R)^T$. It is a simple exercise to check that it also provides a representation of the algebra.

B.3 SU(2)

To warm up before the study of more general Lie algebras, we study the construction of representations for $SU(2)$, the simplest non-abelian group. The Lie algebra is given by

$$[J_a, J_b] = i\epsilon_{abc} J_c \tag{B.32}$$

A familiar representation is provided by the Pauli matrices $J_a = \sigma_a/2$, with

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad ; \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad ; \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{B.33}$$

In this representation, elements of the group correspond to 2×2 unitary matrices. This particular representation arises as the action of the 3d rotation group on spin 1/2 particles. We will be interested in constructing more general representations in a more systematic way.

B.3.1 Roots

We first put the Lie algebra in **Cartan-Weyl form**. To do that, the first step is to choose a maximal set of mutually commuting generators (this is

Figure B.3: Fig. a) shows the root diagram for the $SU(2)$ Lie algebra; Fig b) shows the general structure of the weights for irreducible representations of this algebra.

the so-called **Cartan subalgebra**, whose dimension is called the **rank** of the group/algebra). For $SU(2)$ any pair of generators is non-commuting, there is at most one such generator, say J_3.

Next, se take the remaining generators are form linear combinations

$$J^{\pm} = \frac{1}{\sqrt{2}}(J_1 \pm iJ_2) \tag{B.34}$$

such that they have simple commutation relations with the Cartan generator J_3

$$[J_3, J^+] = J^+ \quad ; \quad [J_3, J^-] = -J^- \tag{B.35}$$

In intuitive terms, this tells us the charges of $J^{\pm}$ with respect to the $U(1)$ subgroup generated by the Cartan J_3. In the adjoint representation, we have the relation $J_3|J^{\pm}\rangle = \pm|J^{\pm}\rangle$; upon exponentiation, $g(\xi)|J^{\pm}\rangle = e^{\pm i\xi J_3}|J^{\pm}\rangle$, namely $|J^{\pm}\rangle$ tranform with charges $\pm$ under the $U(1)$ generated by J_3. By abuse of language we use the same language for $J^{\pm}$ themselves.

We also have

$$[J_3, J_3] = 0 \quad ; \quad [J^+, J^-] = J_3 \tag{B.36}$$

This are the commutation relations for the algebra written in the Cartan-Weyl form. The charges of the different generators with respect to the $U(1)$ generated by the Cartan J_3 are called the **roots** of the algebra. In our case we have the roots -1, 0, $+1$ for J^-, J_3, J^+ respectively.

The roots of an algebra are drawn in a **root diagram**, as in figure B.3a). Such picture encodes all the information about the algebra.

B.3.2 Weights

Let us now discuss the construction of irreps. The representation space is a vector space spanned by a set of basis vectors. It is natural to take a basis where the representative of J_3 is diagonal, and then it is natural to label each vector in the basis by its J_3 eigenvalue, $|\mu\rangle$. Hence we have by construction

$$J_3|\mu\rangle = \mu|\mu\rangle \tag{B.37}$$

The eigenvalues μ are in principle real numbers, which give us the charge of the corresponding eigenstate with respect to the $U(1)$ generated by J_3. Such charges are called **weights** of the representation. The irrep is essentially defined by giving the set of weights for all basis vector in the representation space, and it is usual to draw the weights in a **weight diagram** (see below) that encodes all information about the representation.

We define the **highest weight** as the highest of all eigenvalues, and call it j. Soon we will see that the complete irrep is defined just in terms of its highest weight.

An important fact is that **weights** in an irrep differ by roots. Starting with a state of weight $|\mu\rangle$, we can build the states $J^\pm|\mu\rangle$, which are eigenstates of J_3 with eigenvalues $\mu \pm 1$

$$J_3 J^\pm|\mu\rangle \;=\; ([J_3, J^\pm] + J^\pm J_3)|\mu\rangle \;=\; (\pm J^\pm + \mu J^\pm)|\mu\rangle \;=\; (\mu \pm 1)J^\pm|\mu\rangle \tag{B.38}$$

So the states $J^\pm|\mu\rangle$ must be either zero or they are part of our basis vectors. Hence there should exist weigths which are equal to $\mu \pm 1$, namely weights differ by roots.

Since by definition $\mu = j$ was the highest weight, the structure of the basis vectors is

$$|j\rangle\,,\;|j-1\rangle\,,\;|j-2\rangle\;\ldots \tag{B.39}$$

On the other hand, the representations we are interested in are finite dimensional, so the representation should end. To compute when, we must realize that $J^-|\mu\rangle \simeq |\mu - 1\rangle$ up to a normalization factor. Namely, one has

$$\begin{aligned} J^-|\mu\rangle \;&=\; N_\mu\,|\mu-1\rangle \\ J^+|\mu\rangle \;&=\; N_\mu\,|\mu+1\rangle \end{aligned} \tag{B.40}$$

and the coefficent can be computed to be

$$N_\mu = \frac{1}{\sqrt{2}}\sqrt{(j+\mu)(j-\mu+1)} \tag{B.41}$$

which means that the representation is finite-dimensional if some $\mu = -j$

$$J^-|-j\rangle = 0 \tag{B.42}$$

Since μ's differ by integers, j and $-j$ must differ by an integers, which implies the constraint that j must be integer or half odd.

Hence irreps of $SU(2)$ are characterized by a highest weight, which must be an integer of half-odd number. The representation space is spanned by the basis vectors

$$|j\rangle\,,\,|j-1\rangle\,,\,|j-2\rangle\,\ldots\,|-j\rangle \tag{B.43}$$

which is $(2j+1)$-dimensional. The matrices representing generators in this space are easy to obtain from the actions of $J^{\pm}$, J_3 on the basis vectors. All the information of the irrep with highest weight j is encoded in a weight diagram as in figure B.3b.

B.4 Roots and weights for general Lie algebras

The idea is to generalize to any Lie algebra the procedure introduced for $SU(2)$.

B.4.1 Roots

First we put the Lie algebra in the **Cartan-Weyl form**. The first step is to pick a maximal set of mutually commuting hermitian [2] generators, which we call H_i, $i = 1\ldots,r$. The number of such generators is called the **rank** r of the group; they generate the **Cartan subalgebra** of the Lie algebra. Upon exponentiation, they generate a $U(1)^r$ subgroup of the Lie group.

The second step is to take linear combinations of the remaining operators so that they have easy commutators with the H_i. To do that, we go to the adjoint representation, with basis vectors $|T_a\rangle$, and construct the matrix

$$M_{ab}^{(i)} = \langle T_a|H_i|T_b\rangle \tag{B.44}$$

[2] By abuse of language we talk about a hermitian generator in the abstract algebra, as a generator which is represented by a hermitian operator/matrix in any unitary representation.

Diagonalizing simultaneously the matrices $M^{(i)}$ (they commute since they represent the Cartan generators, which commute in the abstract algebra), we get a new basis of vectors $|E_\alpha\rangle$, which are eigenstates of the H_i (better, of their representatives in the adjoint representation). We label each such state by its r eigenvalues α_i with respect to H_i.

$$H_i|E_\alpha\rangle = \alpha_i|E_\alpha\rangle \tag{B.45}$$

At the level of the abstract algebra, this induces some linear combinations of the original generators T_a into some generators E_α with commutation relations

$$[H_i, E_\alpha] = \alpha_i E_\alpha \tag{B.46}$$

These are not hermitian, rather $E_\alpha^\dagger = E_{-\alpha}$

Using the Jacobi identity it is also possible to show that

$$\begin{aligned} [E_\alpha, E_{-\alpha}] &= \sum_i \alpha_i H_i \\ [E_\alpha, E_\beta] &= E_{\alpha+\beta} \;\text{ if }\; \alpha+\beta \text{ is root} \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{B.47}$$

The r-dimensional vectors α are called the **roots of the Lie algebra**, and they provide the charges of the E_α with respect to the $U(1)^r$ generated by the Cartan subalgebra.

B.4.2 Weights

To describe irreps, we choose a basis of the representation space where all matrices representing the Cartan generators are diagonal, and we label the vectors in the basis (eigenstates of the matrix representing H_i) by the corresponding eigenvalues. By abuse of language, we denote H_i the matrix representing the abstract H_i in the representation. We have

$$H_i|\mu\rangle = \mu_i|\mu\rangle \qquad i = 1, \ldots, r \tag{B.48}$$

The r-dimensional vectors μ are called **weights of the representation**. The set of weights of a representation characterize the representation.

OBS: Notice that the weights are a property of the representation, while the roots are a property of the algebra. Notice also that the weights of the adjoint representation are the roots of the Lie algebra (this is beacuse the adjoint is a very canonical representation, built into the structure of the algebra itself).

OBS: Notice that in an irreducible representation there may be different states with the same weight vectors. One (special) example is the states $|H_i\rangle$ in the adjoint representation, which all have weight equal to zero. One must be careful in dealing with situations where different vectors have same weights.

In a given representation, weights are not arbitrary. Rather, as in $SU(2)$, **weights differ by roots**. Namely, starting with an state $|\mu\rangle$ we can construct $E_{\pm\alpha}|\mu\rangle$ which is an eigenstate of the H_i, with eigenvalue $\mu_i \pm \alpha_i$, as follows

$$H_i E_{\pm\alpha}|\mu\rangle \; = \; (\alpha_i E_{\pm\alpha} + E_{\pm\alpha} H_i)\,|\mu\rangle \; = \; (\mu_i \pm \alpha_i) E_{\pm\alpha}|\mu\rangle \qquad \text{(B.49)}$$

So there must in principle exist a weight in the representation given by the vectors $\mu + \alpha$, and a corresponding state $|\mu \pm \alpha\rangle$. In fact, as in $SU(2)$ we have a relation modulo a coefficient

$$E_{\pm\alpha}|\mu\rangle \; = \; N_{\mu,\pm\alpha}|\mu \pm \alpha\rangle \qquad \text{(B.50)}$$

and for some μ we will have $N_{\mu,\pm\alpha} = 0$, which ensures that representations are finite-dimensional, and impose some additional constraints on the possible values of the weights μ. The sets of allowed irreps and the corresponding weights is difficult to analyze in general, and we leave their discussion for specific examples, see sections B.6.

It is worth pointing out that the analogy with $SU(2)$ is quite precise. In fact, for any non-zero root α, the generators $E_{\pm\alpha}$, $\sum_i \alpha_i H_i$ form an $SU(2)$ subalgebra of the Lie algebra. Defining $E^{\pm} = \frac{1}{|\alpha|} E_{\pm\alpha}$, $E_3 = \frac{1}{|\alpha|^2} \sum_i \alpha_i H_i$ we have the commutators

$$[E_3, E^{\pm}] = \pm E^{\pm} \quad ; \quad [E^+, E^-] = E_3 \qquad \text{(B.51)}$$

which is an $SU(2)$ algebra in the Cartan-Weyl form. This means that for any μ the states $|\mu + k\alpha\rangle$ form an irrep of this $SU(2)$.

For future convenience, we use this a bit further. This irrep will contain some highest and lowest $SU(2)$ weight states $|j\rangle$ and $|-j\rangle$, namely there exist integers p, q such that

$$E_\alpha|\mu + p\alpha\rangle = 0 \quad ; \qquad j = \frac{\alpha \cdot \mu}{|\alpha|^2} + p$$
$$E_{-\alpha}|\mu - q\alpha\rangle = 0 \quad ; \qquad -j = \frac{\alpha \cdot \mu}{|\alpha|^2} - q \tag{B.52}$$

so we get $\frac{\alpha\cdot\mu}{|\alpha|^2} = -\frac{1}{2}(p-q)$. This is the master formula extensively used in the classification of Lie algebras, see section B.5.

The basic strategy to build irreps is therefore as follows. We need to introduce the concept of a highest weight. To do so, we define a positive vector in the r-dimensional space of roots/weights/charges, $v > 0$ if $v_1 > 0$; if $v_1 = 0$ we say that $v > 0$ if $v_2 > 0$; etc. We say that one vector v is higher than other vector w, $v > w$, if $v - w > 0$. This allows to define the **highest weight** μ_0 of a representation the weight such that $\mu_0 > \mu$ for any other weight μ.

The concept of positivity allows to split the set of non-zero roots into the set of positive roots and of negative roots. For $\alpha > 0$ the E_α are raising operators and the $E_{-\alpha}$ are lowering operators. The highest weight vector is characterized by the fact that it is annihilated by the raising operators (if not, we would get states $|\mu_0 + \alpha\rangle$ with weight higher than $|\mu_0\rangle$, which was defined as the highest!).

The representation is build by applying lowering operators to the highest weight state, in all possible inequivalent ways, until we exhaust the representation (namely, until we start finding zeroes upon application of lowering operators). That this happens is guaranteed because states form representations of the $SU(2)$'s associated to each α, and such representations are finite-dimensional from our experience with $SU(2)$

B.4.3 SU(3) and some pictures

Instead of giving the commutation relations of the $SU(3)$ algebra, all the relevant information is provided by the root diagram of the algebra, shown in figure B.4. Namely, the rank is two; the Cartan subalgebra is spanned by two generators H_1, H_2, which are mutually commuting. The remaining eight generators are labelled E_α, $E_{-\alpha}$ for $\alpha = (1,0)$, $(1/2, \sqrt{3}/2)$, $(1/2, -\sqrt{3}/2)$, and have commutation relations

$$[H_i, E_{\pm\alpha}] = \pm\alpha_i E_{\pm\alpha} \tag{B.53}$$

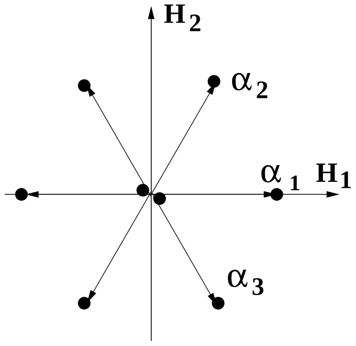


Figure B.4: The root system of the $SU(3)$ Lie algebra. The positive roots are $\alpha_1 = (1,0)$, $\alpha_2 = (1/2, 1/(2\sqrt{3})$, $\alpha_3 = (1/2, -1/(2\sqrt{3})$. The two roots at $(0,0)$ correspond to the Cartan generators.

Notice the $SU(2)$ subalgebras along the different α's, which graphically correspond to lines along which the roots reproduce the root diagram of $SU(2)$.

Some representations

Instead of writing the explicit matrices providing a particular representation of the $SU(3)$ algebra, we can instead provide the weight diagram of the corresponding representation.

A familiar representation is the fundamental representation, which is 3-dimensional, and on which the generators are represented as 3×3 hermitian matrices (the Gell-Mann matrices). Upon exponentiation, the group elements are represented as 3×3 unitary matrices.

This representation can be equivalentely described by the weights in picture B.5a. The action of the Cartans on the states $|\mu = (\pm 1/2, 1/(2\sqrt{3}))$, $(0, -1/\sqrt{3})$ is

$$H_i|\mu\rangle = \mu_i|\mu\rangle \tag{B.54}$$

The action of non-zero root generators E_α is

$$E_\alpha|\mu\rangle = N_{\mu,\alpha}|\mu + \alpha\rangle \tag{B.55}$$

Notice that the states form representations under the $SU(2)$ subalgebras of the non-zero roots. That is, weights along lines parallel to the root diagram of the corresponding $SU(2)$ subgroup differ by the corresponding root.

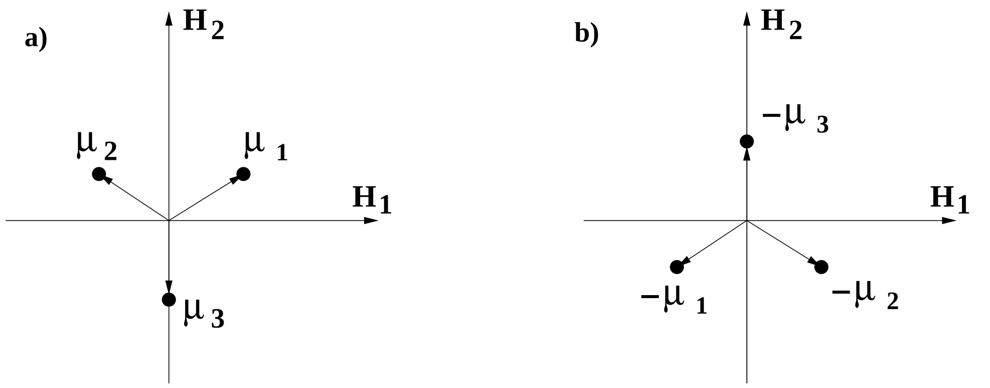


Figure B.5: The weight diagram for the fundamental (a) and antifundamental representations of $SU(3)$.

The construction of the irrep is as follows. The highest weight is $|(1/2, 1/(2\sqrt{3})\rangle$, so this is annihilated by the positive roots $\alpha_1 = (1,0)$, $\alpha_2 = (1/2, 1/(2\sqrt{3})$, $\alpha_3 = (1/2, -1/(2\sqrt{3})$. The remaining states are obtained as

$$\begin{aligned} E_{-\alpha_1}|(1/2, 1/(2\sqrt{3})\rangle &\simeq |(-1/2, 1/(2\sqrt{3})\rangle \\ E_{-\alpha_2}|(1/2, 1/(2\sqrt{3})\rangle &\simeq |(0, -1/\sqrt{3})\rangle \end{aligned} \tag{B.56}$$

The conjugate representation, the antifundamental, which is obtained by minus the transposed GellMann matrices, has weights opposite to those of the fundamental. Namely, conjugation of the representation flips the charges of objects. The weights are shown in figure B.5b

B.5 Dynkin diagrams and classification of simple groups

The discussion in this section will be very sketchy. For more information, see chapter VIII of [120]. However, the discussion is not too relevant, one can jump to the results directly.

The information we have obtained is also useful in yielding information that can be used to classify all possible Lie algebras. In fact in the study of representations we obtained some interesting constraints. For instance, recall the master formula that for any representation, the fact that $|\mu + k\alpha$ for a representation of $SU(2)_\alpha$ implied that the weights satisfy

$$\frac{\alpha \cdot \mu}{|\alpha|^2} = -\frac{1}{2}(p-q) \tag{B.57}$$

In particular we may apply this to the adjoint representation, where the weight μ is a root. Requiring that the states $|\beta + k\alpha\rangle$ form a representation of $SU(2)_\alpha$, and that the states $|\alpha + k\beta\rangle$ form a representation of $SU(2)_\beta$, we get

$$\frac{\alpha \cdot \beta}{|\alpha|^2} = -\frac{1}{2}m \quad ; \quad \frac{\beta \cdot \alpha}{|\beta|^2} = -\frac{1}{2}m \quad ; m.m' \in \mathbf{Z} \tag{B.58}$$

We obtain a constraint on the relative angle of the roots

$$\cos^2 \theta_{\alpha,\beta} = \frac{(\alpha \cdot \beta)^2}{|\alpha|^2|\beta|^2} = \frac{mm'}{4} \tag{B.59}$$

The angle is constrained to be 0, 30, 45, 60, 90, 120, 135, 150 or 180 degrees.

B.5.1 Simple roots

We now define a **simple root** as a positive root which cannot be written as a sum of positive roots with positive coefficients. Simple roots have nice properties, in particular the set of simple roots of an algebra is linearly independent, and there are r simple roots; so simple roots provide a basis of root space.

Moreover, the angles between simple roots are more constrained. To see this, notice that if α and β are simple roots, then $\alpha - \beta$ is *not* a root [3]. Now going to the adjoint representation, $E_{-\alpha}$ must annihilate E_β (since otherwise it would create a state $|E_{\beta-\alpha}\rangle$, but $\beta - \alpha$ is not a root!), so $|E_\beta\rangle$ is the lower weight state $|-j\rangle$ for the subalgebra $SU(2)_\beta$, and we get

$$2\frac{\alpha \cdot \beta}{|\alpha|^2} = -p \quad ; p \in \mathbf{Z}^+ \tag{B.60}$$

[3] If it were, it would be positive or negative; if it is positive then $\alpha = \beta + (\alpha - \beta)$ contradicts the fact that α is simple; if it is negative, then $\beta = \alpha + (\beta - \alpha)$ contradicts that β is simple

Hence the quantities $2\frac{\alpha\cdot\beta}{|\alpha|^2}$ are non-positive integers for simple roots. Using

$$2\frac{\alpha\cdot\beta}{|\alpha|^2} = -p \quad ; \quad 2\frac{\alpha\cdot\beta}{|\beta|^2} = -p' \tag{B.61}$$

we get $\cos\theta_{\alpha,\beta} = -\frac{1}{2}\sqrt{pp'}$, and this forces the angles between simple roots to be 90, 120, 135 or 150 degrees.

B.5.2 Cartan classification

The only invariants of the set of simple roots are the relative lengths and angles of the simple roots. Use of this information is enough to recover the complete system of roots, since simple roots provide a basis. Hence the problem of classification of Lie algebras is the problem of classifying sets of r linearly independent vectors in r-dimensional space with non-positive integer values of $2\alpha\cdot\beta/|\alpha|^2$.

In the classification it is important to note the following. Two r_1- resp r_2-dimensional systems of simple roots, satisfying the above properties, can always be combined into a new (r_1+r_2)-dimensional simple root system, by simply joining orthogonally the two initial systems. Clearly we are interested in root systems which cannot be split into orthogonal subsystems.

This is related to the concept of invariant subalgebra. Given and algebra A, an invariant subalgebra B is a subalgebra such that the commutator of any element in B with any element in A is still in A. Upon exponentiation, Lie algebras with invariant subalgebras lead to non-simple groups, namely groups which split as product of groups, $G = G_1 \times G_2$.

So one is in principle interested in classifying simple groups (as any other is obtained by taking products) and Lie algebras without invariant subalgebras (simple Lie algebras). Lie algebras with invariant subalgebras manifest as root systems which split into two orthgonal subsystems. Hence we are interested in classifying simple root systems without such subsystems. Any other can be obtained by simple adjunction.

The problem of classifying simple root systems of this kind has been solved. The result, called the Cartan classification can be recast is giving the relative lengths and angles between the simple roots. This is conveniently codified in a picture called the Dynkin diagram. The classification of **Dynkin diagrams** for simple Lie algebras is given in figure B.6. The rules to obtain the simple root system from the diagram are as follows.

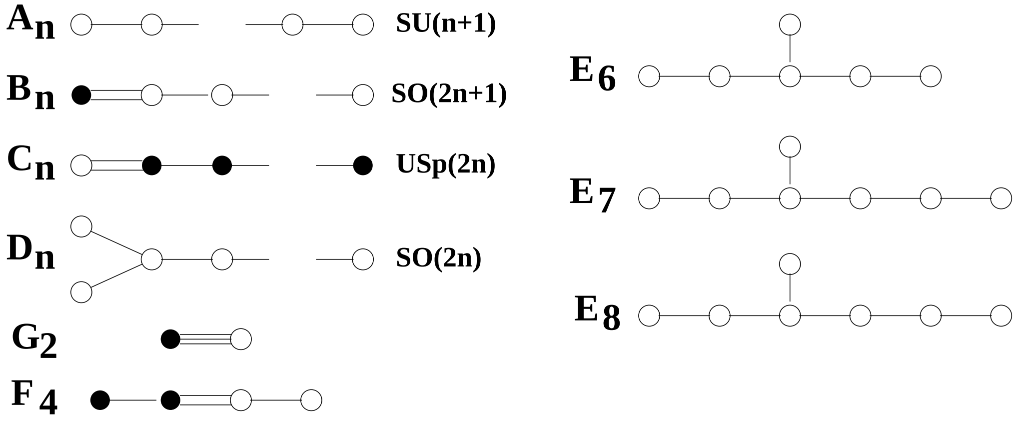


Figure B.6: Dynkin diagrams for simple Lie algebras. There are four infinite series (labeled by a positive integer r giving the number of nodes), and some exceptional algebras. Notice that for small rank some algebras are isomorphic and have the same Dynking diagram (e.g. $A_3 = D_3$, namely $SU(4) \simeq SO(6)$. The groups arising from the A, B, C, and D series were known in classical mathematics before Cartan and are known as classical Lie groups, they are listed to the right of the corresponding diagram.

• Each node corresponds to a simple root (hence the number of nodes is the rank of the Lie algebra/group)

• The number of lines joining two nodes gives us the angle between the two simple roots: no line means 90^o, one line means 120^o, two lines means 135^o, three lines means 150^o.

• Dark nodes correspond to shorter roots (the relative lengths can be found from (B.59)

Clearly, Dynking diagrams corresponding to non-simple algebras are obtained by adjoining in a disconnected way Dynkin diagrams for simple algebras (so that we adjoin orthogonally the two subsystems of simple roots).

B.6 Some examples of useful roots and weights

There are some systems of roots and weights that we will encounter in our study of string theory. In this section we list some of them. A more complete reference, which includes a systematic discussion of tensor products or irreps,

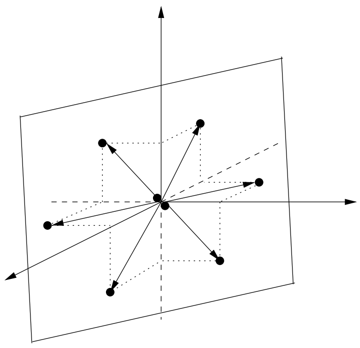


Figure B.7: The root system of $SU(3)$ described as a set of vectors lying in a 2-plane in 3-dimensional space.

and decomposition of representations under subgroups, is the appendices of [124].

B.6.1 Comments on $SU(k)$

Roots

Although $SU(k)$ (or its algebra A_{k-1}) has rank $k-1$, it is convenient and easier to describe its roots as k-dimensional vectors, which lie on an $(k-1)$-plane. Besides the $k-1$ zero roots associated to the Cartan generators, the non-zero roots are given by the k-dimensional vectors

$$(\underline{+,-,0,\ldots,0}) \tag{B.62}$$

where $+$, $-$ denote $+1$, -1, and where underlining means permutation, namely the $+$ and $-$ can be located in any (non-coincident) positions. Note that all roots satisfy one relation $\sum_{i=1}^{n} v_i = 0$, so they live in a $(k-1)$-plane Π in $\mathbf{R^n}$. There are a total of k^2-1 roots, which is the number of generators of $SU(k)$.

Fixing a basis within the $(k-1)$-plane it is straightforward to read out the roots as $(k-1)$-dimensional vectors. The picture of the root system of $SU(3)$ in this language is given in figure B.7.

The extra direction in the diagram can be regarded as associated to the extra $U(1)$ generator in $U(k) = SU(k)\times U(1)$. Hence, $SU(k)$ weight diagrams embedded in $(k-1)$-planes parallel to Π but not passing through the origin are associated to states which, in addition to being in a representation of $SU(k)$, also carry some charge under the additional $U(1)$.

Weights

A familiar representation is the fundamental representation. The corresponding weights, given as k-dimensional vectors but inside the $(k-1)$-plane Π are,

$$\frac{1}{n}\,(\underline{n-1,-1,\ldots,-1}) \tag{B.63}$$

Notice that weights differ by roots, so application of generators associated to non-zero roots relate states with different weights (or give zero if they take us out of the representation).

In situations where the gauge group is $U(k)$ so there is an additional $U(1)$ generator, the fundamentals of $SU(k)$ may carry some charge, so the weights satisfy the relation $\sum_{i=1}^{n} v_i = q$ for some non-zero constant q giving (up to normalization) the charge under the additional $U(1)$. Very often one finds fundamentals from weights of the form

$$(\underline{+,0,\ldots,0}) \tag{B.64}$$

or

$$\frac{1}{2}(\underline{+,-,\ldots,-}) \tag{B.65}$$

Notice that the weights (B.63) can be written as

$$(\underline{+,0,\ldots,0})\;-\;(1/n,\ldots,-1/n) \tag{B.66}$$

where the second term removes the piece corresponding to the additional $U(1)$ charge. By abuse of language, we will often use things expressions like (B.64) or (B.65) to denote the fundamental even in situations where there is no additional $U(1)$, removing implicitly the piece corresponding to this charge.

The weights for the antifundamental representation are the opposite to those for the fundamental, namely

$$(\underline{-,0,\ldots,0}) \tag{B.67}$$

By this, we mean

$$\frac{1}{n}\,(\underline{-(n-1), 1, \ldots, 1}) \tag{B.68}$$

or any other shifted version, with the understanding that the additional $U(1)$ charge should be removed.

Other representations can be obtained by taking tensor products of the fundamental (using the techniques of Yound tableaux, not discussed in this lecture, see [120] for discussion). The corresponding weights are obtained by adding the weights of the fundamental representation.

For instance, the two-index antisymmetric representation has $k(k-1)/2$ weights

$$(\underline{+, +, 0, \ldots, 0}) \tag{B.69}$$

while the two-index symmetric representation has $k(k+1)/2$ weights

$$(\underline{+, +, 0, \ldots, 0}) \quad ; \quad (\underline{\pm 2, 0, \ldots, 0}) \tag{B.70}$$

They are obtained by adding two times weights of the fundamental representation in a way consisten with antisymmetry or symmetry of the representation.

It is straightforward to derive familiar facts like the equivalence of the antifundamental representation and the $(k-1)$-index antisymmetric representation. They have the same weights.

B.6.2 Comments on $SO(2r)$

Roots

Besides the n zero roots, the non-zero roots for the D_r Lie algebra are given by the r-dimensional vectors

$$(\underline{\pm, \pm, 0, \ldots, 0}) \tag{B.71}$$

Meaning that the + and - can be choses arbitrarily in any non-coincident position. The total number of roots is $2r(2r-1)/2$.

The root system of $SO(4)$ is shown in figure B.8. The fact that there are two subsets of orthogonal roots means that there are invariant subalgebras.

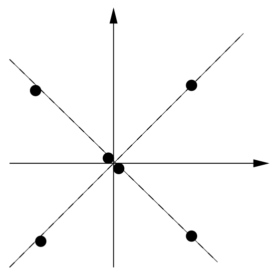


Figure B.8: Root diagram for $SO(4)$. In fact it splits as two orthogonal $SU(2)$ root systems.

In fact, $SO(4) \simeq SU(2) \times SU(2)'$, with non-zero roots of the latter being given by

$$SU(2) \,:\, (++), (--) \quad ; \quad SU(2)' \,:\, (+-), (-+) \tag{B.72}$$

Notice also that the Dynkin diagram for D_2 are two disconnected nodes, so is the same as two A_1 Dynkin diagrams.

It is important to notice that the root system of $SO(2r)$ contains the roots of $SU(r)$, so by exponentiation the group $SO(2r)$ contains a subgroup $SU(r)$.

Weights

An important representation is the **vector representation**, which is $2r$-dimensional and has weights

$$(\underline{\pm, 0, \ldots, 0}) \tag{B.73}$$

Notice that it is a real representation, since its conjugate has opposite weights, but the representation (as a whole) is invariant under such change.

When the group is regarded as the group of rotational isometries of a $2r$ dimensional euclidean space, the vector representation in which vectors of this space transform.

More representations can be obtained by taking tensor products of the vector representation. These are the respresentations under which tensors in the euclidean space transform under rotations.

There are some additional representations which cannot be obtained from tensor products of the vector representation. These are the spinor representations. For D_r there are two inequivalent irreducible spinor representations, both with dimension 2^{r-1}, and weights

$$\begin{aligned} \text{spinor} &: \quad (\pm\frac{1}{2},\ldots,\pm\frac{1}{2}) \quad , \#- = \text{even} \\ \text{spinor}' &: \quad (\pm\frac{1}{2},\ldots,\pm\frac{1}{2}) \quad , \#- = \text{odd} \end{aligned} \tag{B.74}$$

These spinor representations are said to have different chirality [4].

Spinor representations and Clifford algebras

There is a canonical and very useful way to describe the spinor representations of $SO(2r)$, related to representations of **Clifford algebras**. We briefly review this here, since it will appear in our construction of string spectra.

Consider the algebra of objects Γ^i, $i = 1, \ldots, 2r$, satisfying

$$\{\Gamma^i, \Gamma^j\} \, = \, 2\delta_{ij} \tag{B.75}$$

It is called a Clifford algebra. It is important to remark that this is not a Lie algebra! In particular it is not defined in terms or commutators.

The important point is that this algebra is invariant under the group of transformations

$$\Gamma'^i \, = \, R^i_j \, \Gamma^j \tag{B.76}$$

where R is a $2r \times 2r$ orthogonal matrix. This group is precisely $SO(2r)$, and we have found it acting on the set of Γ^i in the fundamental representation.

The fact that the Clifford algebra (B.75) has an $SO(2r)$ invariance menas that any representation of the Clifford algebra must also form a representation of $SO(2r)$. In fact, given a hermitian matrix representation for the Γ^i, the hermitian matrices $J^{ij} = \frac{-i}{4}[\Gamma^i, \Gamma^j]$ can be seen to form a (possibly reducible) hermitian matrix representation of the $SO(2r)$ algebra, which is

$$[J^{ij}, J^{kl}] \, = \, i\,(\delta^{ik}J^{jl} + \delta^{jl}J^{ik} - \delta^{il}J^{jk} - \delta^{jk}J^{il}) \tag{B.77}$$

So our purpose is to build a representation of the Clifford algebra, and the resulting representations of $SO(2r)$. The standard technique to build a

[4] Clearly there discussion of spinors under the Lorentz group in even dimensional space can be recovered from the group theory of spinor representations of $SO(2r)$ (with a few subtleties arising from the non-compactness of the Lorentz group). A nice discussion of Lorentz spinors can be found in the appendices of [71].

representation of the Clifford algebra is to form linear combinations of the Γ^i which can act as raising and lowering operators. We define

$$A_a = \frac{1}{\sqrt{2}}(\Gamma_{2a} + i\Gamma_{2a-1}) \quad ; \quad A_a^\dagger = \frac{1}{\sqrt{2}}(\Gamma_{2a} - i\Gamma_{2a-1}) \quad , \quad a = 1, \ldots, r \tag{B.78}$$

They satisfy the relations

$$\{A_a^\dagger, A_b^\dagger\} = \{A_a, A_b\} = 0 \quad ; \quad \{A_a^\dagger, A_b\} = \delta_{ab} \tag{B.79}$$

So they behave as fermionic oscillator ladder operators. Notice that in this language only an $SU(r)$ invariance is manifest, with the $A_a^\dagger$, A_a transforming in the fundamental resp. antifundamental representations.

To build a representation of the Clifford algebra, we introduce a 'groundstate' for the harmonic oscillator

$$A_a|0\rangle = 0 \tag{B.80}$$

The representation is built by applying raising operators to this 'groundstate' in all possible inequivalent ways. We have

$$\begin{array}{cl} \text{states} & \text{number} \\ |0\rangle & 1 \\ A_a^\dagger|0\rangle & r \\ A_a^\dagger A_b^\dagger|0\rangle & r(r-1)/2 \\ \ldots & \ldots \\ A_{a_1}^\dagger \ldots A_{a_k}^\dagger|0\rangle & \binom{r}{k} \\ \ldots & \ldots \\ A_1^\dagger \ldots A_r^\dagger|0\rangle & 1 \end{array} \tag{B.81}$$

The bunch of $\binom{r}{k}$ states arising from applying k operators to the groundstate clearly forms a k-index completely antisymmetric tensor representation of the $SU(r)$ invariance group.

The total number of states is 2^r. Constructing the Lorentz generators, it is possible to check that the weights are of the form

$$(\pm\frac{1}{2}, \ldots, \pm\frac{1}{2}) \tag{B.82}$$

Moreover, it is easy to realize that the weights among the above with has k $+1/2$'s correspond to the weights of a k-index completely antisymmetric tensor representation of $SU(r)$, in agreement with our above statement.

The above weights therefore define a representation of the $SO(2r)$ group (although only $SU(r)$ invariance was manifest in intermediate steps). Now this representation is reducible. REcalling that the $SO(2r)$ generators are constructed with products of two Γ^i's, it is clear that they are unable to relate states (B.81) with even number of Γ's to states with odd number of Γ's. More formally, one can introduce the chirality operator $\Gamma = \Gamma^1 \ldots \Gamma^{2r}$ which commutes with all $SO(2r)$ generators (and anticommutes with the Γ^i), and can be used to distinguish the two subsets of states.

This means that the 2^r-dimensional representation is reducible into two 2^{r-1}-dimensional irreducible representations, with weights given in (B.74), called the chiral spinor representations.

B.6.3 Comments on $SO(2r+1)$

We will not say much about $SO(2r+1)$, since most of the relevant facts about its representations can be obtained by noticing that it is a subroup of $SO(2r+2)$ and that it contains an $SO(2r)$ subgroup.

Let us simply say that it has an $(2r+1)$-dimensional vector representation, out of which other tensor representations can be obtained by tensor produce. It also has a unique spinor representation, of dimension 2^r which is irreducible [5].

The tensor product of representations and decomposition under subgroups can be found in standard tables, like the appendices in [124].

B.6.4 Comments on $USp(2n)$

We will not say much about these, since these groups rarely appear in particle physics or in string theory. Moreover, most of its properties can be derived from the trick that it can be constructed from $U(2n)$ by keeping the subset of roots invariant under an involution. We will see more of this as we need it.

[5] This underlies the fact that there are no chiral spinors in euclidean spaces of odd dimension.

B.6.5 Comments on exceptional groups

The most interesting one is E_8, since it appears automatically in the construction of the heterotic superstring. Moreover, properties of E_6, E_7 etc are easy to derive since they are subgroups of E_8. For details we refer to the properties listed in tables like the appendices in [124].

For the moment, the only data we need is the root system of E_8. This has rank 8 and dimension 248, and the 240 non-zero roots are of the form

$$\begin{array}{l} (\underline{\pm,\pm,0,0,0,0,0,0}) \\ (\pm\tfrac{1}{2},\pm\tfrac{1}{2},\pm\tfrac{1}{2},\pm\tfrac{1}{2},\pm\tfrac{1}{2},\pm\tfrac{1}{2},\pm\tfrac{1}{2},\pm\tfrac{1}{2}) \quad , \quad \#- = \text{even} \end{array} \tag{B.83}$$

Notice that there is a nice subset of $SO(16)$ roots, given by the first line of non-zero roots (along with the 8 Cartan generators). With respect to this $SO(16)$ subalgebra, the states associated with the vectors in the second line are transforming in a 2^{8-1}-dimensional chiral spinor representation of $SO(16)$.

We will find good application of these facts for instance in the identification of the spectrum of the heterotic theory.

Appendix C

Appendix: Rudiments of Supersymmetry

In this appendix we provide the basic ideas on the construction of supersymmetric field theories. The emphasis is in providing some basic results to be used in the general lectures. We mainly follow the notation and discussion in [125], to which we refer the reader interested in more details and proofs. For useful tables of supermultiplet components, for diverse extended supersymmetries in diverse dimensions, see [127, 126].

C.1 Preliminaries: Spinors in 4d

Before discussing supersymmetry, it is useful to briefly review two-component 4d spinors (Weyl spinors), their properties, some useful notation, and their relation to the more familiar four-component Dirac spinors. It is important to realize that the following discussion has nothing to do with supersymmetry, but just with spinor representations of the 4d Lorentz group, and that two-component spinors appear in many contexts, for instance in the Standard Model.

The 4d Lorentz group contains two inequivalent spinor representations, usually denoted left- and right-handed spinors. These representations are two-dimensional, so the spinors are denoted two-component, and sometime Weyl spinors. The two representations are exchanged under (Dirac) conjugation (transposition and complex conjugation), namely the conjugate of a left-handed object transforms as a right-handed spinor.

We use the following notation, we denote a left-handed spinor as ψ_α, a right-handed spinor as $\bar{\psi}^{\dot{\alpha}}$. We also denote the conjugate of a right-handed spinor by ψ^α and the conjugate of a left-handed spinor by $\bar{\psi}_{\dot{\alpha}}$.

A Lorentz transformation is represented on spinors in terms of a matrix M in $SL(2,\mathbf{C})$ (Notice that it contains six independent real parameters).

Spinors transform as

$$\begin{aligned} \psi'_\alpha = M_\alpha{}^\beta \psi_\beta \quad &; \qquad \bar{\psi}'_{\dot{\alpha}} = (M^*)_{\dot{\alpha}}{}^{\dot{\beta}} \bar{\psi}_{\dot{\beta}} \\ \psi'^\alpha = \psi^\beta (M^{-1})_\beta{}^\alpha \quad &; \qquad \bar{\psi}'^{\dot{\alpha}} = \bar{\psi}^{\dot{\beta}} (M^{*-1})_{\dot{\beta}}{}^{\dot{\alpha}} \end{aligned} \tag{C.1}$$

Namely, ψ_α and ψ^α are rotated by M as column and row vectors, while $\bar{\psi}_{\dot{\alpha}}$ and $\bar{\psi}^{\dot{\alpha}}$ are rotated by M^*.

Thus, contractions of the form $(\ldots)^\alpha(\ldots)_\alpha$ and $(\ldots)_{\dot{\alpha}}(\ldots)^{\dot{\alpha}}$ are invariant.

Vector representations can be constructed from the spinor representations. For that purpose, we introduce the matrices $\sigma^\mu{}_{\alpha\dot{\alpha}}$

$$\sigma^0 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \quad ; \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad ; \quad \sigma^0 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad ; \quad \sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{C.2}$$

Considering linear combinations of the form $P = P_\mu \sigma^\mu$, the inherited action of M is

$$P'_{\alpha\dot{\alpha}} = M_\alpha{}^\beta P_{\beta\dot{\beta}} (M^*)^{\dot{\alpha}\dot{\beta}} = (MPM^\dagger)_{\alpha\dot{\alpha}} \tag{C.3}$$

Indeed this is a Lorentz transformation on the 4-vector (P_μ), since the transformation preserves $\det P = -[-(P_0)^2 + (P_1)^2 + (P_2)^2 + (P_3)^2]$, which is precisely (minus) the norm of P_μ. Hence, any vector can be expressed in terms of bi-spinor components (and vice-versa).

It is useful to introduce the tensors

$$(\epsilon^{\alpha\beta}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad ; \qquad (\epsilon_{\alpha\beta}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{C.4}$$

(and similarly for dotted indices). They are Lorentz invariant, namely they satisfy

$$\epsilon_{\alpha\beta} \,=\, M_\alpha{}^\gamma\, M_\beta{}^\delta\, \epsilon_{\gamma\delta} \qquad ; \qquad \epsilon^{\alpha\beta} \,=\, \epsilon^{\gamma\delta}\, (M^{-1})_\gamma{}^\alpha\, (M^{-1})_\delta{}^\beta \tag{C.5}$$

as may be checked by using their explicit expressions.

These properties imply that the tensors can be used to raise and lower indices

$$\psi^\alpha = \epsilon^{\alpha\beta}\psi_\beta \qquad ; \qquad \psi_\alpha = \epsilon_{\alpha\beta}\psi^\beta \tag{C.6}$$

(and similarly for dotted indices). What this means is that e.g. the object $\epsilon^{\alpha\beta}\psi^\beta$ transforms as an object $()^\alpha$, (i.e. as a column vector on which M acts), which we denote ψ^α. We introduce the shorthand notation

$$\chi\psi = \chi^\alpha\psi_\alpha \qquad ; \qquad \bar{\chi}\bar{\psi} = \chi_{\dot{\alpha}}\psi^{\dot{\alpha}} \tag{C.7}$$

Using the ϵ tensors, we can also define

$$(\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} = \epsilon^{\dot{\alpha}\dot{\beta}}\epsilon^{\alpha\beta}\sigma^\mu{}_{\beta\dot{\beta}} \tag{C.8}$$

They satisfy

$$(\sigma^\mu\bar{\sigma}^\nu + \sigma^\nu\bar{\sigma}^\mu)_\alpha{}^\beta = -2\eta^{\mu\nu}\delta_\alpha{}^\beta \quad ; \quad (\bar{\sigma}^\mu\sigma^\nu + \bar{\sigma}^\nu\sigma^\mu)^{\dot{\alpha}}{}_{\dot{\beta}} = -2\eta^{\mu\nu}\delta^{\dot{\alpha}}{}_{\dot{\beta}} \tag{C.9}$$

In terms of them, the generators of the Lorentz group are given by

$$(\sigma^{\mu\nu})_\alpha{}^\beta = \frac{1}{4}[\sigma^\mu_{\alpha\dot{\alpha}}\bar{\sigma}^{\nu\dot{\alpha}\beta} - \sigma^\nu_{\alpha\dot{\alpha}}\bar{\sigma}^{\mu\dot{\alpha}\beta}] \quad ; \quad (\bar{\sigma}^{\mu\nu})^{\dot{\alpha}}{}_{\dot{\beta}} = \frac{1}{4}[\dot{\sigma}^{\mu\dot{\alpha}\alpha}\sigma^\nu{}_{\alpha\dot{\beta}} - \dot{\sigma}^{\nu\dot{\alpha}\alpha}\sigma^\mu{}_{\alpha\dot{\beta}}] \tag{C.10}$$

Given two Weyl spinors of opposite chiralities χ_α, $\bar{\psi}^{\dot{\alpha}}$ (and equal global and gauge quantum numbers), one can construct a four-component Dirac spinor by superposing them as a column vector

$$\Psi_D = \begin{pmatrix} \chi_\alpha \\ \bar{\psi}^{\dot{\alpha}} \end{pmatrix} \tag{C.11}$$

on which the Dirac matrices are realized as

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \tag{C.12}$$

which satisfy the Clifford algebra relations, as follows from (C.9). Also, given a single Weyl spinor, say χ_α, in a real representation of all all global and gauge symmetries, one can construct a four-component fermion, by taking its conjugate to play the role of the right-handed piece, as follows

$$\Psi_M = \begin{pmatrix} \chi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix} \tag{C.13}$$

Such spinors Ψ_M are thus subject to a reality condition, and are denoted Majorana. Notice that Weyl spinors in complex representations of the global or gauge symmetries cannot be turned into Majorana spinors, since the spinor and its conjugate cannot belong to the same multiplet.

C.2 4d $N = 1$ Supersymmetry algebra and representations

In this section we discuss the basic structure of 4d $N = 1$ supersymmetry algebra, and its realization in terms of fields.

C.2.1 The supersymmetry algebra

The 4d $N = 1$ supersymmetry algebra contains two spinorial generators Q_α, $\bar{Q}_{\dot{\alpha}}$, which behave as Grassman variables, and hence obey anticommutation relations. The algebra is given by

$$\begin{aligned} \{Q_\alpha, \bar{Q}_{\dot{\alpha}}\} &= 2\sigma^\mu{}_{\alpha\dot{\alpha}} P_\mu \\ \{Q_\alpha, Q_\beta\} &= \{\bar{Q}_{\dot{\alpha}}, \bar{Q}_{\dot{\beta}}\} = 0 \\ \{P_\mu, Q_\alpha\} &= \{P_\mu, \bar{Q}_{\dot{\alpha}}\} = 0 \end{aligned} \qquad (C.14)$$

(in addition, we have the natural commutators that imply that the Q's are in the spinor representations).

OBS: The above algebra is invariant under $U(1)$ transformations rotating the supercharges Q_α, $\bar{Q}_{\dot{\alpha}}$ by opposite phases.

$$Q_\alpha \to e^{i\lambda} Q_\alpha \quad ; \quad \bar{Q}_{\dot{\alpha}} \to e^{-i\lambda} \bar{Q}_{\dot{\alpha}} \qquad (C.15)$$

This symmetry is known as R-symmetry.

Since the supergenerators Q_α, $\bar{Q}_{\dot{\alpha}}$, are Grassman quantities, when realized on quantum fields they relate bosons and fermions. Each multiplet providing a representation of the supersymmetry algebra (supermultiplet) thus contains bosons and fermions. Since the operator P^2, which is the mass square operator, commutes with the Q's, bosons and fermions in the same multiplet are mass degenerate. Similarly, the supergenerators commute with any global and gauge symmetry of the theory [1], so all fields in a supermultiplet belong to the same representation of global and gauge symmetries.

An important property is that the total number of *physical* bosonic and fermionic degrees of freedom is equal within a supermultiplet. To show this, we define the operator $(-1)^F$, which is equal to $+1$ for bosons and -1 for fermions, and hence satisfies $(-1)^F Q_\alpha = -Q_\alpha (-1)^F$. We can then compute,

[1] Except for R-symmetries, see below.

in two different ways, $\text{Tr}\,[(-1)^F\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\}]$, where the trace is taken over states of fixed momentum in a supermultiplet,

$$
\begin{array}{lll}
1) & \text{Tr}\,[(-1)^F\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\}] & = \text{Tr}\,[(-1)^F Q_\alpha \bar{Q}_{\dot{\alpha}} + (-1)^F \bar{Q}_{\dot{\alpha}} Q_\alpha] = \\
 & & = \text{Tr}\,[-Q_\alpha(-1)^F \bar{Q}_{\dot{\alpha}} + Q_\alpha(-1)^F \bar{Q}_{\dot{\alpha}}] = 0 \\
2) & \text{Tr}\,[(-1)^F\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\}] & = 2\sigma^\mu{}_{\alpha\dot{\alpha}} P_\mu \text{Tr}\,[(-1)^F]
\end{array}
\tag{C.16}
$$

Hence $\text{Tr}\,[(-1)^F] = 0$ in a supermultiplet.

C.2.2 Structure of supermultiplets

Let us consider the construction of the supermultiplet for massive fields of mass M. Going to the rest frame for such particles, the relevant piece of the algebra (C.14) becomes

$$
\begin{array}{rcl}
\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\} & = & 2M\delta_{\alpha\dot{\alpha}} \\
\{Q_\alpha, Q_\beta\} & = & \{\bar{Q}_{\dot{\alpha}}, \bar{Q}_{\dot{\beta}}\} = 0
\end{array}
\tag{C.17}
$$

By defining $a_\alpha = Q_\alpha/\sqrt{2M}$, $a_\alpha^\dagger = \bar{Q}_{\dot{\alpha}}/\sqrt{2M}$, these are the anticommutators for two decoupled fermionic harmonic oscillators. The supermultiplet is built by starting with a lowest helicity state $|\Omega\rangle$, obeying $a_\alpha|\Omega\rangle = 0$, and appying operators $a_\alpha^\dagger$, namely

State	Helicity
$\|\Omega\rangle$	j
$a_\alpha^\dagger\|\Omega\rangle$	$j \pm \frac{1}{2}$
$a_1^\dagger a_2^\dagger\|\Omega\rangle$	j

In building a quantum field theory with the corresponding fields, it is important to notice that CPT flips the chirality (and conjugates the global and gauge representations), so a CPT-invariant supermultiplet may require using two of the above basic multiplets.

Two of the most useful supermultiplets are the following:

- The massive scalar supermultiplet is obtained by starting with a $j = 0$ state $|\Omega\rangle$. It contains states of helicities 0, $\pm 1/2$, 0. It thus contains a Weyl spinor and a complex scalar. This is CPT-invariant if the supermultiplet belongs to a real representation of the gauge and global symmetries. If not, two of these basic multiplets, in conjugate representations, must be combined to form a CPT-invariant set.

- The massive vector multiplet is obtained by starting with a $j = 1/2$ state $|\Omega\rangle$. It contains states of helicities 1/2, 1, 0, 1/2. Combining it with its CPT conjugate, the total multiplet contains one massive vector boson, one real scalar and two Weyl fermions.

Let us now consider the construction of supermultiplets for massless fields. Since they have light-like momentum $P^2 = 0$, they do not have rest frame, but we may use a reference system where $P = (-E, 0, 0, E)$. In this frame, the supersymmetry algebra is

$$\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\} = 2\begin{pmatrix} 2E & 0 \\ 0 & 0 \end{pmatrix} \tag{C.18}$$

Defining the rescaled operators

$$a = \frac{1}{2\sqrt{E}} Q_1 \qquad ; \qquad a^\dagger = \frac{1}{2\sqrt{E}} \bar{Q}_{\dot{1}} \tag{C.19}$$

they correspond to a fermionic harmonic oscillator. The multiplet is constructed by starting with a lowest helicity state $|\Omega\rangle$, satisfying

$$a|\Omega\rangle = Q_2|\Omega\rangle = \bar{Q}_{\dot{2}}|0\rangle = 0 \tag{C.20}$$

Hence the multiplet contains the states $|\Omega\rangle$ and $a^\dagger|\Omega\rangle$, with helicities j and $j+1/2$, respectively. As before, one may need to combine this multiplet with its CPT conjugate to formulate a quantum field theory.

Some of the most useful massless supermultiplets are:

- The chiral supermultiplet, obtained by taking $|\Omega\rangle$ of helicity $j = 0$, so it contains states of helicity $j = 0, 1/2$. This should be combined with its CPT conjugate, with helicities $j = 0, -1/2$. This complete chiral supermultiplet contains a complex scalar and a 4d Weyl fermion. This multiplet can transform in an arbitrary representation of the gauge and global symmetries, hence contains a chiral fermion, which is necessarily massless. If the multiplet happens to transform in a real representation, it is possible to write a mass term for it (see later), so it is equivalent to a massive scalar supermultiplet.

- The massless vector supermultiplet, obtained by taking $|\Omega\rangle$ of helicity $j = 1/2$, so it contains states of helicities $j = 1/2, 1$. Combined with its CPT conjugate, with helicities $j = -1, -1/2$, the multiplet contains a 4d Weyl spinor and a massless vector boson. The multiplet transforms in the adjoint representation of the gauge group, which is real, so the 4d Weyl spinor can be recast as a 4d Majorana spinor.

- The supergravity multiplet, containing states of helicity $j = 3/2, 2$. Combined with its CPT conjugate, of helicities $j = -2, -3/2$, it contains a graviton and a gravitino (a spin 3/2 particle). We will not discuss it in detail, since interacting theories involving this multiplet have spacetime diffeomorphism invariance, and include gravity (and in fact local supersymmetry), they are known as supergravity theories, and lie beyond the scope of this lecture

C.3 Component fields, chiral multiplet

The supersymmetry transformation parameters are anticommuting spinors ξ^α, $\bar{\xi}_{\dot\alpha}$. Formally, the supersymmetry variation

is $\delta_\xi = \xi Q + \bar{\xi}\bar{Q}$. The supersymmetry algebra can be expressed as

$$\begin{aligned} [\xi Q, \bar{\eta}\bar{Q}] &= 2\,\xi\sigma^\mu\bar{\eta}\,P_\mu \\ [\xi Q, \eta Q] &= [\bar{\xi}\bar{Q}, \bar{\eta}\bar{Q}] = 0 \end{aligned} \qquad \text{(C.21)}$$

We would like to construct a representation of the supersymmetry algebra, using the massive scalar multiplet, which contains as *physical* degrees of freedom a 4d Weyl spinor ψ_α and a complex scalar Φ. The supersymmetry transformations of these fields are

$$\begin{aligned} \delta_\xi \Phi &= \sqrt{2}\,\xi\psi \\ \delta_\xi \psi_\alpha &= i\sqrt{2}\,\sigma^\mu{}_{\alpha\dot\alpha}\bar{\xi}^{\dot\alpha}\,\partial_\mu\Phi + \sqrt{2}\,\xi_\alpha F \end{aligned} \qquad \text{(C.22)}$$

Namely

$$\begin{aligned} \bar{Q}_{\dot\alpha}\Phi = 0 \qquad\qquad & Q_\alpha\Phi = \sqrt{2}\psi_\alpha \\ \bar{Q}_{\dot\alpha}\psi_\alpha = -i\sqrt{2}\sigma^\mu{}_{\alpha\dot\alpha}\partial_\mu\Phi \qquad\qquad & Q_\alpha\psi_\beta = \sqrt{2}\epsilon_{\alpha\beta}F \end{aligned} \qquad \text{(C.23)}$$

The field F appearing in the transformation of the fermions is discussed below.

The transformations acting on Φ satisfy the supersymmetry algebra. In order for the transformations acting on ψ to satisfy the supersymmetry algebra, we have two choices

i) Take $F = -m\Phi^*$, and use the equation of motion of a free massive fermion for ψ, namely $-i\bar{\sigma}^\mu \partial_\mu \psi = m\psi$. Since we are using equations of motion, the algebra closes on-shell.

ii) Consider F to be and independent field, and require $\delta_\xi F = i\sqrt{2}\bar{\xi}\bar{\sigma}^\mu \partial_\mu \psi$. Since the equations of motion are not involved, the algebra closes off-shell.

Notice that the viewpoint i) is disadvantageous, since the equations of motion are different for different theories, and this complicates the construction of interacting theories. On the other hand, from the viewpoint ii) the transformations obey the supersymmetry algebra relations, no matter what the dynamics of the theory is. It is important to notice that the field F does not really describe a new physical degree of freedom. Since the dimension of ξ, $\bar{\xi}$ is 1/2, F has dimension 2, and it is not possible to write a kinetic term for it, and it is called an auxiliary field. Hence we still have equality of the number of bosonic and fermionic *physical* degrees of freedom in the supermultiplet.

In principle, one can construct supersymmetry transformations for fields in other supermultiplets. However it is non-trivial to do so for more complicated supermultiplets. The task is facilitated by a technique, known as superfield formalism.

C.4 Superfields

C.4.1 Superfields and supersymmetry transformations

Let us consider the set of component fields in a supermultiplet. Since they form an irreducible representation, the whole set can be generated from any one of them, say A, by acting with the supergenerators. It is useful to consider the following formal expression

$$F(x,\theta,\bar{\theta}) = e^{\theta Q + \bar{\theta}\bar{Q}} A \tag{C.24}$$

Different component fields in the supermultiplet arise as coefficients in the power-expansion of F in θ, $\bar{\theta}$. Since the latter are Grassman variables, the power-expansion is a finite expression, of the form

$$\begin{aligned} F(x,\theta,\bar{\theta}) \;&=\; f(x) \,+\, \theta\phi(x) \,+\, \bar{\theta}\bar{\xi}(x) \,+\, \theta\theta\, m(x) \,+\, \bar{\theta}\bar{\theta}\, n(x) \,+ \\ &+\; \theta\sigma^\mu\bar{\theta}\, v_\mu(x) \,+\, \theta\theta\,\bar{\theta}\bar{\lambda}(x) \,+\, \bar{\theta}\bar{\theta}\,\theta\psi(x) \,+\, \theta\theta\,\bar{\theta}\bar{\theta}\, d(x) \end{aligned} \tag{C.25}$$

where all the fields are related to each other by the action of Q, $\bar{Q}$. Expressions of the form (C.25), providing a formal sum of the component fields in a supermultiplet, are refered to as superfields. Formally, they are functions over a superspace parametrize by the supercoodinates $z=(x,\theta,\bar{\theta})$. A whole branch of mathematical physics is the study of the geometry of superspace (supergeometry), but we will not need much of its machinery.

The use of superfields facilitates the computation of supersymmetry transformations of the component fields. Let us introduce a formal sum of such variations

$$\begin{aligned}\delta_\xi F(x,\theta,\bar{\theta}) &= \delta_\xi f(x) + \theta\delta_\xi\phi(x) + \bar{\theta}\delta_\xi\bar{\xi}(x) + \theta\theta\,\delta_\xi m(x) + \bar{\theta}\bar{\theta}\,\delta_\xi n(x) + \\ &+ \theta\sigma^\mu\bar{\theta}\,\delta_\xi v_\mu(x) + \theta\theta\,\bar{\theta}\delta_\xi\bar{\lambda}(x) + \bar{\theta}\bar{\theta}\,\theta\delta_\xi\psi(x) + \theta\theta\,\bar{\theta}\bar{\theta}\,\delta_\xi d(x) \quad \text{(C.26)}\end{aligned}$$

We formally write $\delta_\xi F \equiv (\xi Q + \bar{\xi}\bar{Q}) \times F$. The operation $(\xi Q + \bar{\xi}\bar{Q})\times$ thus maps a superfield to the superfield constructed using the susy variations of the component fields. Notice that it does not interfere with the θ, $\bar{\theta}$.

We would like to represent the action of $(\xi Q+\bar{\xi}\bar{Q})\times$ in terms of differential operators in superspace. The simplest operators in superspace are derivatives $\partial_\alpha = \frac{\partial}{\partial\theta^\alpha}$ and $\bar{\partial}_{\dot{\alpha}} = \frac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}}$ (in addition to the familiar $\partial_\mu = \frac{\partial}{\partial x^\mu}$. Using Hausdorff formula, $e^{A+B} = e^A e^B e^{-[A,B]/2}$ (for A, B, commuting with $[A,B]$), we have

$$\begin{aligned}\xi^\alpha\partial_\alpha\,(\,e^{\theta Q+\bar{\theta}\bar{Q}}\times\,) &= \xi^\alpha\partial_\alpha\,e^{\theta Q}e^{\bar{\theta}\bar{Q}}e^{-\theta\sigma^\mu\bar{\theta}P_\mu}\times = \\ &= (\xi Q + i\sigma^\mu\bar{\theta}\partial_\mu)\times e^{\theta Q+\bar{\theta}\bar{Q}}\times \\ \bar{\xi}_{\dot{\alpha}}\bar{\partial}^{\dot{\alpha}}\,e^{\theta Q+\bar{\theta}\bar{Q}}\times &= (\bar{\xi}\bar{Q} - i\theta\sigma^\mu\bar{\xi}\partial_\mu)\times e^{\theta Q+\bar{\theta}\bar{Q}}\times \end{aligned} \quad \text{(C.27)}$$

From this we learn that the action of ξQ, $\bar{\xi}\bar{Q}$ on component fields can be represented in terms of differential operators acting on superfields. By abuse of notation, these differential operators are also denoted Q_α and $\bar{Q}_{\dot{\alpha}}$

$$\begin{aligned}Q_\alpha &= \partial_\alpha - i\sigma^\mu{}_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\partial_\mu \\ \bar{Q}_{\dot{\alpha}} &= \bar{\partial}_{\dot{\alpha}} - i\theta^\alpha\sigma^\mu{}_{\alpha\dot{\alpha}}\partial_\mu\end{aligned} \quad \text{(C.28)}$$

Namely, given a superfield $F(x,\theta,\bar{\theta})$, we can compute the supersymmetry variation of its components, which are encoded in the superfield of variations (C.26) $\delta_\xi F = (\xi Q + \bar{\xi}\bar{Q}) \times F$, by computing $\delta_\xi F$ using the action of the differential operators (C.28), namely $\delta_\xi F = (\xi Q + \bar{\xi}\bar{Q})F$. Comparing terms in both θ-expansions leads to the supersymmetry variations.

An important observation is that the component field corresponding to highest power in θ, $\bar{\theta}$ in the expansion, always transforms as a total divergence. This is because θ, $\bar{\theta}$ have dimension $-1/2$, so that this component field is the one of highest dimension in the supermultiplet. On the other hand, the supergenerators Q_α, $\bar{Q}_{\dot{\alpha}}$, have dimension 1/2. Thus the supersymmetry variation of the highest-dimension component field is necessarily the derivative of a lower-dimension component field. This observation will be the key idea in the construction of supersymmetric field theory actions.

Superfields are useful since they provide linear representations of the supersymmetry algebra. Actually, a completely general superfield corresponds to a reducible representation. Different irreducible representations correspond to superfields satisfying different constraints, consistent with the action of the operators (C.28). This will be discussed below. For that purpose, it is useful to define the differential operators

$$D_\alpha = \partial_\alpha + i\sigma^\mu{}_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\partial_\mu \quad ; \quad \bar{D}_{\dot{\alpha}} = -\bar{\partial}_{\dot{\alpha}} - i\theta^\alpha\sigma^\mu{}_{\alpha\dot{\alpha}}\partial_\mu \tag{C.29}$$

They anticommute with the operators (C.28)

$$\{D_\alpha, Q_\beta\} = \{D_\alpha, \bar{Q}_{\dot{\beta}}\} = \{\bar{D}_{\dot{\alpha}}, Q_\beta\} = \{\bar{D}_{\dot{\alpha}}, \bar{Q}_{\dot{\beta}}\} = 0 \tag{C.30}$$

C.4.2 The chiral superfield

A chiral superfield $\Phi(x, \theta, \bar{\theta}$ is characterized by $\bar{D}_{\dot{\alpha}}\Phi = 0$. It is useful to describe it in terms of a new position variable $y^\mu = x^\mu + i\theta\sigma^\mu\bar{\theta}$, in terms of which the differential operators (C.29) read

$$D_\alpha = \frac{\partial}{\partial\theta^\alpha} + 2i\sigma^\mu{}_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\frac{\partial}{\partial y^\mu} \quad ; \quad \bar{D}_{\dot{\alpha}} = -\frac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}} \tag{C.31}$$

Hence a chiral superfield has the expansion

$$\Phi(y, \theta, \bar{\theta}) = \Phi(y) + \sqrt{2}\theta\psi(y) + \theta\theta F(y) \tag{C.32}$$

We can readily identify that this describes a chiral (or scalar) supermultiplet (by abuse of language, one often uses the same notation for the superfield and for its complex scalar component field, hoping the context will disentangle any possible ambiguity). Indeed we can reproduce the supersymmetry transformations of the component fields, by using the differential operators (C.28), which in these coordinates read

$$Q_\alpha = \frac{\partial}{\partial\theta^\alpha} \quad ; \quad \bar{Q}_{\dot\alpha} = \frac{\partial}{\partial\bar\theta^{\dot\alpha}} - 2i\theta^\alpha \sigma^\mu{}_{\alpha\dot\alpha}\frac{\partial}{\partial y^\mu} \tag{C.33}$$

and comparing

$$\begin{aligned}(\xi Q + \bar\xi\bar Q)\times\Phi(y,\theta,\bar\theta) &= \delta_\xi\Phi(y) + \sqrt{2}\theta^\alpha\delta_\xi\psi_\alpha(y) + \theta\theta\delta_\xi F(y) \qquad (C.34)\\ (\xi Q + \bar\xi\bar Q)\Phi(y,\theta,\bar\theta) &= \xi^\alpha\frac{\partial}{\partial\theta^\alpha}\Phi(y,\theta,\bar\theta) + (\frac{\partial}{\partial\bar\theta^{\dot\alpha}} - 2i\theta^\alpha\sigma^\mu{}_{\alpha\dot\alpha}\frac{\partial}{\partial y^\mu})\xi^{\dot\alpha}\Phi(y,\theta,\bar\theta) = \\ &= \sqrt{2}\xi\psi + \sqrt{2}\theta^\alpha(-i\sqrt{2}\sigma^\mu{}_{\alpha\dot\alpha}\bar\xi^{\dot\alpha}\partial_\mu\Phi + \xi_\alpha F) + \theta\theta\, i\sqrt{2}\bar\xi\bar\sigma^\mu\partial_\mu\psi\end{aligned}$$

In terms of the original coordinates, we have

$$\begin{aligned}\Phi(x,\theta,\bar\theta) &= \Phi(x) + i\theta\sigma^\mu\bar\theta\,\partial_\mu\Phi(x) + \frac{1}{4}\theta\theta\,\bar\theta\bar\theta\,\Box\,\Phi(x) + \\ &+\sqrt{2}\theta\psi(x) - \frac{i}{\sqrt{2}}\theta\theta\partial_\mu\psi(x)\sigma^\mu\bar\theta + \theta\theta F(x)\end{aligned} \tag{C.35}$$

Notice that the highest-dimension component is the same, expressed in terms of x or y.

An antichiral field satisfies the condition that D_α annihilates it. Clearly the the adjoint superfield $\Phi^\dagger$ of a chiral superfield is antichiral. In terms of x, θ, $\bar\theta$, it reads

$$\begin{aligned}\Phi^\dagger(x,\theta,\bar\theta) &= \Phi^*(x) - i\theta\sigma^\mu\bar\theta\,\partial_\mu\Phi^*(x) + \frac{1}{4}\theta\theta\bar\theta\bar\theta\Box\,\Phi^*(x) + \\ &+\sqrt{2}\bar\theta\bar\psi(x) + \frac{i}{\sqrt{2}}\bar\theta\bar\theta\theta\sigma^\mu\partial_\mu\bar\psi(x) + \theta\theta F^*(x)\end{aligned} \tag{C.36}$$

The supermultiplet has a simpler expression in terms of the variable $y^{\dagger\mu} = x^\mu - i\theta\sigma^\mu\bar\theta$, it reads

$$\Phi^\dagger(y^\dagger,\theta,\bar\theta) = \Phi^*(y^\dagger) + \sqrt{2}\bar\theta\bar\psi(y^\dagger) + \theta\theta F^*(y^\dagger) \tag{C.37}$$

An important property of chiral multiplets is that their product is also a chiral superfield. This is straightforward using the expression in terms of y coordinates. By using power-series, one can show that any holomorphic function of chiral multiplets $W(\Phi_k(x,\theta,\bar\theta)$ is also a chiral multiplet. For future convenience, one can show that its highest-dimension component is given by

$$W(\Phi)|_{\theta\theta} \;=\; \frac{\partial^2 W}{\partial\Phi_k\partial\Phi_l}\psi_k\psi_l + F_k\left(\frac{\partial W}{\partial\Phi_k}\right)^* + F_k^*\frac{\partial W}{\partial\Phi_k} \tag{C.38}$$

where in the right-hand side Φ denotes the scalar component field, not the superfield.

On the other hand, non-holomorphic functions like $\Phi^\dagger\Phi$ are not chiral superfields. For future convenience, we list the highest-dimension component of the latter

$$\begin{aligned}\Phi_1^\dagger(x,\theta,\bar\theta)\Phi_2(x,\theta,\bar\theta)|_{\theta\theta\bar\theta\bar\theta} \;=\;& F_1^*F_2 + \frac{1}{4}\Phi_1^*\Box\,\Phi_2 + \frac{1}{4}\Box\,\Phi_1^*\,\Phi_2 - \frac{1}{2}\partial_\mu\Phi_1^*\partial^\mu\Phi_2 + \\ &+\frac{i}{2}\partial_\mu\bar\psi_1\bar\sigma^\mu\psi_2 - \frac{i}{2}\bar\psi_1\bar\sigma^\mu\partial_\mu\psi_2\end{aligned} \tag{C.39}$$

We are now ready to construct supersymmetric lagragians for fields in chiral supermultiplets. The key idea is that, since the highest-dimensional component of a supermultiplet (usually a product of basic supermultiplets) transforms as a total derivative, its spacetime integral is invariant under supersymmetry transformations. The strategy then is to construct product superfields whose highest-dimensional component corresponds to kinetic and interactions terms. Finally, recalling the rules of integration over Grassman variables,

$$\int d\theta = 0 \qquad ; \qquad \int d\theta\,\theta = 1 \tag{C.40}$$

an efficient way to extract the highest component of a supermultiplet is to integrate it over the supercoordinates θ and/or $\bar\theta$. For instance

$$\int d^2\theta\Phi(x,\theta,\bar\theta) = F(x) \tag{C.41}$$

A typical supersymmetric action for a set of chiral supermultiplets has the structure

$$S \;=\; \int d^4x\,d^2\theta\,d^2\bar\theta\,\Phi_i^\dagger\Phi_i \,+\, \int d^4x\,d^2\theta\,W(\Phi_i) \,+\, \int d^4x\,d^2\bar\theta\,W(\Phi_i)^* \tag{C.42}$$

The first term can be generalized to $\int d^4x\,d^2\theta\,d^2\bar\theta\,K(\Phi_i,\Phi_i^\dagger)$, with K a real function, known as Kahler potential. Expanding in components, this implies that the space parametrized by scalars in chiral multiplets is Kahler (in the geometric sense). We will however stick to the canonical kinetic term above, but occasionally refer to these more general possible actions.

Using (C.38), (C.39), the action in component fields reads (integrating by parts in certain terms)

$$S = -\left[\partial_\mu\Phi_i^*\partial^\mu\Phi_i + i\bar{\psi}_i\bar{\sigma}^\mu\partial_\mu\psi_i - F_i^*F_i - \frac{\partial^2 W}{\partial\Phi_i\partial\Phi_j}\psi_i\psi_j - F_i\left(\frac{\partial W}{\partial\Phi_i}\right)^* - F_i^*\frac{\partial W}{\partial\Phi_i}\right] \quad (C.43)$$

We see that the auxiliary fields F_i are indeed non-dynamical. We can use their equations of motion, to obtain $F_i = -\partial W/\partial\Phi_i$. Replacing in the above expression, we have

$$S = -\left[\partial_\mu\Phi_i^*\partial^\mu\Phi_i + i\bar{\psi}_i\bar{\sigma}^\mu\partial_\mu\psi_i - \sum_i\left|\frac{\partial W}{\partial\Phi_i}\right|^2 - \frac{\partial^2 W}{\partial\Phi_i\partial\Phi_j}\psi_i\psi_j\right] \quad (C.44)$$

The first two pieces are standard kinetic terms. The fourth describes scalar-fermion interactions, and the third is a scalar potential

$$V(\Phi_i) = \sum_i\left|\frac{\partial W}{\partial\Phi_i}\right|^2 \quad (C.45)$$

It is positive-definite, and vanishes for scalar vevs such that

$$F_i = -\frac{\partial W}{\partial\Phi_i} = 0 \quad (C.46)$$

These are know as F-term constraints, which are a necessary condition for a supersymmetric vacuum of the theory.

An important property of supersymmetric field theories is that the superpotential is not renormalized in perturbation theory. That is, because of the relations imposed by supersymmetry, all radiative corrections to the terms arising from the superpotential vanish to all orders in perturbation theory. The proof of this statements involves the structure of Feynman diagrams in superspace, and we will not discuss it. In particular examples (for instance for the Wess-Zumino model, i.e. a theory with one chiral multiplet and a cubic superpotential), one can show it very explicitly exploiting the holomorphy of the superpotential, see [?] for detailed discussion. Both arguments show that there are important non-renormalization theorems involving terms in the action which involve intergration over half the superspace coordinates. Another important observation is that the non-renomalization theorem in general does not hold beyond perturbation theory, hence non-perturbative corrections to the superpotential may appear. In some situations, they may be exactly computable using the constraints from supersymmetry and reasonable assumptions about the field theory dynamics. These non-perturbative corrections usually have a nice physical interpretation (like instanton effects or gaugino condensation). See [?] for more complete discussion.

C.4.3 The vector superfield

A vector superfield V is characterized by the condition $V = V^\dagger$. The expansion in component fields can be expressed as

$$\begin{aligned} V(x,\theta,\bar\theta) &= C(x) + i\theta\chi(x) i\bar\theta\bar\chi(x) + \frac{i}{2}\theta\theta M(x) - \frac{i}{2}\bar\theta\bar\theta M^*(x) - \theta\sigma^\mu\bar\theta\, V_\mu(x) + \\ &+ \ i\theta\theta\,\bar\theta\,[\,\bar\lambda(x) + \frac{i}{2}\bar\sigma^\mu\partial_\mu\xi(x)\,] - i\bar\theta\bar\theta\,\theta\,[\,\lambda(x) + \frac{i}{2}\sigma^\mu\partial_\mu\bar\xi(x)\,] + \frac{1}{2}\theta\theta\,\bar\theta\bar\theta\, D(x) \end{aligned} \tag{C.47}$$

The peculiar choice of components in the $\theta^2\bar\theta$, $\bar\theta^2\theta$ and $\theta^2\bar\theta^2$ terms, is for future convenience.

As we will see, the content of component fields of the vector superfield is that of a massless vector superfield. Thus, it should describe the supersymmetric version of a gauge boson. Hence there is a supersymmetric version of a gauge transformation. For vector multiplets associated to $U(1)$, it is given by

$$V \longrightarrow V + (\Lambda + \Lambda^\dagger) \tag{C.48}$$

where $\Lambda(y,\theta,\bar\theta) = A + \sqrt{2}\theta\psi + \theta\theta F$ is a chiral superfield. Since

$$\begin{aligned} \Lambda + \Lambda^\dagger &= \Lambda + \Lambda^* + \sqrt{2}(\theta\psi + \bar\theta\bar\psi) + \theta\theta F + \bar\theta\bar\theta F^* + i\theta\sigma^\mu\bar\theta\partial_\mu(\Lambda - \Lambda^*) + \\ &+ \frac{i}{\sqrt{2}}\theta\theta\bar\theta\bar\sigma^\mu\partial_\mu\psi + + \frac{i}{\sqrt{2}}\bar\theta\bar\theta\theta\sigma^\mu\partial_\mu\bar\psi + \frac{1}{4}\theta\theta\bar\theta\bar\theta\Box(\Lambda + \Lambda^*) \end{aligned} \tag{C.49}$$

the transformation of component fields is

$$\begin{array}{rl} V_\mu \to V_\mu - i\partial_\mu(\Lambda - \Lambda^*) & ; \quad C \to C + \Lambda + \Lambda^+ \\ \lambda \to \lambda & \quad\ \ \xi \to \xi - i\sqrt{2}\psi \\ D \to D & \quad\ \ M \to M - 2iF \end{array} \tag{C.50}$$

So one can use the gauge transformation parameters $\Lambda + \Lambda^*$, ψ, F to gauge away C, ξ and M. The vector supermultiplet then reduces to [2]

$$V(x,\theta,\bar{\theta}) \;=\; -\theta\sigma^\mu\bar{\theta}V_\mu + i\theta\theta\,\bar{\theta}\bar{\lambda} - i\bar{\theta}\bar{\theta}\,\theta\lambda + \frac{1}{2}\theta\theta\,\bar{\theta}\bar{\theta}\,D \qquad \text{(C.51)}$$

This partial gauge fixing, known as Wess-Zumino gauge, still allows for standard gauge transformations $V_\mu \to V_\mu - i\partial_\mu(\Lambda - \Lambda^*)$. Hence the vector supermultiplet provides the supersymmetric generalization of the Yang-Mills gauge potential V_μ. In order to build gauge-invariant kinetic terms, we introduce the field-strength superfields

$$W_\alpha = -\frac{1}{4}\bar{D}\bar{D}\,D_\alpha V \qquad ; \qquad \bar{W}_{\dot{\alpha}} = -\frac{1}{4}DD\,\bar{D}_{\dot{\alpha}}V \qquad \text{(C.52)}$$

They are chiral superfields, which are invariant under the gauge transformations (C.48). In terms of components fields (in coordinates y, θ, $\bar{\theta}$), we have

$$W_\alpha = -i\lambda_\alpha(y) + \theta_\alpha D(y) - \frac{i}{2}(\sigma^\mu\bar{\sigma}\nu)_\alpha{}^\beta\theta_\beta\,F_{\mu\nu}(y) + \theta\theta\,\sigma^\mu{}_{\alpha\dot{\alpha}}\partial_\mu\bar{\lambda}^{\dot{\alpha}}(y) \quad \text{(C.53)}$$

where $F_{\mu\nu} = \partial_{[\mu}V_{\nu]}$. There is a similar expression for $\bar{W}_{\dot{\alpha}}$ in terms of $y^\dagger$. Hence the above superfields provide the supersymmetric completion of the gauge-invariant field strength.

The gauge and Lorentz invariant expression $W^\alpha W_\alpha$ has a highest-dimension component

$$W^\alpha W_\alpha = \ldots + \theta\theta\,(\,-2i\lambda\sigma^\mu\partial_\mu\bar{\lambda} - \frac{1}{2}F^{\mu\nu}F_{\mu\nu} + D^2 + \frac{i}{2}\varepsilon_{\mu\nu\sigma\rho}F^{\mu\nu}F^{\sigma\rho}\,) \quad \text{(C.54)}$$

precisely of the form of the kinetic term (and theta-term) for the $U(1)$ gauge boson, and the gauginos. Hence the action for the gauge boson can be constructed as

$$S = \int d^4x\,d^2\theta\,W^\alpha W_\alpha \;+\; \int d^4x\,d^2\bar{\theta}\,\bar{W}_{\dot{\alpha}}\bar{W}^{\dot{\alpha}} \qquad \text{(C.55)}$$

OBS: For $U(1)$ gauge group, it is also possible to introduce an additional term in the action, known as Fayet-Illiopoulos term, of the form

$$S_{FI} = \chi_{FI}\int d^4x \int d^2\theta\,d^2\bar{\theta}\,V \;=\; \int d^4x\,D \qquad \text{(C.56)}$$

[2] Notice that supersymmetry transformations do not preserve the WZ gauge. Hence any supersymmetry transformation should be followed by a compensating gauge transformation to bring the supermultiplet to the WZ gauge.

where χ_{FI} is a constant.

The discussion of non-abelian gauge bosons is similar, with slightly more general definitions. Vectors superfields have the same structure, but transform in the adjoint representation. The gauge parameters are given by a set of chiral multiplets in the adjoint representation of the gauge group G,

$$\Lambda_{ij} = T^a{}_{ij}\Lambda_a \tag{C.57}$$

The gauge transformation is given by

$$e^V \to e^{-i\Lambda^\dagger} e^V e^{i\Lambda} \tag{C.58}$$

This also allows for a WZ gauge, leaving V_μ^a, λ^a, D^a as degrees of freedom, with the standard gauge transformations for V_μ^a. The non-abelian field-strength superfields are given by

$$W_\alpha = -\frac{1}{4}\bar{D}\bar{D}e^{-V}D_a e^V \tag{C.59}$$

which transforms under (C.58) as

$$W_\alpha \to e^{-i\Lambda^\dagger} W_\alpha e^{i\Lambda} \tag{C.60}$$

The supersymmetric Yang-Mills action is given by (C.55), with an implicit trace over gauge indices.

C.4.4 Coupling of vector and chiral multiplets

We would like to discuss the construction of actions describing the interaction of gauge and chiral supermultiplets. As expected, the coupling of chiral multiplets to gauge vector multiplets is obtained by a suitable modification of the chiral multiplet kinetic term so as to make it gauge invariant.

Let us start with the case of a $U(1)$ vector multiplet, and several chiral multiplets ϕ_i, transforming under $U(1)$ with charges q_i. Namely, under a gauge transformation $V \to V + i(\Lambda - \Lambda^\dagger)$,

$$\Phi_i \to e^{-iq_i\Lambda}\Phi_i \quad ; \quad \Phi_i^\dagger \to e^{iq_i\Lambda^\dagger}\Phi_i^\dagger \tag{C.61}$$

Hence the expression $\Phi_i^\dagger e^{q_i V}\Phi_i$ is gauge invariant, and is the gauge-invariant generalization of $\Phi^\dagger\Phi_i$.

The full lagrangian for the vector and chiral multiplet interactions is

$$S = \int d^4x\, d^2\theta\, W^\alpha W_\alpha + \int d^4x\, d^2\bar\theta\, \bar W_{\dot\alpha} \bar W^{\dot\alpha} + \\ + \int d^4x\, d^2\theta\, d^2\bar\theta\, \Phi_i^\dagger\, e^{q_i V}\, \Phi_i + \int d^4x\, d^2\theta\, W(\Phi_i) + \int d^4x \int d^2\bar\theta\, W(\Phi_i^*) \quad (C.62)$$

In fact, one can generalize the gauge kinetic term to an expression $\int d^4x d^\theta f(\Phi) W^\alpha W_\alpha$, where f is a holomorphic function (known as gauge kinetic function) and Φ are chiral multiplets. Notice that this can be regarded as promoting the gauge coupling to a chiral superfield. In the following we however stick to the simplest situation of constant f.

The term containing the chiral-vector coupling is

$$\int d^4x\, d^2\theta\, d^2\bar\theta\, \Phi^\dagger e^{qV} \Phi = FF^* + \Phi \Box \Phi^* + i\partial_\mu \bar\psi \bar\sigma^\mu \psi + \frac{1}{2}(\, qD\Phi^*\Phi\,) + \quad (C.63)$$

$$+qV_\mu\,(\,\frac{1}{2}\bar\psi\bar\sigma^\mu\psi + \frac{1}{2}\Phi^*\partial_\mu\Phi - \frac{i}{2}\partial_\mu\Phi^*\Phi\,) - \frac{i}{\sqrt{2}}q\,(\,\Phi\bar\lambda\bar\psi - \Phi^*\lambda\psi\,) - \frac{1}{4}q^2\, V_\mu V^\mu \Phi^*\Phi$$

One can integrate out the auxiliary field D, by using its equations of motion. The field D appears in

$$\mathcal{L}_D = \frac{1}{2}D^2 + \frac{1}{2}\sum_i q_i D\Phi_i^*\Phi_i + \chi_{FI} D \quad (C.64)$$

so the equations of motion give $D = -1/2\, \sum_i q_i \Phi_i^*\Phi_i + \chi_{FI}$. The D-term lagrangian becomes a potential term

$$V_D = \frac{1}{2}(\frac{1}{2}\sum_i q_i \Phi_i^*\Phi_i - \chi_{FI})^2 \quad (C.65)$$

The condition $D = 0$ that it vanishes is a necessary condition for a supersymmetric vacuum, known as D-term condition.

For non-abelian gauge symmetries, chiral multiplets transform in a representation R of the gauge group,

$$\Phi \to e^{-i\Lambda}\Phi \quad (C.66)$$

where Φ is regarded as a column vector and $\Lambda_{ij} = (t_a^R)_{ij}\Lambda^a$ is a matrix acting on it. The action for the complete system is given by

$$S = \frac{1}{4g^2}\int d^4x\, d^2\theta\, W^\alpha W_\alpha + \frac{1}{4g^2}\int d^4x\, d^2\bar\theta\, \bar W_{\dot\alpha}\bar W^{\dot\alpha} + \\ + \int d^4x \int d^2\theta\, d^2\bar\theta\, \Phi_i^\dagger e^{t_a^R V_a}\Phi_i + \int d^4x\, d^2\theta\, W(\Phi_i) + \int d^4x\, d^2\bar\theta\, W(\Phi_i^*) \quad (C.67)$$

After integrating out the D field, the D-term potential has the explicit expression

$$V_D = \frac{1}{2}\sum_a \left(\frac{1}{2}\sum_k \Phi_k^\dagger (t_a^{R_k}) \Phi_k\right)^2 \qquad \text{(C.68)}$$

where the sum in k runs over all chiral multiplets in non-trivial representations (denoted R_k) of the gauge group G.

In conclusion, the most general $N = 1$ supersymmetric action (up to two derivatives) for a system of chiral and vector multiplets is specified by three functions: the Kahler potential $K(\Phi, \Phi^\dagger)$, which is a real function and defines the chiral multiplet kinetic term, the superpotential $W(\Phi)$, which is holomorphic and defines chiral multiplet interactions, and the gauge kinetic functions $f(\Phi)$, which are holomorphic and define the gauge boson kinetic term.

C.4.5 Moduli space

Supersymmetric gauge field theories often contain flat directions in the scalar potential, namely there is a continuous set of (inequivalent) supersymmetric vacuum states of the theory, parametrized by the vacuum expectation values (vevs) for scalar fields. The scalars parametrizing flat directions in the scalar potential are known as moduli (moduli fields in string theory, like the dilaton etc, are indeed examples of such fields), and are massless. The set of vevs corresponding to supersymmetric minima of the theory is known as moduli space.

The conditions that scalar vevs should satisfy to belong to the moduli space are that the F-terms and D-terms vanish, namely

$$\begin{aligned} \frac{\partial W}{\partial \Phi_i} &= 0 \\ \sum_i \Phi_i^\dagger (t_a^{R_i}) \Phi_i &= 0 \end{aligned} \qquad \text{(C.69)}$$

where i runs through the chiral multiplets in the theory (in a representation R_i of the gauge group) and a runs through the generators of the gauge group.

Notice that supersymmetry is essential in maintaining the direction flat after quantum corrections. Indeed the F-term conditions are obtained from the superpotential, which is protected against quantum corrections by supersymmetry. On the other hand, the D-term conditions follow from gauge invariance, and are uncorrected as well. In non-supersymmetric theories, fields which look like moduli at tree level typically acquire mass terms from radiative corrections, and moduli space is lifted (a non-trivial scalar potential develops).

Let us provide some typical examples of theories with flat directions.

Consider a $U(1)$ gauge theory with one neutral chiral multiplet Φ, and two chiral multiplets Φ_1, Φ_2 with charge $+1$, and two Φ_1, Φ_2 with charge -1. We introduce a superpotential

$$W = \Phi\Phi_1\Phi_1' - \Phi\Phi_2\Phi_2' \tag{C.70}$$

The F-term conditions on scalars give

$$\Phi_1\Phi_1' = \Phi_2\Phi_2' \quad ; \quad \Phi\Phi_i = 0 \quad ; \quad \Phi\Phi_i' = 0 \tag{C.71}$$

while the D-term conditions read

$$|\Phi_1|^2 + |\Phi_2|^2 - |\Phi_1'|^2 - |\Phi_2'|^2 = 0 \tag{C.72}$$

These equations are satisfied for the choice of vevs

$$\langle\Phi\rangle = 0 \quad ; \quad \langle\Phi_1\rangle = v \quad ; \quad \langle\Phi_1'\rangle = w \quad ; \quad \langle\Phi_2\rangle = w \quad ; \quad \langle\Phi_2'\rangle = v \quad ; \tag{C.73}$$

So the moduli space is parametrized by two complex parameters. There is a complex two-dimensions manifold of vacuum configurations for this theory [3].

Let us provide a second example, with non-abelian gauge symmetry. Consider a $U(N)$ supersymmetric gauge theory with three chiral multiplets Φ_i in the adjoint representation (thus regarded as $N \times N$ matrices, and superpotential

$$W = \mathrm{tr}\,(\,\Phi_1\Phi_2\Phi_3\,-\,\Phi_1\Phi_3\Phi_2\,) \tag{C.74}$$

[3] As we will see later, this theory is in fact $N = 2$ supersymmetric, with V and Φ forming an $N = 2$ vector multiplet, and Φ_i, Φ_i' forming two hypermultiplets. The moduli space is parametrized by the vevs of a hypermultimplet, given by a combination of the latter.

This theory has a very non-trivial moduli space [4]. The F-term conditions read

$$[\Phi_i, \Phi_j] = 0 \qquad \text{(C.75)}$$

This implies that the matrices of vevs for these fields should be commuting. Then one can use gauge transformations to simultaneously diagonalize them, so that the vevs are

$$(\Phi_i)_{mn} = (v_i)_n \delta_{mn} \qquad \text{(no sum)} \qquad \text{(C.76)}$$

For adjoint multiplets expressed as $n \times n$ matrices, the D-term condition is

$$\sum_i \left[\, (\Phi_i^\dagger)_{mn} (t_a^{\rm fund})_{np} (\Phi_i)_{pm} - (\Phi_i^\dagger)_{mn} (t_a^{\rm fund})_{mq} (\Phi_i)_{nq} \,\right] \,=\, 0 \qquad \text{(C.77)}$$

These are automatically satisfied, upont substitution of the above vevs.

Hence the moduli space is parametrized by the n triples of complex eigenvalues $(v_i)_n$. Some realizations of this gauge theory in string theory (in terms of configurations of D-branes) allow for a simple geometric interpretation of this moduli space.

C.5 Extended 4d supersymmetry

C.5.1 Extended superalgebras

N-extended supersymmetry is generated by N Weyl spinor supercharges Q_α^I, $\bar{Q}_{\dot{\alpha} I}$, with $I = 1, \ldots, N$. Since each supercharge contains two-components, the number of supercharges is $4N$. The algebra that they satisfy is

$$\begin{aligned} \{Q_\alpha^I, \bar{Q}_{\dot{\alpha} J}\} &= 2\sigma^\mu_{\alpha\dot{\alpha}} P_\mu \delta^I{}_J \\ \{Q_\alpha^I, Q_\beta^J\} &= \varepsilon_{\alpha\beta} Z^{IJ} \\ \{\bar{Q}_{\dot{\alpha} I}, \bar{Q}_{\dot{\beta} J}\} &= \varepsilon_{\dot{\alpha}\dot{\beta}} (Z^*)_{IJ} \end{aligned} \qquad \text{(C.78)}$$

with Z^{IJ} antisymmetric in its indices.

This is the most general superalgebra consistent with 4d Lorentz invariance. The Z^{IJ} (and their conjugates Z^*) commute with all supercharges Q, $\bar{Q}$, and are known as central charges. Each state (each supermultiplet) has

[4] As we will see later, this theory is in fact $\mathcal{N} = 4$ supersymmetric.

a particular value for the corresponding operators. For the most familiar supermultiplets, the value of the central charges is zero, so we ignore them for most of our discussion (however, the supermultiplets describing soliton states of certain supersymmetric theories have non-trivial central charges. Thus, we will make some useful comments on this case, towards the end).

Some remarks are in order: Notice that the R-symmetry of the superalgebra is (for zero central charges) $U(N)$, where the $SU(N)$ acts on the indices I (in the fundamental or antifundamental representation), while the $U(1)$ acts on supercharges as an overall phase rotation (just like in $N=1$ supersymmetry). Notice also the fact that the N-extended supersymmery algebra contains the supersymmetry algebras of M-extended supersymmetry, for $M<N$. This implies that the supermultiplets of extended supersymmetries naturally decompose as sums of supermultiplets of their subalgebras.

C.5.2 Supermultiplet structure

Let us start by considering the construction of supermultiplets, in a sector of zero central charges, so that the superalgebra reads

$$\begin{aligned} \{Q^I_\alpha,\bar{Q}_{\dot{\alpha}J}\} &= 2\sigma^\mu_{\alpha\dot{\alpha}}P_\mu\delta^I{}_J \\ \{Q^I_\alpha,Q^J_\beta\} &= 0 \qquad ; \qquad \{\bar{Q}_{\dot{\alpha}I},\bar{Q}_{\dot{\beta}J}\}=0 \end{aligned} \tag{C.79}$$

Let us start discussing massless supermultiplets. In the reference frame where the momentum is $(P_\mu)=(-E,0,0,E)$, the non-trivial piece of the superalgebra reads

$$\{Q^I_\alpha,\bar{Q}_{\dot{\beta}J}\}=2\begin{pmatrix}2E & 0\\ 0 & 0\end{pmatrix}\delta^I{}_J \tag{C.80}$$

As in the $N=1$ case, the supercharges Q^I_2, $\bar{Q}_{\dot{2}J}$ are realized as zero, and we introduce

$$a^I=\frac{1}{2\sqrt{2}}Q^I_1 \qquad ; \qquad a^\dagger_I=\frac{1}{2\sqrt{2}}\bar{Q}_{\dot{1}I} \tag{C.81}$$

which satisfy

$$\{a^I,a^\dagger_J\}=\delta^I{}_J \quad ; \quad \{a^I,a^J\}=\{a^\dagger_I,a^\dagger_J\}=0 \tag{C.82}$$

We construct the supermultiplet by starting with a state $\Omega\rangle$ of lowest helicity j, annihilated by the a^I (and the Q_2, $\bar{Q}_{\dot{2}}$), and applying the operators $a^\dagger_I$ to it. The number of states in such multiplet is 2^N. As in the $N = 1$ case, CPT invariance may require to combine these basic multiplets with their conjugates to be realized in a local field theory.

We will discuss some explicit examples of massless supermultiplets below.

The construction of massive supermultiplets is also a simple generalization of the $N = 1$ case. In the rest frame, we have

$$\begin{aligned}\{Q^I_\alpha, \bar{Q}_{\dot{\alpha}J}\} &= 2M\delta_{\alpha\dot{\alpha}}\delta^I{}_J \\ \{Q^I_\alpha, Q^J_\beta\} &= 0 \quad ; \quad \{\bar{Q}_{\dot{\alpha}I}, \bar{Q}_{\dot{\beta}J}\} = 0\end{aligned} \tag{C.83}$$

Rescaling the operators as

$$a^I_\alpha = \frac{1}{\sqrt{2M}} Q^I_\alpha \quad ; \quad a^{I\,\dagger}_\alpha = \frac{1}{\sqrt{2M}} \bar{Q}_{\dot{\alpha}I} \tag{C.84}$$

we have a set of $2N$ decoupled fermionic harmonic oscillators, which lead to a supermultiplet of 2^{2N} degrees of freedom.

Finally, let us briefly sketch the construction of massive multiplets in a sector of non-zero central charges. Using the R-symmetry of the theory, we may bring the antisymmetric matrix Z^{IJ} to a block form e.g. for N even (on which we center in what follows)

$$Z = \varepsilon \otimes D = \begin{pmatrix} 0 & D \\ D & 0 \end{pmatrix} \tag{C.85}$$

with $D = \mathrm{diag}\,(Z_1, \ldots, Z_{N/2})$. Splitting the indices I as (a, m), with $a = 1, 2$ and $m = 1, \ldots, N/2$, the central charges read $Z^{am,bn} = \epsilon^{ab}\delta^{mn}\, Z_n$ (no sum). The superalgebra reads

$$\begin{aligned}\{Q^{am}_\alpha, \bar{Q}_{\dot{\alpha}bn}\} &= 2M\delta_{\alpha\dot{\alpha}}\delta^a{}_b\delta^m{}_n \\ \{Q^{am}_\alpha, Q^{bn}_\beta\} &= \varepsilon_{\alpha\beta}\varepsilon^{ab}\delta^{mn} Z_n \\ \{\bar{Q}_{\dot{\alpha}am}, \bar{Q}_{\dot{\beta}bn}\} &= \varepsilon_{\dot{\alpha}\dot{\beta}}\varepsilon_{ab}\delta_{mn} Z_n\end{aligned} \tag{C.86}$$

We can define the linear combinations

$$\begin{aligned}a^m_\alpha &= \frac{1}{\sqrt{2}}[Q^{1m}_\alpha + \varepsilon_{\alpha\beta}\bar{Q}_{\dot{\beta}2m}] \\ b^m_\alpha &= \frac{1}{\sqrt{2}}[Q^{1m}_\alpha - \varepsilon_{\alpha\beta}\bar{Q}_{\dot{\beta}2m}]\end{aligned} \tag{C.87}$$

and their adjoints. They satisfy

$$\begin{aligned}
\{a^m_\alpha, a^n_\beta\} &= \{b^m_\alpha, b^n_\beta\} = \{a^m_\alpha, b^n_\beta\} = 0 \\
\{a^m_\alpha, (a^n_\beta)^\dagger\} &= \delta_{\alpha\beta}\delta^{mn}(2M + Z_n) \\
\{b^m_\alpha, (b^n_\beta)^\dagger\} &= \delta_{\alpha\beta}\delta^{mn}(2M - Z_n)
\end{aligned} \tag{C.88}$$

From this it follows that in a sector of given charges Z_n, the masses of the states satisfy $2M \geq |Z_n|$, for all n. This conditions is known as the BPS bound.

For generic mass M, we have $2\times 2\times N/2$ fermionic harmonic oscillators, so that supermultiplets contain 2^{2N} states. On the other hand if $2M = \pm Z_n$ for some n, then some of the operators anticommute, and are realized as zero, so there are $2N-1$ harmonic oscillators, and the representation contains 2^{2N-1} states, less than the generic supermultiplet. Supermultiplets saturating the BPS bound are known as BPS states, and contain less states than generic supermultiplets. This guarantees that BPS states cannot cease to be BPS, and their mass is given by the central charge, which is part of the algebra. Hence, for BPS states the mass is controlled by the symmetry of the theory and is protected agains quantum corrections by supersymmetry.

C.5.3 Some useful information on extended supersymmetric field theories

There is no simple superfield formalism for theories with extended supersymmetry, hence supersymmetry transformations must be checked on-shell. The simplest way to describe the supermultiplets and the supersymmetric actions is thus to phrase them in terms of the supermultiplets and superfield formalism of an $N = 1$ subalgebra.

In the following we discuss some basic features of $N = 2, 4$ supersymmetric theories. $N = 8$ supersymmetry also appears in some applications, but the smallest supermultiplet already contains spin-2 particles, namely gravitons. They can be realized in theories describing gravitational interactions, namely supergravity theories. Their discussion is beyond the scope of this lecture. Finally, for even higher degree of supersymmetry, even the smallest massless supermultiplet already contains fields with spin higher than 2. It is not known how to write interacting theories for such fields, hence they are not usually considered.

$N = 2$ supersymmetric theories

The basic supermultiplets of $N = 2$ supersymmetric field theories are most simply described by specifying their decomposition under a $N = 1$ subalgebra of the theory. We describe some useful massless supermultiplets.

- The hypermultiplet: It decomposes as two chiral multiplets (in conjugate representations of the gauge and global symmetries) of $N = 1$ supersymmetry. Hence, one hypermultiplet contains two complex scalars and two Weyl fermions. Notice that the latter have same chirality and conjugate quantum numbers, hence the supermultiplet is non-chiral. It is possible to write supersymmetric mass terms for hypermultiplets, hence the massive hypermultiplet has the same supermultiplet structure.

- The $N = 2$ vector multiplet: It decomposes as one $N = 1$ vector multiplet, and a chiral multiplet (in the adjoint representation). Hence, it contains a gauge boson, two Majorana fermions, and one complex scalar.

Let us describe the general action (up to two derivatives) for an $N = 2$ supersymmetric theory with hyper- and vector multiplets. For $N = 2$, the action is fully determined by the gauge quantum numbers of the hypermultiplets. Let us denote V, Σ the $N = 1$ vector and chiral multiplets in the $N = 2$ vector multiplets of the gauge group G, and Φ_i, Φ'_i the two chiral multiplets in the i^{th} hypermultiplet, in the representation R_i. The $N = 2$ action has the standard $N = 1$ form, with a superpotential fully determined by gauge symmetry and supersymmetry

$$W(\Phi_i, \Phi'_i, \Sigma) = \sum_{i,a} \Phi_i \Sigma_a (t_a^{R_i}) \Phi'_i \qquad \text{(C.89)}$$

The $N = 2$ supersymmetry implies additional non-renormalization theorems beyond those in the $N = 1$ theory. For instance, in $N = 1$ language the Kahler potential for the chiral multiplets splits in two pieces, $K(\Sigma, \Sigma^\dagger)$ and $K(\Phi, \Phi', \Phi^\dagger, \Phi'^\dagger)$. This implies that the kinetic terms for scalars in vector multiplets do not depend on scalars in hypermultiplets, and viceversa. This implies that the scalar field space (and hence the moduli space) factorizes as the vector multiplet scalar field space times the hypermultiplet scalar field space. Moreover, the former is a Kahler space, while the latter is even more constrained, and is hyperKahler [5].

[5]Namely, admits three Kahler forms, with their product obeying the rules of quaternionic product.

$N = 4$ **supersymmetric theories**

Let us now describe some facts on $N = 4$ supersymmetric theories [6].

The smallest supermultiplet is the $N = 4$ vector multiplet. Under an $N = 2$ subalgebra, it contains one $N = 2$ vector multiplet and one hypermultiplet in the adjoint representation. In terms of $N = 1$, it contains a vector multiplet and three chiral multiplets in the adjoint representation. Finally, in component fields, it contains one gauge boson, four Majorana fermions, and six real scalars.

Other supermultiplets contain spin-2 particles, namely gravitons, and so appear only in supergravity theories. Their discussion is beyond the scope of this lecture.

The general action for an $\mathcal{N} = 4$ theory is extremely constrained. It has the structure of an $N = 2$ theory, but with the gauge representation of hypermultiplets fixed by the $N = 4$ supermultiplet structure. Using $N = 1$ language, we denote V, Φ_1, Φ_2, Φ_3 the vector and chiral multiplets of the $N = 4$ vector multiplet. The superpotential is given by

$$W(\Phi_i) = \mathrm{Tr}\,\Phi_1\Phi_2\Phi_3 - \mathrm{Tr}\,\Phi_1\Phi_3\Phi_2 \tag{C.90}$$

Again, the action is protected by even more powerful non-renomalization theorem. In particular, the Kahler potential for scalar fields are forced to be canonical, and the gauge kinetic functions are non-renormalized. This implies that $N = 4$ supersymmetric theories are finite (this in fact holds even non-perturbatively).

C.6 Supersymmetry in several dimensions

C.6.1 Some generalities

In this section we sketch the basic structure of supermultiplets in theories in more than four dimensions. The basic ideas are completely analogous to those discussed for four-dimensional supersymmetry. The main difference arises because of the larger number of components of higher-dimensional spinors, as compared with four-dimensional ones.

[6] The supermultiplet structure and low-energy effective action of $N = 3$ is exactly as in $N = 4$, so $N = 3$ supersymmetry is not so interesting.

A detailed discussion of the construction of irreducible spinor representation of the Lorentz group in an arbitrary number of dimensions can be found in appendix B of [71]. For our present purposes, it will be enough to just mention that in an even number of dimensions, $D = 2n$, the representation of the Clifford algebra has dimension 2^n.

This spinor representation of $SO(2n-1,1)$ is reducible into two Weyl spinor representations, of opposite chiralities, and with 2^{n-1} components each. Also, for odd n, namely $D = 2k+4$ it is possible to define Majorana spinors, which satisfy a reality condition, and thus have 2^{n-1} components. In general, Majorana and Weyl conditions are incompatible (namely, the conjugation operation flips the chirality, so Majorana spinors contain components with opposite chiralities). However, for $D = 2k+8$, the conjugation operation does not flip the chirality, and one can define spinors satisfying both the Majorana and Weyl conditions, and thus have 2^{n-2} components.

The basic features of supersymmetric theories in different dimensions mainly depend only on the total number of supercharges. Indeed, any superalgebra in a given dimension can be regarded as a superalgebra of lower dimensional supersymmetry, simply obtained by decomposing the Lorentz representations of supergenerators with respect to the lower-dimensional Lorentz group. This is usually knows as dimensional reduction. Notice that since spinor representations in higher dimensions have larger number of components than in lower dimensions, the original superalgebra in general descends to an extended superalgebra in the lower dimension. Clearly, the same kind of relation follows for representations of the superalgebras. Namely, supermultiplets of the higher-dimensional supersymmetry can be recast as supermultiplets of the lower-dimensional one. An important point is that, since higher-dimensional superalgebras are related to extended superalgebras in 4d, there is no superfield formalism to describe the structure of higher-dimensional supermultiplets.

In each dimension, it is conventional to define $N = 1$ supersymmetry as that generated by supercharges in the smallest spinor representation. Hence, N-extended supersymmetry corresponds to that generated by N supercharges in the smallest spinor representation. Since the number of components of spinors jumps with dimension in a non-trivial way, it is sometimes more useful to refer to the theories by its total number of supercharges, although we will use the conventional N-extended susy notation as well.

C.6.2 Some useful superalgebras and supermultiplets in higher dimensions

In this section we provide some useful supermultiplets of certain superalgebras in six and ten dimensions. It is by no means a complete classification, but rather a list of some structures which will appear in the main text. A detailed classification of superalgebras and supermultiplets may be found in [126] and [127].

Minimal Supersymmetry in six dimensions

In six dimensions $D = 6$, the Weyl spinor has $2^3/2 = 4$ complex components, hence the minimal supersymmetry, denoted $N = 1$, is generated by 8 supercharges. Thus $D = 6$ N-extended supersymmetry is generated by $8N$ supercharges.

Let us center on the minimal supersymmetry, with 8 supercharges, denoted $N = 1$ (sometimes also $N = (1,0)$ or $(0,1)$ to indicate the left or right chirality of the chosen supergenerators; clearly, both such superalgebras are isomorphic). The R-symmetry of the theory is $SU(2)_R$. Let us describe some useful massless supermultiplets of this theory, providing their quantum numbers under the Lorentz (massless) little group $SO(4)_L = SU(2) \times SU(2)$ and the R-symmetry $SU(2)_L$.

Vector multiplet: It contains fields transforming under $SU(2) \times SU(2) \times SU(2)_R$ as

$$(2,2;1) \,+\, (1,2;2) \tag{C.91}$$

namely a massless vector boson and a chiral right-handed Weyl spinor.

Hypermultiplet: It contains fields transforming as

$$(2,1;1) \,+\, (1,1;2) \tag{C.92}$$

Unless it transforms in a pseudoreal representation of the gauge and global symmetries, it must be combined with its CPT conjugate to form a physical field. Then it contains two complex scalar fields, and a chiral left-handed Weyl spinor.

Tensor multiplet: It contains fields transforming as

$$(3,1;1) \,+\, (1,1;1) \,+\, (2,1;2) \tag{C.93}$$

namely a self-dual two-form, a real scalar fields and a chiral left-handed Weyl spinor.

Graviton multiplet: It contains fields transforming as

$$(3,3;1) + (1,3;1) + (2,3;2) \tag{C.94}$$

namely a massless graviton, an anti-selfdual 2-form, and two left-handed gravitinos.

This superalgebra can be dimensionally reduced to 4d $N=2$ supersymmetry. It is a simple exercise to match the above 6d supermultiplets with supermultiplets of 4d $N=2$ supersymmetry.

Extended supersymmetry in six dimensions

Let us discuss some features of $N=2$ supersymmetry in six dimensions. The superalgebra is generated by 16 supercharges, organized in two Weyl spinors. There are two possible inequivalent superalgebras, depending on the relative chirality of these two spinors. Namely, there is a 6d $N=(2,0)$ superalgebra, where both supergenerators have the same chirality, and a 6d $N=(1,1)$ superalgebra, where they have opposite chiralities. Let us describe some of their massless multiplets in turn.

The $N=(2,0)$ supersymmetry has a $USp(4)=SO(5)$ R-symmetry. Some interesting massless supermultiplets are

Tensor multiplet: It contains fields transforming under $SU(2)\times SU(2)\times SO(5)_R$ as

$$(3,1;1) + (1,1;5) + (2,1;4) \tag{C.95}$$

namely a self-dual two-form, five real scalar fields and two chiral left-handed Weyl spinors. Notice that it decomposes as a hyper- and a tensor multiplet with respect to the 6d $N=1$ subalgebra.

Graviton multiplet: It contains fields transforming as

$$(3,3;1) + (1,3;5) + (2,3;4) \tag{C.96}$$

namely, a graviton, five anti-selfdual 2-forms and four left-handed gravitinos.

The $N=(1,1)$ supersymmetry has a $SO(4)=SU(2)\times SU(2)$ R-symmetry. Some interesting massless supermultiplets are

Vector multiplet: It contains fields transforming under $SU(2) \times SU(2) \times [SU(2) \times SU(2)]_R$ as

$$(2,2;1,1) + (1,1;2,2) + (2,1;1,2) + (1,2;2,1) \qquad \text{(C.97)}$$

namely a massless vector boson, two complex scalars, one chiral left- and one chiral right-handed Weyl spinors. Notice that it decomposes as a hyper- and a vector multiplet with respect to the 6d $N = 1$ subalgebra.

Graviton multiplet: It contains fields transforming as

$$(3,3;1,1) + (3,1;1,1) + (1,3;1,1) + (1,1;1,1) + (2,2;2,2) + \\ + (3,2;1,2) + (2,3;2,1) + (1,2;1,2) + (2,1;2,1) \qquad \text{(C.98)}$$

namely, a graviton, a two-form, a real scalar, four vector bosons, two left- and two right-handed gravitinos, and one left- and one right-handed spinor.

Supersymmetry in ten dimensions

In ten dimensions $D = 10$, the minimal spinor satisfies the Majorana and Weyl constraints and has $2^5/4 = 8$ complex components, hence the minimal supersymmetry, denoted $N = 1$, is generated by 16 supercharges. Thus $D = 6$ N-extended supersymmetry is generated by $8N$ supercharges. Indeed, for $N > 2$ the smallest massless supermultiplet contains fields with spin higher than two; it is not known how to write interacting theories for such fields, hence they are not usually considered.

Let us center on the minimal $N = 1$ supersymmetry, with 16 supercharges. The R-symmetry of the theory is trivial. Some useful massless supermultiplets of this theory are

Vector multiplet, containing fields in the $8_V + 8_C$ of the $SO(8)$ Lorentz little group. Namely, a massless vector boson and a chiral 10d spinor.

Graviton multiplet, containing fields transforming under $SO(8)$ as

$$35_V + 28_V + 1 + 8_S + 56_S \qquad \text{(C.99)}$$

namely, a graviton, a 2-form, a real scalar, a right-handed gravitino and a right-handed spinor.

Concerning extended supersymmetry, with 32 supercharges organized in two Majorana-Weyl spinors, there are two possibilities, according to their relative chirality. The 10d $N = (2,0)$ supersymmetry is generated by spinors of same chirality. The R-symmetry is $SO(2)_R$. The only relevant massless supermultiplet is the graviton multiplet, with fields transforming as

$$\begin{aligned} & \mathbf{35}_V + \mathbf{28}_V + \mathbf{1} + \mathbf{35}_C + \mathbf{28}_C + \mathbf{1} + \\ & \quad + 2 \times (\mathbf{8}_C + \mathbf{56}_C) \end{aligned} \tag{C.100}$$

Namely, one graviton, two 2-forms, two real scalars, one self-dual 4-form and two right-handed gravitinos and two right-handed spinors.

The 10d $N = (1, 1)$ supersymmetry is generated by spinors of opposite chirality. The R-symmetry is trivial. The only relevant massless supermultiplet is the graviton multiplet, with fields transforming as

$$\begin{aligned} & \mathbf{35}_V + \mathbf{28}_V + \mathbf{1} + \mathbf{8}_V + \mathbf{56}_V + \\ & + + \mathbf{8}_C + \mathbf{56}_C + \mathbf{8}_S + \mathbf{56}_S \end{aligned} \tag{C.101}$$

Namely, one graviton, one 2-form, one real scalar, one 1-form, one 3-form, one left- and one right-handed gravitino and one left- and one right-handed spinor.

Finally, for completeness we provide the basic massless supermultiplet of 11d $N = 1$ supersymmetry, the gravity multiplet. It contains states transforming as $\mathbf{44} + \mathbf{84} + \mathbf{128}$ under the $SO(9)$ Lorentz little group. Notice that it maps to the gravity multiplet of 10d $N = (1, 1)$ supersymmetry upon dimensional reduction.

Notice that going to higher dimensions requires introducing more supercharges, which implies that even the smallest massless supermultiplet already contains fields with spin higher than 2, so these theories are usually not considered. This underlies the statement that eleven is the maximal number of dimensions allowed by supersymmetry (with the extra assumption of not having massless fields with spins higher than 2). The maximal amount of supersymmetry is thus 32 supercharges.

Appendix D

Rudiments of differential geometry/topology

Useful references for this lecture are [117] and sections 12, 14 and 15 of [118].

D.1 Differential manifolds; Homology and cohomology

D.1.1 Differential manifolds

An n-dimensional differential manifold M is a topological space, together with an atlas, that is a collection of charts $(U_\alpha, x_{(\alpha)})$ where U_α are open sets of M and $x_{(\alpha)}$ is a one to one map between U_α and an open set in $\mathbf{R^n}$, such that

i) M is covered by the U_α, that is $\bigcup_\alpha U_\alpha = M$.

ii) If $U_\alpha \bigcap U_\beta$ is non-empty, the map

$$x_{(\beta)} \circ x_{(\alpha)}^{-1} \,:\, x_{(\alpha)}(U_\alpha \bigcap U_\beta) \in \mathbf{R^n} \;\rightarrow\; \mathbf{x}_{(\beta)}(\mathbf{U}_\alpha \bigcap \mathbf{U}_\beta) \in \mathbf{R^n} \qquad \text{(D.1)}$$

is differentiable.

Namely, the charts attach coordinates to the points in the U_α, such that on intersections $U\alpha \bigcap U_\beta$ the $x_{(\beta)}$ are smooth functions of the $x_{(\alpha)}$. This is illustrated in figure D.1. Namely, a differential manifold is a space that at each point looks locally like $\mathbf{R^n}$ (with respect to differential structures).

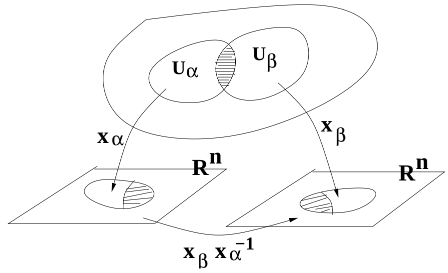


Figure D.1: Charts in a differential manifold.

By abuse of notation, we will often refer to a point $P \in M$ by its coordinates x (in some chart). Also, we will denote the map $x_{(\beta)} \circ x_{(\alpha)}^{-1}$ as $x_{(\beta)}(x_{(\alpha)})$.

We refer to any introductory book on differential geometry for examples of the description of familiar manifolds (like the n-sphere $\mathbf{S^n}$ or the n-torus $\mathbf{T^n}$ in the above language).

In this lecture we will center on orientable manifolds. An orientable manifold is such that the sign of the determinant of the jacobian matrix $J_i^j = \partial x^j_{(\beta)}/\partial x^i_{(\alpha)}$ is the same in all intersections $U_\alpha \bigcap U_\beta$.

In a differential manifold we can introduce the concept of a differentiable (or smooth) function. A function $f : M \to \mathbf{R}$ is differentiable if the functions

$$f \circ x_{(\alpha)}^{-1} : x_{(\alpha)}(U_\alpha) \in \mathbf{R^n} \to \mathbf{R} \tag{D.2}$$

are differentiable. And similarly for functions taking values in $\mathbf{R^n}$, $\mathbf{C}$, $\mathbf{C^n}$, etc.

We denote by $\mathcal{F}$ the set of smooth (real) functions over M. By abuse of language we often write $f(x)$ to denote $f \circ x_{(\alpha)}^{-1}$.

D.1.2 Tangent and cotangent space

A tangent vector to M at a point $P \in U_\alpha$ is a linear mapping from the set of smooth functions $\mathcal{F}$ to $\mathbf{R}$. A basis of tangent vectors is the set $\{\partial_i\}$, $i = 1, \dots, n$, which act as

$$\begin{aligned} \partial_i \ : \ \mathcal{F} &\to \mathbf{R} \\ f &\longmapsto \left.\frac{\partial f}{\partial_{x^i_{(\alpha)}}}\right|_P \end{aligned} \tag{D.3}$$

The tangent space to M at P, denoted $T_P(M)$, is the vector space generated by linear combinations of the ∂_i, acting as

$$\begin{aligned} V = V^i \partial_i \quad : \quad \mathcal{F} &\rightarrow \mathbf{R} \\ f &\longmapsto V^i \left.\frac{\partial f}{\partial_{x^i_{(\alpha)}}}\right|_P \end{aligned} \tag{D.4}$$

A vector field is a set of tangent vectors, one per point of M, smoothly varying with P. Namely, a set of linear combinations with coefficients given by functions, defined on the U_α

$$V_{(\alpha)} = V^i_{(\alpha)}(x_{(\alpha)})\partial_i \tag{D.5}$$

with the conditions that they agree on intersections $U_\alpha \bigcap U_\beta$, namely

$$V^i_{(\alpha)}(x_{(\alpha)}) = \frac{\partial x_{(\alpha)^i}}{\partial x^j_{(\beta)}} V^j_{(\beta)}(x_{(\beta)}) \tag{D.6}$$

We will define analogously the concept of field for other vector spaces below. In section 2.1 we will see that they are simply sections of the corresponding fiber bundle.

The cotangent space $T_P(M)^*$ of M at P is the vector space dual to $T_P(M)$. Namely it is the vector space of linear mappings from $T_P(M)$ to $\mathbf{R}$. We can understand this better by introducing a basis for $T_P(M)^*$, which is given by the set $\{dx^i\}$, which act as

$$\begin{aligned} dx^i \quad : \quad T_P(M) &\rightarrow \mathbf{R} \\ \partial_j &\longmapsto \delta^i_j \end{aligned} \tag{D.7}$$

A general linear combination $u = u_i dx^i$ is hence defined by

$$\begin{aligned} u \quad : \quad T_P(M) &\rightarrow \mathbf{R} \\ \partial_j &\longmapsto u_j \end{aligned} \tag{D.8}$$

The element of $T_P(M)^*$ are also called 1-forms, see below.

A tensor of type (k, l) is a linear mapping from $(T_P(M)^*)^k \times T_P(M)^l$ to $\mathbf{R}$. It is the vector space of linear combinations

$$T \;=\; T^{i_1 \ldots i_k}_{j_1 \ldots j_l}\; dx^{j_1} \otimes \ldots \otimes dx^{j_l} \otimes \partial_{i_1} \otimes \ldots \otimes \partial_{i_k} \tag{D.9}$$

with the obvious definition of the elements of the basis.

A simple examples is given by the metric, which is a tensor field of type $(0, 2)$, $g = g_{ij} dx^i \otimes dx^j$, or $g_{ij} = g(\partial_i, \partial_j)$.

D.1.3 Differential forms

A differential p-form is a tensor of type $(0,p)$, which has completely antisymmetric component (this statement is true in any coordinates). So they are of the form

$$A_{(p)} = A_{i_1\ldots i_p}\, dx^{i_1} \otimes \ldots \otimes dx^{i_p} \tag{D.10}$$

with completely antisymmetric $A_{i_1\ldots i_p}$.

Equivalently, it is the vector space of linear combinations of the basis elements

$$dx^{i_1} \wedge \ldots \wedge dx^{i_p} = \frac{1}{p!}\epsilon_{i_1\ldots i_p} dx^{i_1} \otimes \ldots dx^{i_p} \tag{D.11}$$

(with $i_1 < \ldots < i_p$), namely

$$A_{(p)} = A_{i_1\ldots i_p}\, dx^{i_1} \wedge \ldots \wedge dx^{i_p} \tag{D.12}$$

The vector space of p-forms is denoted $\Lambda^p(M)$. We define p-form fields as usual, which will be denoted p-forms by abuse of language.

We define the wedge product of a p-form $A_{(p)}$ and a q-form $B_{(q)}$ to be the $(p+q)$-form

$$A_{(p)} \wedge B_{(q)} \;=\; \frac{1}{p!q!}\, A_{i_1\ldots i_p}\, B_{j_1\ldots j_q}\, dx^{i_1} \wedge \ldots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_q} \tag{D.13}$$

Notice the property $A_{(p)} \wedge B_q = (-1)^{pq} B_{(q)} \wedge A_{(p)}$. Often, wedge products are assumed and not explicitly displayed.

We define the exterior derivative d as a mapping from p-form fields to $(p+1)$-form fields. For a p-form (field) $A_{(p)}$ its exterior derivative $(dA)_{(p+1)}$ is defined by

$$dA \;=\; \partial_{i_0} A_{i_1\ldots i_p}\, dx^{i_0} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_p} \tag{D.14}$$

Notice the property

$$d(A_p \wedge B_{(q)}) = dA_{(p)} \wedge B_{(q)} + (-1)^p A_{(p)} \wedge dB_{(q)} \tag{D.15}$$

However, the main property of exterior differentiation for this lecture is

$$d^2 = 0 \tag{D.16}$$

in the sense that for any p-form $A_{(p)}$, $d(dA) = 0$. This follows easily from the symmetry of double partial derivation $\partial_i\partial_j = \partial_j\partial_i$.

We refer to introductory books on differential forms to check that d reproduces the familiar formulae for the gradient, divergence and curl of 3d vector calculus.

D.1.4 Cohomology

A p-form field $A_{(p)}$ is said to be closed if $dA = 0$. A p-form $A_{(p)}$ is said to be exact if there exists a $(p-1)$-form B_{p-1} (globally defined on M, see below) such that $A_p = dB_{(p-1)}$. Clearly, because $d^2 = 0$ every exact form is also a closed form.

$$A_{(p)} = dB_{(p-1)} \rightarrow dA = ddB = 0 \tag{D.17}$$

It is natural to ask to what extent the reverse is true. In general, it is not. There exist manifolds where there are closed forms which are not exact. We will see one example below.

However, there is one important case where the reverse is true, and every closed form is also exact:

Poincare lemma: In $\mathbf{R^n}$, any closed p-form, $p > 0$, is also exact.

(since there are no (-1)-forms, clearly 0-forms can never be exact). A simple example is provided by 1-forms in $\mathbf{R}$. Any 1-form $A = f(x)dx$ in $\mathbf{R}$ can be written as $A = dF$, where F is the 0-form (i.e. function)

$$F(x) = \int_0^x f(y)\, dy \tag{D.18}$$

This is very important, and can be exploited to define a topological invariant for any differentiable manifold M, the cohomology of M. The argument is as follows.

Recall that M is a bunch of open sets U_α isomorphic to $\mathbf{R^n}$, glued in some way (specified by the transition functions $x_{(\beta)}(x_{(\alpha)})$). A p-form (field) $A_{(p)}$ is a bunch of p-forms $A^\alpha_{(p)}$ defined on the U_α's, which agree on the intersections $U_\alpha \bigcap U_\beta$

$$A^\alpha_{i_1 \ldots i_p} = \frac{\partial x^{i_1}_{(\alpha)}}{\partial x_{(\beta)^{j_1}}} \cdots \frac{\partial x^{i_p}_{(\alpha)}}{\partial x_{(\beta)^{j_p}}} A^\beta_{j_1 \ldots j_p} \tag{D.19}$$

A closed p-form satisfies $dA = 0$ globally, hence $dA^\alpha = 0$ on every U_α. Since each U_α is essentially $\mathbf{R^n}$, Poincare ensures that there always exists some $(p-1)$-form B^α in U_α such that $A^\alpha = dB^\alpha$. However, there is no guarantee that the B^α glue in the right way at intersections to define a global $(p-1)$-form B satisfying $A = dB$ globally. If this is not the case then A is closed but not exact.

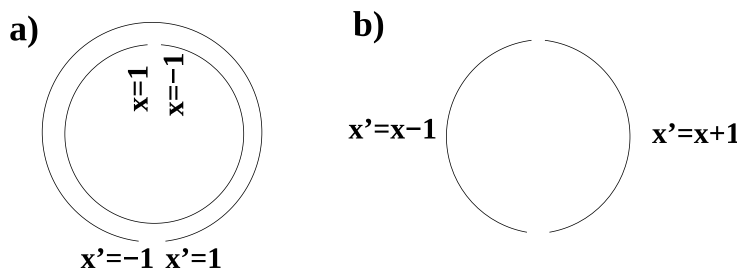


Figure D.2: Covering the circle with two charts.

In this argument, the local structure of M is not relevant, only the global structure, defined by how the U_α patch together, is relevant. Therefore, the existence of closed forms which are not exact is an statement which depends only on the global topology of M, and not on its local properties.

To give a simple example, consider the circle $\mathbf{S^1}$, described using two charts with local coordinates x, x', as shown in figure D.2, running in $(-1, 1)$, each covering $\mathbf{S^1}$ except the norht and south poles repectively. The intersection is disjoint, and on its two disconneted pieces the transition functions are $x' = x + 1$ and x'=x-1. Let us construct a global 1-form A, by glueing together the 1-form dx on U and dx' on U'; note they glue nicely with the above transition functions. The global 1-form is closed, and on U and U' is is locally exact, it reduces to dx or dx'. However, it is not possible to patch together x and x' to form a 0-form f such that $A = df$ globally (this would be as much as finding a coordinate valid globally on $\mathbf{S^1}$, which is not possible). By a strong and misleading abuse of language, the global 1-form is often referred to as dx, although we know that x is not a global 0-form.

The natural object which can be defined from these observations, and which depends only on the global structre of M is the de Rahm cohomology groups. Let $\mathcal{Z}^p$ be the set of closed p-form on M

$$\mathcal{Z}^p = \{ A_{(p)} \, | dA_{(p)} = 0 \, \} \tag{D.20}$$

and $\mathcal{B}_p$ the set of exact p-forms on M

$$\mathcal{B}^p = \{ A_{(p)} \, | \, A_{(p)} = dB_{(p-1)} \text{ for some } B_{(p-1)} \} \tag{D.21}$$

Since $\mathcal{B}^p \subset \mathcal{Z}^p$, we can define the quotient

$$H^p(M, \mathbf{R}) \; = \; \frac{\mathcal{Z}^{\mathbf{p}}}{\mathcal{B}^{\mathbf{p}}} \tag{D.22}$$

known at p^{th} de Rahm cohomology group of M. It is the set of closed forms of M modulo the equivalence relation

$$A_{(p)} \simeq A_{(p)} + dB_{(p-1)} \tag{D.23}$$

Namely, two closed p-forms define the same equivalence class in cohomology if they differ by an exact form. Notice that exact forms are also closed, they correspond to the zero (or trivial) class in cohomology (the class corresponding to an identically vanishing form). We denote by $[A]$ the cohomology class of a closed form A.

The sets $H^p(M, \mathbf{R})$ have the structure of finite-dimensional vector spaces (so in particular they are groups with respect to addition). Their structure depens only on the topology of M. Their dimensions, denoted b_p and known as Bettin numbers of M, are the simplest topological invariants of manifolds.

D.1.5 Homology

We now aim at defining a related class of topological quantities. To define them we need some additional concepts.

An m-dimensional submanifold N of M ($m < n$) is a subset of M which has the structure of an m-dimensional differential manifold. We will be interested in allowing for submanifolds with boundary, so we define the concept of boundary of a manifold.

A manifold M with boundary is a topological set together with an atlas with two kinds of charts: the familiar $(U_\alpha, x_{(\alpha)}$ and charts $(V_\beta, x_{(\beta)})$, where V_β is isomorphic to an open set in 'half' $\mathbf{R^n}$. As before, the charts cover M, and the $x_{(\alpha)}$, $x_{(\beta)}$ define differentiable transition functions. By 'half' $\mathbf{R^n}$ we mean the set of point

$$\mathbf{R^n_+} = \{(\mathbf{x^1}, \ldots, \mathbf{x^n}) | \mathbf{x^1} \geq \mathbf{0}\} \tag{D.24}$$

The boundary ∂M of M is the set of points which are anti-images of the points $x^1 = 0$ in the maps $x_{(\beta)}$. See figure D.3. It is important, although we do not discuss it in detail, to notice that the orientation in a manifold induces a natural orientation on its boundary.

A p-chain $a_{(p)}$ is a formal linear combination (with real coefficients) of p-dimensional submanifolds N_k (possibly with boundary) of M, namely $a = c_k N_k$.

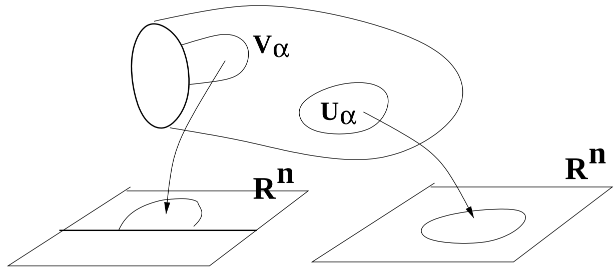


Figure D.3: Manifolds with boundary are described by two kinds of charts.

The operation of taking the boundary, which we call ∂, can be regarded as a linear operator mapping a p-chain to a $(p-1)$-chain, by

$$\partial a_{(p)} = c_k \, \partial N_k \tag{D.25}$$

An essential property of ∂, which is geometrically obvious is that

$$\partial^2 = 0 \tag{D.26}$$

In the sense that for any p-chain, $\partial(\partial a) = \emptyset$ is empty.

A p-chain $a_{(p)}$ without boundary is called a p-cycle, $\partial a_{(p)} = 0$. A p-chain is called trivial if it is the boundary of a $(p+1)$-chain, namely $a_{(p)} = \partial b_{(p+1)}$. Clearly, because $\partial^2 = 0$ any trivial p-chain is a p-cycle.

$$a_{(p)} = \partial b_{(b+1)} \;\rightarrow\; \partial a = \partial^2 b = 0 \tag{D.27}$$

It is natural to wonder to what extent the reverse is true. In general it is not: there exist manifolds M where there are p-cycles which are not trivial. An example of non-trivial 1-cycles is shown in figure D.4.

However, there is an important n-dimensional manifold where any p-cycle ($p < n$) is trivial [1]. This is the case for $\mathbf{R^n}$, see figure D.5. Again, this implies that the existence of non-trivial p-cycles on a manifold M is determined by the global structure of M, how it is patched together. It is ia features insensitive to the local structure of M, since locally it looks like $\mathbf{R^n}$, where all p-cycles are trivial.

We are now ready to define the p^{th} homology group $H_p(M, \mathbf{R})$. Let $\mathcal{Z}_p$ be the set of p-cycles

$$\mathcal{Z}_p = \{a_{(p)} | \partial a_{(p)} = 0\} \tag{D.28}$$

[1]Since there are no $(n+1)$-cycles in an n-dimensional space, n-chains cannot be trivial.

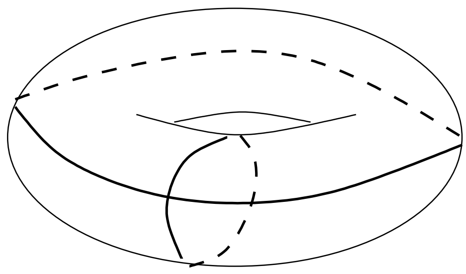


Figure D.4: Non-trivial 1-cycles in a two-torus.

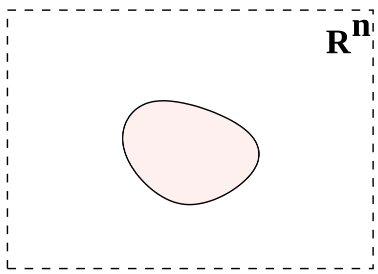


Figure D.5: All cycles in $\mathbf{R^n}$ are boundaries of some higher dimensional chain.

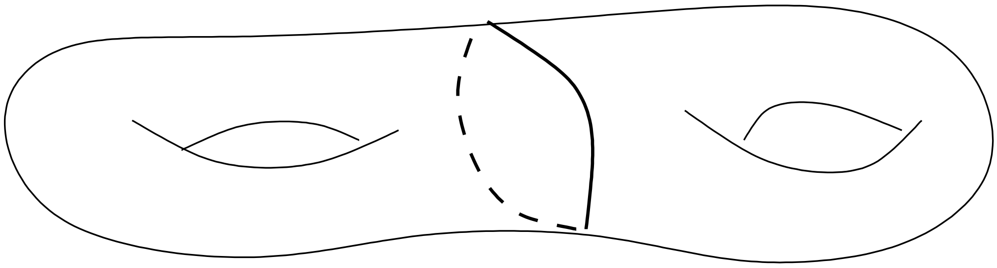


Figure D.6: A homologically trivial 1-cycles which is not homotopically trivial.

and let $\mathcal{B}_p$ be the set of trivial p-chains

$$\mathcal{B}_p = \{a_{(p)} | a_{(p)} = \partial b_{(p+1)}\} \tag{D.29}$$

Since $\mathcal{B}_p \subset \mathcal{Z}_p$, we can define the quotient

$$H_p(M, \mathbf{R}) = \frac{\mathcal{Z}_\mathbf{p}}{\mathcal{B}_\mathbf{p}} \tag{D.30}$$

known as the p^{th} homology group of M. It is formed by the set of p-cycles modulo the equivalence

$$a_{(p)} = a_{(p)} + \partial b_{(b+1)} \tag{D.31}$$

namely two p-cycles define the same homology class if they differ by a boundary. Trivial p-cycles correspond to the zero class in homology. We denote by $[a]$ the homology class of a cycle a. The spaces $H_p(M, \mathbf{R})$ have the structure of vector spaces, and their structure depends only on the topology of M. The dimension of $H_p(M, \mathbf{R})$ will be seen to be equal to b_p, i.e. the dimension of $H^p(M, \mathbf{R})$.

Examples of non-trivial 1-homology classes on $\mathbf{T^2}$ are shown in figure D.4. It is important to point out that homology is not the same as homotopy. In particular, homotopically trivial cycles (contractible cycles) are always homologically trivial (boundaries), but homologically trivial cycles may not be homotopically trivial. One example is shown in figure D.6.

D.1.6 de Rahm duality

We can notice a close analogy between the construction of cohomology and homology groups, as follows

closed form	cycle
exact form	trivial chain
d	∂
$H^p(M > \mathbf{R})$	$H_p(M, \mathbf{R})$

Indeed this is not accidental. There is a duality between the vector spaces $H^p(M, \mathbf{R})$ and $H_p(M, \mathbf{R})$ which explains the analogies in their construction. The duality is obtained via the operation of integration of forms over chains.

We define the integral of a p-form $A_{(p)}$ over a p-dimensional submanifold N of M, by splitting A into pieces A^α in the U_α, and integrating the components of A over the U_α in the usual calculus sense

$$\int_N A_{(p)} = \sum_\alpha \int_{U_\alpha} A^\alpha_{i_1 \ldots i_p} dx^1 \, \ldots dx^n \tag{D.32}$$

In fact, we should define this more carefully so as to make sure that we do not overcount the points of M, because of overlapping of the patches U_α. Each point in M should count only once in the integral. This can be done by using partitions of unity (see e.g. [119], but we will not enter into this detail, hoping the idea is clear. Note that on the overlaps it does not matter which coordinates we use, since the integrand is invariant under coordinate transformations (the change of the form component is an inverse jacobien which cancels agains the change of the differential calculus measure).

One can now define the integral of a p-form $A_{(p)}$ over a p-chain $a_{(p)} = \sum_k c_k N_k$ by

$$\int_{a_{(p)}} A_{(p)} = \sum_k c_k \int_{N_k} A_{(p)} \tag{D.33}$$

An important property is Stokes theorem, which states that for any $(p-1)$-form $B_{(p-1)}$ and p-chain $a_{(p)}$,

$$\int_{a_{(p)}} dB_{(p-1)} = \int_{\partial a_{(p)}} B_{p-1} \tag{D.34}$$

A simple example is provided by 0-forms (functions) and the 1-chain $[0, 1]$ (or other similar chains of closed sets in $\mathbf{R}$)

$$\int_{[0,1]} df \stackrel{\text{def}}{=} \int_{[0,1]} \frac{\partial f}{\partial x} dx = f(x)|_{x=0}^{x=1} = f(0) - f(1) = \int_{\partial[0,1]} f \tag{D.35}$$

(since the natural definition of an integral of a 0-form f over a 0d space (point) is simply evaluation of f at the point; the sign is due to opposite induced orientations).

Very interestingly, the integral of a closed p-form $A_{(p)}$ over a p-cycle $a_{(p)}$ depends only of their cohomology and homology classes, $[A]$ and $[a]$, respectively. Namely, the integral is unchanged if we take a different closed p-form $A'_{(p)}$ and a different p-cycle $a'_{(p)}$ in the same class $A'_(p) = A_{(p)} + dB_{(p-1)}$, $a'_{(p)} = a_{(p)} + \partial b_{(p+1)}$.

$$\begin{aligned} \int_a A' &= \int_a A + \int_a dB = \int_a A + \int_{\partial a} B = \int_a A \\ \int_{a'} A &= \int_a A + \int_{\partial b} A = \int_a A + \int_b dB = \int_a A \end{aligned} \tag{D.36}$$

This is often called the period of $[A]$ over $[a]$.

This implies that integration is well defined for cohomology and homology classes, since it does not depend on the particular representatives chosen. Thus integration define a linear mapping $H^p(M,\mathbf{R}) \times \mathbf{H_p}(\mathbf{M},\mathbf{R}) \to \mathbf{R}$. Equivalently, this shows that $H^p(M,\mathbf{R})$ is the vector space dual to $H_p(M,\mathbf{R})$, and vice versa. Namely, a p-cohomology class $[A_{(p)}]$ can be regarded as a linear mapping

$$\begin{aligned} [A_{(p)}] \, : \quad H_p(M,\mathbf{R}) &\longrightarrow \mathbf{R} \\ [a_{(p)}] \quad &\longmapsto \int_{a_{(p)}} A_{(p)} \end{aligned} \tag{D.37}$$

This implies the promised result that the dimensions of the p^{th} cohomology and homology groups are the same.

Notice that the duality implies that it is always possible to choose basis of cycles $\{a_i\}$ and forms $\{A_j\}$ such that $\int_{[a_i]}[A_j] = \delta_{ij}$. An example in $\mathbf{T^2}$ is given by the 1-forms dx, dy on the two independent circles, and the nontrivial 1-cycles.

D.1.7 Hodge structures

Now consider that M is a Riemannian manifold, i.e. it is endowed with a metric g of euclidean signature. The previous structures are topological and independent of the metric (they were constructed without any metric at all). In the presence of a metric, we can define some additional structures which are important, but not topologically invariant.

We define the Hodge operation $*$ as the map between p-forms and $(n-p)$-forms defined by the action on the basis

$$*(dx^{i_1} \wedge \ldots \wedge dx^{i_p}) \; = \; \frac{1}{(n-p)!}\sqrt{\det g}g^{i_1 j_1} \ldots g^{i_p j_p} \epsilon_{j_1 \ldots j_p j_{p+1} \ldots j_n} \, dx^{j_{p+1}} \wedge \ldots \wedge dx^{j_n} \tag{D.38}$$

It has the property that for a p-form $A_{(p)}$, $* * A_{(p)} = (-1)^{p(n-p)} A_{(p)}$.

The Hodge operator defines an positive-definite inner product betwen p-forms

$$(A_{(p)}, B_{(p)}) = \int_M A_{(p)} \wedge *B_{(p)} \tag{D.39}$$

Notice that this is not topological (however it is very important in physics, since it corresponds to

$$(A_{(p)}, B_{(p)}) = \int_M \sqrt{\det g} A_{i_1 \ldots i_p} B^{i_1 \ldots i_p} dx^1 \ldots dx^p \tag{D.40}$$

which is used to define the kinetic term of $(p-1)$-form gauge fields $C_{(p-1)}$ by taking $A_{(p)} = B_{(p)} = dC_{(p-1)}$ the gauge invariant field strength).

It is natural to define the adjoint $d^\dagger$ of d with respect to this inner product, i.e. it is defined by

$$(A_{(p)}, dB_{(p-1)}) = (d^\dagger A_{(p)}, B_{(p-1)}) \tag{D.41}$$

Hence $d^\dagger$ maps p-forms to $(p-1)$-forms. One can check that $d^\dagger = *d*$ for n even and $d^\dagger = (-1)^p * d*$ for n odd.

There is a theorem that ensures that any p-form $A_{(p)}$ has a *unique* decomponsition (known as Hodge decomposition) as

$$A_{(p)} = B_{(p)} + dC_{(p-1)} + d^\dagger D_{(p+1)} \tag{D.42}$$

with $B_{(p)}$ a harmonic form, namely obeys $dB_{(p)} = 0$, $d^+ B_{(p)} = 0$.

For closed p-forms, $dA_{(p)} = 0$ implies $dd^\dagger D_{(p+1)} = 0$. Taking the inner product with $D_{(p+1)}$,

$$(D_{(p+1)}, dd^\dagger D_{(p+1)}) = 0 \to (d^\dagger D_{(p+1)}, d^\dagger D_{(p+1)}) = 0 \tag{D.43}$$

the positive definiteness of the product implies $d^\dagger D_{(p+1)} = 0$. Then

$$A_{(p)} = B_{(p)} + dC_{(p-1)} \tag{D.44}$$

Thus in the cohomology class $[A]$ there is a unique harmonic p-form representative.

Namely, for each p-cohomology class, there exists a unique harmonic representative. Namely the p^{th} Betti number b_p is the number of independent harmonic p-forms on M. These are interesting statements: although the metric determines which particular p-form in the class is the harmonic one, the statement that there exists a unique one is independent of the metric. This is one simple example of a result which is topological invariant, but which is reached using additional non-topolocial structures, like a metric (there is no paradox, the result is independent of the metric chosen). Later on we will find more involved topological invariants which are easily defined using additional structures, althogh they are independent of the particular choices of these additional structures.

Harmonic p-forms will be quite useful in the study of KK compactification on curved spaces. Namely, the harmonic forms will provide the internal part of wavefunctions of the zero modes in the KK reduction of 10d p-form gauge fields. See lecture on Calabi-Yau compactification.

Another useful property due to Hodge operation is Poincare duality. The Hodge operator induces a homomorphism between $H^p(M, \mathbf{R})$ and $H^{n-p}(M, \mathbf{R})$. This can be seen by starting with a p-cohomology class, taking its harmonic representative, taking its Hodge dual (which is also harmonic) and finally taking the corresponding $(n-p)$-cohomology class.

This implies in particular $b_p = b_{n-p}$. Again this is an statement which we reach by using a metric, but is a topological statement.

Another consequence is that for any p-homology class $[a_{(p)}]$ we can define the Poincare dual $(n-p)$-cohomology class $[A_{(n-p)}]$, such that for any p-form $B_{(p)}$

$$\int_{a_{(p)}} B_{(p)} = \int_M B_{(p)} \wedge A_{(n-p)} \qquad \text{(D.45)}$$

Intuitively, $[A_{(n-p)}]$ can be considered as the class of a $(n-p)$-form 'delta function' with support on the volume of any p-cycle $a_{(p)}$ in the class $[a_{(p)}]$, see figure D.7.

Finally for completeness we define the intersection numbers of a p-cycle and $a_{(p)}$ and an $(n-p)$-cycle $b_{(n-p)}$ to be

$$\#(a_{(p)}, b_{(n-p)}) = \int_M A_{(n-p)} \wedge B_{(p)} \qquad \text{(D.46)}$$

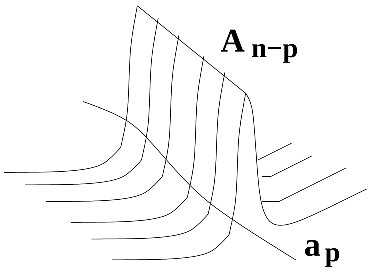


Figure D.7: The Poincare dual form of a cycle can be though of as a delta function (bump form) with support on the cycle.

where $A_{(n-p)}$, $B_{(p)}$ are the Poincare dual forms. Recalling the interpretation of Poincare dual forms as 'delta functions' localized on the corresponding cycles, the above number is an integer which counts the number of intersections (weighted with signs due to orientations) of the cycles $a_{(p)}$ and $b_{(n-p)}$.

D.2 Fiber bundles

Fiber bundles are a useful geometric concept in physics when studying fields that transform not only with respect to spacetime coordinate changes, but also have some particular behaviour under some internal gauge symmetries.

D.2.1 Fiber bundles

A vector bundle or fiber bundle E over a differential manifold M is a family of vector spaces V_P for each $P \in M$ (all isomorphic to an m-dimensional vector space V), which varies smoothly with P. V_P is called the fiber of E over the point P. The spaces M and V are referred to as the base and fiber of the bundle.

Equivalently, E can be defined with a set of charts $(U_\alpha \times V, (v_{(\alpha)}, x_{(\alpha)}))$, with $(U_\alpha, x_{(\alpha)})$ being charts describing M, and $v_{(\alpha)}$ being coordinates in V, such that on $U_\alpha \bigcap U_\beta$ coordinates on the base and fiber are related by

$$x_{(\beta)} \circ x_{(\alpha)})^{-1}$$

$$v_{(\beta)} = R_{(\alpha\beta)}(x_{(\alpha)}) \cdot v_\alpha \qquad \text{(D.47)}$$

Figure D.8: Construction of the Moebius strip as a nontrivial bundle with fiber **R** over $\mathbf{S^1}$.

where $R_{(\alpha\beta)}$ are (point dependent) matrices in $GL(m, \mathbf{R})$, known as transition functions. Notice that coordinate indices in V are implicit here (α, β denote the patches).

Intuitively, a bundle is locally identical to $\mathbf{R^n} \times \mathbf{V}$, and different local patches are glued on the base, and on the fiber, up to a linear transformation on the fiber.

A bundle E is therefore specified by the set of patches $U_\alpha \times V$ and the transition functions for the base and fiber, the latter satisfying the consistency condition $R_{(\alpha\gamma)} R_{(\gamma\beta)} R_{(\beta\alpha)} = 1$.

The total bundle E has a natural projection π to the base M given by the map defined by 'forgetting the fiber'

$$\begin{aligned} \pi \quad : \quad & E \longrightarrow M \\ & (P, v) \longmapsto P \end{aligned} \tag{D.48}$$

The simplest example of bundle is a trivial bundle, which is simply a space of the form $M \times V$. All transition functions $R = 1$ in this case.

A less trivial example is given by a Moebius strip. Consider $M = \mathbf{S^1}$, and $V = \mathbf{R}$. To build the bundle, cover $\mathbf{S^1}$ with two patches U, U' with coordinates x, x', as in section 1.4 and put coordinates y, y' on **R**, and use the glueing conditions

$$x' = x + 1 \; y' = y \quad \text{and} \quad x' = x - 1 \; y' = -y \tag{D.49}$$

on the two disconnected pieces of $U \bigcap U'$. The result is a non-trivial bundle. This is schematically shown in figure D.8.

A richer example is provided by the bundle formed by all tangent spaces $T_P(M)$ to a manifold M. The base is M, the fiber over $P \in M$ is $T_P(M)$, and the transition functions on the fiber on $U_\alpha \bigcap U_\beta$ are

$$v^i_{(\beta)} = \frac{\partial x^i_{(\beta)}}{\partial x^j_{(\alpha)}} v^j_{(\alpha)} \tag{D.50}$$

Similarly one can define the cotangent bundle, the tensor bundles, the p-form bundle, etc...

A section σ of a bundle E is a mapping, such that $\pi \circ \sigma = 1$, i.e. of the form

$$\begin{aligned} \sigma \quad : \quad & M \longrightarrow E \\ & P \longmapsto (P, \sigma(P)) \end{aligned} \tag{D.51}$$

That is for each point $P \in M$ we pick a point (vector) in V_P.

A simple example is a vector field, which is a section of the tangent bundle: $V^i(x)\partial_i$ defines a tangent vector for each point x on M. Similarly the cotangent vector fields, tensor fields, p-form fields,... are sections of the corresponding bundles.

D.2.2 Principal bundles, associated bundles

It is useful to extend the notion of vector bundle to other possible fibers with some structure.

A principal G-bundle is a bundle where the fiber is a group G [2]. Namely, on the overlaps of the patches of the base $U_\alpha \bigcap U_\beta$, the fibers (which are isomorphic to G) are glued up to an (point dependent) tranformation in G. The elements of the fiber G in U_α and U_β, denoted $g_{(\alpha)}$ and $g_{(\beta)}$ are related by

$$g_{(\beta)} = f_{\alpha\beta}(x_{(\alpha)}) g_{(\alpha)} f_{\alpha\beta}(x_{(\alpha)})^{-1} \tag{D.52}$$

This kind of bundle underlies the geometric description of gauge theories. For instance, a gauge transformation is nothing but a section of a principal G-bundle: $g(x)$ a group element for each point of M.

[2] We will center on compact Lie groups in this lecture.

When we have a group G, we can consider its representations R and the representation vector spaces on which the group acts. Given a principal G-bundle we can define the associated fiber bundles, which are vector bundles with the fiber the represention space of a representation R of G, and transition functions on the fiber

$$v_{(\beta)} = R(f_{(\alpha\beta)}) \cdot v_{(\alpha)} \tag{D.53}$$

In a gauge theory, fields in a representation R of the gauge group are sections of the corresponding associated bundle. The fact that the transition funcions for different associated bundles are simply different representations of the same transitions function of the principal G-bundle reflects the fact that the gauge group is unique, and we only have different fields charged differently under it. With the above definitions, all the gauge transformation properties of fields charged under a gauge group are recovered.

Notice that a general vector bundle can be regarded as the associated bundle of a principal $GL(m,\mathbf{R})$-bundle (corresponding to the vector representation of $GL(m,\mathbf{R})$. (since the transition funcions are matrices, which represent the action of the group $GL(m,\mathbf{R})$ on vectors of V).

D.3 Connections

In physics, vector bundles usually come equipped with an additional structure, a connection. The main idea is that in a vector bundle there is in principle no canonical way to compare two basis of the fiber at different points. A connection is an additional structure which allows to do so.

In a bundle with connection, in a patch where the point P has coordinates x^i, the canonical change of a basis $\{e^a\}$ of V_P as P changes in the direction i is given by

$$D_i e^a(x) = \partial_i e^a(x) + \omega_i{}^a{}_b(x) e^b(x) \tag{D.54}$$

where ω is the connection. On overlaps $U_\alpha \bigcap U_\beta$ the connection transforms not just as a 1-form, but has the additional transformation

$$\omega_{i,(\beta)} = R_{(\alpha\beta)}\omega_i R^{-1}_{(\alpha\beta)} - (\partial_i R_{(\alpha\beta)})R^{-1}_{(\alpha\beta)} \tag{D.55}$$

which ensures that for a section σ of E, its covariant derivative $D_i\sigma(x)$ transforms as a section of E as well.

There are two classes of physical theories where fiber bundles with connections appear. The first is the case of gauge theories, where charged fields are sections of bundles associated to a principal G-bundle, and carry connections given by the representation of the connection of the principal G-bundle. For a representation R of G, the associated bundle has connection $\omega_i{}^a{}_b = A_i^m (T_R)^a_b$, where A is the connection on the principal G-bundle, m runs over the generators of the Lie algebra, and T_R is the representation of a generator in the representation R.

The second situation is in theories of gravity. The introduction of a metric g in a manifold M can be described in terms of fiber bundles as follows. At each point $x \in M$ introduce a set of tangent vectors $\{e^a(x)\}$, orthonormal with respect to the metric g

$$g_{ij} e^{a,i} e^{b,j} = \delta^{ab} \tag{D.56}$$

which also implies $e_i^a e_{a,j} = g_{ij}$. All the information of the metric is encode in the 'tetrad' $\{e^a\}$.

The tetrad is however defined up to $SO(N)$ rotations at each point, so this behaves as a local gauge invariance of the system. Indeed, such local rotations are sections of a principal $SO(N)$-bundle, and the tangent bundle is an associated bundle (for the vector representation).

Clearly one can construct other associated bundles; one of the most interesting ones is the spinor bundle, whose associated connection (see below) is known as the spin connection.

The metric induces a preferred connection on the tangent bundle, namely the Christoffel connection on vectors. One can then obtain a connection in terms of the tetrad, from the condition

$$D_i e_j^a = \partial_i e_j^a - \Gamma_{ij}^k e_k^a + \omega_i{}^a{}_b e_j^b = 0 \tag{D.57}$$

which defines a connection in the principal $SO(N)$-bundle. The latter then defines connections in all associated bundles, like the spinor bundle. In fact the tetrad formalism was originally deviced to be able to define parallel transport of spinors.

Given a general connection on a fiber bundle, we define its curvature by

$$R_{ij}^a{}_b = \partial_i \omega_j{}^a{}_b - -\partial_j \omega_i{}^a{}_b + [\omega_i, \omega_j]^a_b \tag{D.58}$$

they behave as 2-form with respect to coordinate reparametrizations, and transform covariantly under gauge transformations.

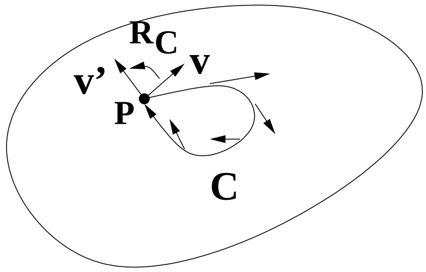


Figure D.9: The holonomy group is given by the set of rotation R_C suffered by a vector under parallel transport around all possible closed loops in the manifold.

In gauge theories, the curvature of the connection of the principal bundle are denoted $F_{ij} = F_{ij}^m t^m$, where t^m are the Lie algebra generators. In a vector bundle associated to a representation R, it is given by $F_{ij}{}^a{}_b = F_{ij}^m (T_R^m)^a_b$.

D.3.1 Holonomy of a connection

We start with a vector bundle E (with fiber V over a base manifold M) with connection. Consider a point $P \in M$, and consider the set of closed loops which start and end at P. It is a group under the operation of adjoining loops. Consider a vector v in the fiber V_P and parallel transport it along a loop C with the connection. It will come back to a vector v' in V_P, related to the original v by some $GL(m, \mathbf{R})$ rotation R_C. The set of such rotations for all closed loops is a group, known as the holonomy group of the connection. See figure D.9.

For a connection induced from a metric, the holonomy of the connection is often referred to as the holonomy of the metric or of the Riemannian manifold.

For a metric connection, the norm of the tangent vector is preserved during parallel transport, hence the holonomy of the connection is necessarily a subgroup of $SO(n)$. For a principal G-bundle, and its associated bundles, like in gauge theories, the holonomy of a connection is necessarily a subgroup of G.

D.3.2 Characteristic classes

Our motivation is to construct topological quantities for fiber bundles, that characterize non-trivial bundles. In this section we see that there are certain quantities, which are computed using additional structures, like metrics or connections, but which at the end turn out to be independent of the particular metric or connection chosen. They are therefore topological. Before constructing them, it will be useful to give a simple example of a non-trivial fiber bundle.

The Wu-Yang magnetic monopole

Consider a $U(1)$ gauge theory on $M = \mathbf{S^2}$. The underlying geometry is a principal $U(1)$-bundle over the base $\mathbf{S^2}$. Let us classify all topologically inequivalent non-trivial gauge bundles. To do so, we cover $\mathbf{S^2}$ with two open sets, U_+ and U_-, which cover the North and South hemispheres, see figure D.10. The bundle over each patch is trivial, so all the information about the bundle over $\mathbf{S^2}$ is encoded in the transition function in $U_+ \bigcap U_-$, which is an $\mathbf{S^1}$, the equator. For a principal $U(1)$-bundle, the transition function $g(\phi)$ takes values on $U(1)$ which is also a circle. Therefore the topologically different bundles are classified by topologically different maps from the equator $\mathbf{S^1}$ to the fiber $\mathbf{S^1}$. Such topologically different maps are classified by the homotopy group $\Pi_1(\mathbf{S^1}) = \mathbf{Z}$. Namely, there exist inequivalent classes of maps labeled by an integer. Simple representatives of these maps are

$$\begin{aligned} g_n \quad : \quad & \mathbf{S^1} \longrightarrow \mathbf{S^1} \\ & e^{i\phi} \longmapsto e^{in\phi} \end{aligned} \tag{D.59}$$

Namely, the label n corresponds to how many time one goes around the target $\mathbf{S^1}$ when going once around the origin $\mathbf{S^1}$.

This example is simple enough to be more explicit about the connections we can put on these bundles (that is, the gauge field configurations corresponding to these bundles). Here we describe a simple case.

Consider polar coordinates θ, ϕ, and introduce the $U(1)$ gauge potentials on $U_\pm$

$$A_\pm = \frac{1}{2}\frac{\pm 1 - \cos\theta}{\sin\theta}\, d\phi \tag{D.60}$$

On the intersection of $U_\pm$, namely at $\theta = \pi/2$ they differ by a gauge transformation

$$A_+ - A_- = d\phi \tag{D.61}$$

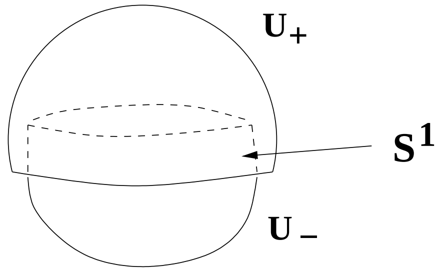


Figure D.10: .

so they define a global connection for the bundle. The curvatures on $U_\pm$ agree on the intersection $F_+ = F_-$.

The above expression shows that the transition function for this bundle is the map

$$\begin{aligned} g_1 \quad : \quad & \mathbf{S^1} \longrightarrow \mathbf{S^1} \\ & e^{i\phi} \longmapsto e^{i\phi} \end{aligned} \tag{D.62}$$

So the bundle is non-trivial.

There is a nice general relation between the winding of the map $g(\phi)$ and the flux of F on $\mathbf{S^2}$. This provides a way of characterizing non-trivial bundles which we will generalize in next section. In a bundle defined by the transition function g_n, the gauge potentials $A_\pm$ satisfy $A_+ - A_- = nd\phi$ on the equator. Hence we have

$$\int_{\mathbf{S^2}} F = \int_{U_+} F_+ + \int_{U_-} F_- = \int_{U_+} dA_+ + \int_{U_-} dA_- = \int_{\mathbf{S^1}} A_+ - \int_{\mathbf{S^1}} A_- = \int_{\mathbf{S^1}} n\, d\phi = 2\pi n \tag{D.63}$$

This example is familiar in the study of magnetic monopoles: When the $\mathbf{S^2}$ is taken to describe the angular part of 3d space, the gauge configuration describes a magnetic monopole sitting at the origin or $\mathbf{R^3}$.

Since F is closed and its integral over $\mathbf{S^2}$ does not vanish, it defines a non-trivial cohomology class. Indeed, $\frac{F}{2\pi}$ defines an integer cohomoly class $[F/2\pi]$, which characterizes the bundle. Notice that although we used a connection to define this quantity, it finally depends only on the transition functions, and therefore is a topological invariant of the bundle. It is known as first Chern class of the bundle.

Another simple example of non-trivial bundle is obtained by considering a $U(1)$ gauge field configuration on $\mathbf{T^2}$, with a constant magnetic field; abusing of language, this can be described by a gauge potential $A = kxdy$.

A final example, familiar from nonabelian 4d gauge theories, is the classification of topological sectors of gauge configurations by the value of

$$k = \frac{1}{8\pi^2} \int_{4d} \mathrm{tr}\, F \wedge F \tag{D.64}$$

known as the instanton number of the configuration.

All these topological invariants are simple examples of characteristic classes. Let us generalize the $U(1)$ case for a general manifold M. To do that, on each U_α we introduce the local form of the connection A_α, such that on overlaps $U_\alpha \bigcap U_\beta$ we have

$$A_\beta = A_\alpha + d\phi_{(\alpha\beta)} \tag{D.65}$$

Then $F = dA_\alpha$ is globally defined, and satisfies $dF = 0$, hence defines a cohomology class $[F]$. We know show that his class is a topological invariant of the bundle. Namely, although to define it we have used a connection, the final class depends only on the transition functions of the bundle $\phi_{(\alpha\beta)}$, and is independent of the particular connection chosen.

To show that, consider a different connection defined by A'_α, still with the same transition functions

$$A'_\beta = A'_\alpha + d\phi_{(\alpha\beta)} \tag{D.66}$$

From (D.65) and (D.66), it follows that $A_\alpha - A'_\alpha = A_\beta - A'_\beta$ so the differences are patch independent and define a global 1-form B. Then $F - F' = dB$ globally, which implies that they define the same cohomology class $[F]$, as we wanted to show.

More sophisticated tools can be used to show that $[F/2\pi]$ is in fact an integer cohomology class, known as first Chern class of the $U(1)$ bundle.

The generalization to principal G-bundles with arbitrary group is analogous. One simply constructs polynomials in wedge products of the curvatures of the connection, tracing in the Lie algebra indices. The resulting form is closed and the corresponding cohomology class is a topological invariant of the bundle. These are known as characteristic classes.

We now give some examples appearing often for $SU(N)$ and $SO(N)$. Consider the closed $2k$-form

$$\Omega_{2k} = \sum_{m_1,\ldots,m_k} (F^{m_1} \wedge \ldots \wedge F^{m_k})\mathrm{Str}(t^{m_1} \ldots t^{m_k}) \tag{D.67}$$

where Str denoted the symmetrized trace of the generators. This is usually written $\Omega_{2k} = \mathrm{tr}\, F^k$ (with wedge products implied). The corresponding cohomology class is a topological invariant of the corresponding bundle. For $U(N)$ it is known as the k^{th} Chern class, and has the generating function

$$ch(E) = \mathrm{tr}\,(e^{F/2\pi}) \tag{D.68}$$

known as the Chern character. For $SO(2N)$, Ω_{2k} automatically vanishes unless k is even $k = 2r$; the cohomology class is in this case known as r^{th} Pontryagin class. The Pontryagin classes also appear often in a generating function

$$\hat{A} = 1 + \frac{1}{8\pi^2}\mathrm{tr}\, R^2 + \ldots \tag{D.69}$$

known as A-roof genus.

Characteristic classes are very useful in characterizing the topology of nontrivial bundles [3]. Clearly much more can be said about bundles and their characterization. However, this will be enough for our purposes and applications.

[3] Although this characterization is not complete, different gauge bundles may still have all characteristic classes equal, and differ in some additional topological quantities. We may see some of this in the discussion of K-theory when discussing stable non-BPS branes.

Bibliography

[1] L. E. Ibanez, 'Recent developments in physics far beyond the standard model', e-Print Archive: hep-ph/9901292

[2] G. G. Ross, 'Grand unified theories', Reading, Usa: Benjamin/cummings (1984) 497 P. (Frontiers In Physics, 60)

[3] A. Billoire, 'Grand unified theories', IN *GIF-SUR-YVETTE 1984, PROCEEDINGS, COSMOLOGY AND ELEMENTARY PARTICLES. HEAVY ION COLLISIONS*, 37-38.

[4] H. P. Nilles, 'Phenomenological aspects of supersymmetry', hep-ph/9511313;
S. P. Martin, 'A supersymmetry primer', hep-ph/9709356

[5] H. P. Nilles, 'Supersymmetry, supergravity and particle physics', Phys.Rept.110:1,1984;
D.G. Cerdeno, C. Munoz, 'An introduction to supergravity', JHEP PDF conference proceedings server (at http://jhep.sissa.it), Sep 1998.

[6] E. Witten, 'Fermion quantum numbers in Kaluza-Klein theory', Published in Shelter Island II 1983:227

[7] N. Arkani-Hamed, S. Dimopoulos, G.R. Dvali 'The hierarchy problem and new dimensions at a millimiter', Phys.Lett.B429:263-272,1998 , hep-ph/9803315; I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos, G.R. Dvali 'New dimensions at a millimiter to a fermi and superstrings at a TeV', Phys.Lett.B436:257-263,1998, hep-ph/9804398

[8] J. Polchinski, 'String theory', Vol 1.

[9] M.Green, J.Schwarz, E.Witten, 'Superstring theory', vol1.

[10] P. Ginsparg, 'Applied conformal field theory', Lectures given at Les Houches Summer School in Theoretical Physics, Les Houches, France, Jun 28 - Aug 5, 1988. Published in Les Houches Summer School 1988:1-168;
A.N. Schellekens, 'Introduction to conformal field theory', Based on lectures given at Grundlagen und neue Methoden der Theoretischen Physik, Saalburg, Germany, 3-16 Sep 1996 and at the Universidad Autonoma, Madrid, Oct-Dec 1995, Published in Fortsch.Phys.44:605-705,1996

[11] J. Polchinski, 'String theory', vol2, Cambridge univ press.

[12] J. Strathdee, 'Extended Poincare Supersymmetry', Int. J. Mod. Phys. A2 (1987) 273;
W. Nahm, 'Supersymmetries and their representations', Nucl. Phys. B135 (1978) 149.

[13] L. Alvarez-Gaume, E.Witten, 'Gravitational anomalies' Nucl.Phys.B234:269,1984.

[14] Michael B. Green, John H. Schwarz, 'Anomaly cancellation in supersymmetric D=10 gauge theory and superstring theory', Phys. Lett. B149: 117-122, 1984;

[15] Michael B. Green, John H. Schwarz, 'The hexagon anomaly in type I superstring theory', Nucl. Phys. B255: 93-114, 1985.

[16] E. Witten, 'Fermion quantum numbers in Kaluza-Klein theory', Published in Shelter Island II 1983:227

[17] P. Candelas, G. T. Horowitz, A. Strominger, E. Witten, 'Vacuum configurations for superstrings', Nucl. Phys. B258: 46-74, 1985

[18] B. R. Greene, 'The elegant universe: Superstrings, hidden dimensions, and the quest of the ultimate theory', New York, USA: Norton (1999) 448

[19] A. Sen, B. Zwiebach, 'Tachyon condensation in string field theory', JHEP 0003 (2000) 002.

[20] S. Coleman, 'Aspects of symmetry', Cambridge Univ. Press.

[21] J. A. Harvey, 'Magnetic monopoles, duality and supersymmetry' hep-th/9603086

[22] P.K. Townsend, 'P-brane democracy', hep-th/9507048.

[23] Edward Witten, 'String theory dynamics in various dimensions' Nucl. Phys. B443 (1995) 85, hep-th/9503124

[24] C.M. Hull, P.K. Townsend, 'Unity of superstring dualities' Nucl. Phys. B438 (1995) 109, hep-th/9410167.

[25] P. Horava, E, Witten, 'Heterotic and type I string dynamics from eleven-dimensions', Nucl. Phys. B460 (1996) 506, hep-th/9510209.

[26] J. Polchinski, 'Dirichlet Branes and Ramond-Ramond charges', Phys. Rev. Lett. 75 (1995) 4724, hep-th/9510017.

[27] N. Arkani-Hamed, S. Dimopoulos, G.R. Dvali 'The hierarchy problem and new dimensions at a millimiter', Phys.Lett.B429:263-272,1998 , hep-ph/9803315; I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos, G.R. Dvali 'New dimensions at a millimiter to a fermi and superstrings at a TeV', Phys.Lett.B436:257-263,1998, hep-ph/9804398

[28] see a pedagogical discussion in Cheng and Li, 'Gauge theory of elementary particle physics' Oxford, Uk: Clarendon (1984).

[29] T. Banks, W. Fischler, S.H. Shenker, L. Susskind, 'M theory as a matrix model: A Conjecture', Phys. Rev. D55 (1997) 5112, hep-th/9610043.

[30] A. Srominger, 'Massless black holes and conifolds in string theory', Nucl. Phys. B451(1995)96, th/9504090.

[31] B. R. Greene, D. R. Morrison, A. Strominger, 'Black hole condensation and the unification of string vacua', Nucl. Phys. B451 (1995) 109, hep-th/9504145.

[32] M.R. Douglas, G. W. Moore, 'D-branes, quivers, and ALE instantons', hep-th/9603167

[33] C. V. Johnson, A. W. Peet, J. Polchinski, 'Gauge theory and the excision of repulson singularities', Phys. Rev. D61 (2000) 086001, hep-th/9911161.

[34] C. Montonen, D. I. Olive, 'Magnetic monopoles as gauge particles?', Phys. Lett. B72 (1977) 117.

[35] J. M. Maldacena, 'The Large N limit of superconformal field theories and supergravity', Adv. Theor. Math. Phys. 2 (1998) 231, hep-th/9711200.

[36] G. 't Hooft, 'A planar diagram theory for strong interactions', Nucl. Phys. B72 (1974) 461.

[37] E. Witten, 'Strong coupling expansion of Calabi-Yau compactification', Nucl. Phys. B471 (1996) 135, hep-th/9602070.

[38] J. Polchinski, 'String theory', Vol 1.

[39] D. Lust, S. Theisen, 'Lectures on string theory', Lect. Notes Phys. 346 (1989)1

[40] D. Mumford, 'Tata lectures on theta', vols 1, 2, Birkhauser, Basel 1983

[41] S. H. Shenker, 'Another length scale in string theory?', hep-th/9509132; see also the D-brane scattering discussion in 'D-brane dynamics', by C. Bachas, Phys. Lett. B374 (1996) 37, hep-th/9511043.

[42] A. Giveon, M. Porrati, E. Rabinovici, 'Target space duality in string theory', Phys. Rept. 244 (1994) 77, hep-th/9401139.

[43] J. Polchinski, 'String theory', Vol 2.

[44] N. Seiberg, E. Witten, 'Spin structures in string theory', Nucl. Phys. B276 (1986) 272.

[45] L. Alvarez-Gaume, E. Witten, 'Gravitational anomalies' Nucl .Phys. B234 (1984) 269.

[46] J. Strathdee, 'Extended Poincare Supersymmetry' Int. J. Mod. Phys. A2 (1987) 273
W. Nahm, 'Supersymmetries and their representations', Nucl. Phys. B135 (1978) 149.

[47] L.J. Romans, ‘ Massive N=2A supergravity in ten dimensions’, Phys. Lett. B169 (1986) 374.

[48] D. J. Gross, J. A. Harvey, E. J. Martinec, R. Rohm, ‘Heterotic string theory. The free heterotic string’, Nucl. Phys. B256 (1985) 253.

[49] J. Polchinski, ‘String theory’, Vol 2.

[50] L. Alvarez-Gaume, ‘An introduction to anomalies’, Erice School Math. Phys. 1985 009.

[51] A.N. Schellekens, N.P. Warner, ‘Anomalies and modular invariance in string theory’, Phys. Lett. B177 (1986) 317.

[52] L. Alvarez-Gaume, E. Witten, ‘Gravitaional anomalies’ Nucl. Phys. B234 (1984) 269.

[53] J. Polchinski, ‘String theory’, Vol 1.

[54] J. Polchinski, Y. Cai, ‘Consistency of open superstring theories’, Nucl. Phys. B296 (1988) 91.

[55] J. Polchinski, ‘String theory’, Vol 1.

[56] J. Polchinski, ‘String theory’, Vol 2.

[57] R. Slansky, ‘Group theory for unified model building’ Phys. Rept. 79 (1981) 1.

[58] K.S. Narain, ‘New heterotic string theories in uncompactified dimensions ¡ 10’, Phys. Lett. B169 (1986) 41.

[59] K.S. Narain, M.H. Sarmadi, E. Witten, ‘A note on toroidal compactification of heterotic string theory’, Nucl. Phys. B279 (1987) 369.

[60] P. Ginsparg, ‘Comment on toroidal compactification of heterotic superstrings’, Phys.Rev.D35 (1987) 648.

[61] P. Candelas, G. T. Horowitz, A. Strominger, E. Witten, ‘Vacuum configurations for superstrings’, Nucl. Phys. B258 (1985) 46.

[62] E. Witten, ‘Phases of N=2 theories in two-dimensions’, Nucl. Phys. B403 (1993) 159.

[63] B. R. Greene, M.R. Plesser, 'Duality in Calabi-Yau moduli space', Nucl.Phys. B338 (1990) 15.

[64] L. J. Dixon, J. A. Harvey, C. Vafa, E. Witten, 'Strings on orbifolds, 1,2' Nucl. Phys. B261 (1985) 678, Nucl. Phys. B274 (1986) 285.

[65] L. E. Ibanez, 'The search for a standard model $SU(3) \times SU(2) \times U(1)$ superstring: An introduction to orbifold constructions. IN *MADRID 1987, PROCEEDINGS, STRINGS AND SUPERSTRINGS* 74-134 AND CERN GENEVA - TH. 4769 (87,REC.SEP.) 61.

[66] O.J. Ganor, J. Sonnenschein, 'On the strong coupling dynamics of heterotic string theory on C^3/Z_3, JHEP 0205 (2002) 018, hep-th/0202206.

[67] K.S. Narain, M.H. Sarmadi, C. Vafa, 'Asymmetric orbifolds' Nucl. Phys. B288 (1987) 551; Nucl. Phys. B356 (1991) 163.

[68] C. M. Hull, P. K. Townsend, 'Unity of superstring dualities', Nucl. Phys. B438 (1995) 109, hep-th/9410167.

[69] E. Witten,'String theory dynamics in various dimensions', Nucl. Phys. B443 (1995) 85, hep-th/9503124.

[70] N. A. Obers, B. Pioline, 'U duality and M theory, an algebraic approach', hep-th/9812139.

[71] J. Polchinski, 'String theory', Vol 2.

[72] S. Coleman, 'Aspects of symmetry', Cambridge Univ. Press.

[73] J. A. Harvey, 'Magnetic monopoles, duality and supersymmetry' hep-th/9603086

[74] A. Sen, 'Dyon - monopole bound states, selfdual harmonic forms on the multi - monopole moduli space, and SL(2,Z) invariance in string theory', Phys. Lett. B329 (1994) 217, hep-th/9402032.

[75] R.G. Leigh, 'Dirac-Born-Infeld action from Dirichlet sigma model', Mod. Phys. Lett. A4 (1989) 2767.

[76] L. J. Romans, 'Massive N=2A supergravity in ten dimensions', Phys. Lett. B169 (1986) 374.

[77] E. Witten, 'Small instantons in string theory', Nucl. Phys. B460 (1996) 541, hep-th/9511030.

[78] E. G. Gimon, J. Polchinski, 'Consistency conditions for orientifolds and d manifolds', Phys. Rev. D54 (1996) 1667, hep-th/9601038.

[79] T. Banks, W. Fischler, S. H. Shenker, L. Susskind, 'M theory as a matrix model: A Conjecture', Phys. Rev. D55 (1997) 5112, hep-th/9610043.

[80] J. H. Schwarz, 'An SL(2,Z) multiplet of type IIB superstrings', Phys. Lett. B360 (1995) 13, ERRATUM-ibid. B364 (1995) 252, hep-th/9508143;
P. S. Aspinwall, 'Some relationships between dualities in string theory', Nucl. Phys. Proc. Suppl. 46 (1996) 30, hep-th/9508154.

[81] J. Polchinski, E. Witten, 'Evidence for heterotic - type I string duality', Nucl. Phys. B460 (1996) 525, hep-th/9510169.

[82] P. Horava, E, Witten, 'Heterotic and type I string dynamics from eleven-dimensions', Nucl. Phys. B460 (1996) 506, hep-th/9510209.

[83] P. Horava, E, Witten, 'Eleven-dimensional supergravity on a manifold with boundary', Nucl. Phys. B475 (1996) 94, hep-th/9603142.

[84] P. S. Aspinwall, 'K3 surfaces and string duality', hep-th/9611137.

[85] P. S. Aspinwall, , D. R. Morrison, 'String theory on K3 surfaces', hep-th/9404151.

[86] C. M. Hull, P. K. Townsend, 'Unity of superstring dualities', Nucl. Phys. B438 (1995) 109, hep-th/9410167.

[87] E. Witten,'String theory dynamics in various dimensions', Nucl. Phys. B443 (1995) 85, hep-th/9503124.

[88] P. S. Aspinwall, 'Enhanced gauge symmetries and K3 surfaces', Phys. Lett. B357 (1995) 329, hep-th/9507012.

[89] E. Witten, 'Some comments on string dynamics', hep-th/9507121.

[90] A. Srominger, 'Massless black holes and conifolds in string theory', Nucl. Phys. B451(1995)96, th/9504090.

[91] B. R. Greene, D. R. Morrison, A. Strominger, 'Black hole condensation and the unification of string vacua', Nucl. Phys. B451 (1995) 109, hep-th/9504145.

[92] B. R. Greene, D. R. Morrison, C. Vafa, 'A geometric realization of confinement', Nucl. Phys. B481 (1996) 513, hep-th/9608039.

[93] A. Giveon, D. Kutasov, 'Brane dynamics and gauge theory', Rev. Mod. Phys. 71 (1999) 983, hep-th/9802067.

[94] N. Seiberg, E. Witten, 'Electric - magnetic duality, monopole condensation, and confinement in N=2 supersymmetric Yang-Mills theory', Nucl. Phys. B426 (1994) 19,Erratum-ibid. B430 (1994) 485, hep-th/9407087.

[95] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri, Y. Oz, 'Large N field theories, string theory and gravity', Phys. Rept. 323 (2000) 183, hep-th/9905111.

[96] J. M. Maldacena, 'The Large N limit of superconformal field theories and supergravity', Adv. Theor. Math. Phys. 2 (1998) 231, hep-th/9711200.

[97] D. Berenstein, J. M. Maldacena, H. Nastase, 'Strings in flat space and pp waves from N=4 superYang-Mills', JHEP 0204 (2002) 013, hep-th/0202021.

[98] E. Witten, 'Anti-de Sitter space and holography', Adv. Theor. Math. Phys. 2 (1998) 253, hep-th/9802150.

[99] S. S. Gubser, I. R. Klebanov, A. M. Polyakov, 'Gauge theory correlators from noncritical string theory', Phys. Lett. B428 (1998) 105, hep-th/9802109.

[100] J. M. Maldacena, 'Wilson loops in large N field theories', Phys. Rev. Lett. 80 (1998) 4859, hep-th/9803002.

[101] E.Witten, 'Baryons and branes in anti-de Sitter space', JHEP 9807 (1998) 006, hep-th/9805112.

[102] E. Witten, 'Anti-de Sitter space, thermal phase transition, and confinement in gauge theories', Adv. Theor. Math. Phys. 2 (1998) 505, hep-th/9803131.

[103] H. Ooguri, C. Vafa, 'World sheet derivation of a large N duality', Nucl. Phys. B641 (2002) 3, hep-th/0205297.

[104] E.Witten, 'Strong coupling expansion of Calabi-Yau compactification', Nucl. Phys. B471 (1996) 135, hep-th/9602070.

[105] J. D. Lykken, 'Weak scale superstrings', Phys. Rev. D54 (1996) 3693, hep-th/9603133.

[106] N. Arkani-Hamed, S. Dimopoulos, G.R. Dvali 'The hierarchy problem and new dimensions at a millimiter', Phys.Lett.B429:263-272,1998 , hep-ph/9803315; I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos, G.R. Dvali 'New dimensions at a millimiter to a fermi and superstrings at a TeV', Phys.Lett.B436:257-263,1998, hep-ph/9804398

[107] See e.g. R. Donagi, A. Lukas, B. A. Ovrut, D. Waldram, 'Nonperturbative vacua and particle physics in M theory', JHEP 9905 (1999) 018, hep-th/9811168, and other references by the authors.

[108] M. R. Douglas, G. W. Moore, 'D-branes, quivers, and ALE instantons', hep-th/9603167; M. R. Douglas, B. R. Greene, D. R. Morrison, 'Orbifold resolution by D-branes', Nucl. Phys. B506 (1997) 84, hep-th/9704151.

[109] G. Aldazabal, L. E. Ibáñez, F. Quevedo, A. M. Uranga, 'D-branes at singularities: A Bottom up approach to the string embedding of the standard model', JHEP 0008 (2000) 002, hep-th/0005067.

[110] M. Berkooz, M. R. Douglas, R. G. Leigh, 'Branes intersecting at angles', Nucl. Phys. B480 (1996) 265, hep-th/9606139.

[111] A. M. Uranga, 'Chiral four-dimensional string compactifications with intersecting D-branes', hep-th/0301032.

[112]

[113] S. Kachru, M. Schulz, S. Trivedi, 'Moduli stabilization from fluxes in a simple iib orientifold', hep-th/0201028.

[114] E. Witten, 'D-branes and K theory', JHEP 9812 (1998) 019, hep-th/9810188.

[115] A. Sen, 'NonBPS states and Branes in string theory', hep-th/9904207.

[116] M. R. Douglas, 'Branes within branes', hep-th/9512077

[117] P. Candelas, 'Lectures on complex manifolds', Trieste, proceedings Superstrings 87.

[118] M. B. Green, J.H. Schwarz, E. Witten. 'Superstring theory, vol 2: Loop amplitudes, anomalies and phenomenology', Cambridge, Uk: Univ. Pr. (1987) 596 P. (Cambridge Monographs On Mathematical Physics.

[119] Bott and Tu, 'Differential forms in algebraci topology', Springer-Verlag.

[120] H. Georgi, 'Lie algebras in particle physics', Benjamin, 1982.

[121] W.K. Tung, 'Group Theory in physics', World Scientific, Singapore, 1985.

[122] E. Wigner

[123] H. Weyl

[124] Appendices of R. Slansky, 'Group theory for unified model building', Phys. Rept. 79 (1981) 1

[125] J. Wess, J. Bagger, Supersymmetry and supergravity, Princeton Univ. Press, 1992.

[126] W. Nahm, Supersymmetries and their representations, Nucl. Phys. B135 (1978) 149.

[127] J. Strathdee, Extended Poincare supersymmetry, Int. J. Mod. Phys. A2 (1987) 273