26 Compactification of string theory II. Calabi–Yau compactifications

Up to now we have focused on rather simple models involving toroidal compactifications and their orbifold generalizations. But, while by far the simplest, these turn out to be only a tiny subset of the possible manifolds on which to compactify string theories. A particularly interesting and rich set of geometries is provided by the Calabi–Yau manifolds. These are manifolds which are Ricci flat, $R_{MN} = 0$. Their interest arises in large part because these compactifications can preserve some subset of the full ten-dimensional supersymmetry. This is significant if one believes that low-energy supersymmetry has something to do with nature. It is also important at a purely theoretical level since, as usual, supersymmetry provides a great deal of control over any analysis; at the same time there is less supersymmetry than in the toroidal case, so a richer set of phenomena is possible.

This chapter is intended to provide an introduction to this subject. In the first section we will develop some mathematical preliminaries. Unlike the toroidal or orbifold compactifications it is not possible, in most instances, to provide explicit formulas for the underlying metric on the manifold and other quantities of interest. The six-dimensional Calabi–Yau spaces, for example, have no continuous isometries (symmetries), so at best one can construct the metrics by numerical methods. But it turns out to be possible from topological considerations to extract much important information without a detailed knowledge of the metric. The machinery required to define these spaces and to extract at least some of this information includes algebraic geometry and cohomology theory, subjects not part of the training of most physicists. The following mathematical interlude provides a brief introduction to the necessary mathematics. There is much more in the suggested reading.

26.1 Mathematical preliminaries

Two notions are very useful for understanding Calabi–Yau spaces: differential forms and vector bundles. Differential forms have already appeared implicitly in our discussion of IIA and IIB string theory. We start with an antisymmetric tensor field $A_{i_1 i_2 \ldots i_n}$. Suppose that there is a gauge invariance

$$A_{i_1 \ldots i_n} \to A_{i_1 \ldots i_n} + \frac{1}{n}\left[\partial_{i_1}\Lambda_{i_2 \ldots i_n} - \partial_{i_2}\Lambda_{i_1 i_3 \ldots i_n} + \cdots (-1)^r \partial_{i_r}\Lambda_{i_1 \ldots i_{r-1} i_{r+1} \ldots i_n}\right], \tag{26.1}$$

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

where Λ is antisymmetric in all its indices. We can write a shorthand for this,

$$\delta A = d\Lambda, \tag{26.2}$$

where $d\Lambda$ is the "exterior derivative." Acting on an antisymmetric tensor of rank p, the exterior derivative produces a rank-$(p+1)$ antisymmetric tensor, dH:

$$dH_{i_1 \ldots i_{p+1}} = \frac{1}{p+1} \left(\partial_{i_1} H_{i_2 \ldots i_{p+1}} - \partial_{i_2} H_{i_1 i_3 \ldots i_p + 1} + \cdots \right). \tag{26.3}$$

We can think of this object more abstractly as follows. Antisymmetric tensors with p indices are called p-forms. A "basis" for p-forms is provided by the antisymmetrized products of differentials:

$$dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}. \tag{26.4}$$

We can then write

$$H = \frac{1}{p!} H_{i_1 \ldots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p}. \tag{26.5}$$

The product of two forms A, B is known as the wedge product, $A \wedge B$. If A is an n-form and B an m-form then

$$(A \wedge B)_{i_1 \ldots i_{n+m}} = \frac{n!m!}{(n+m)!} A_{i_1 \ldots i_n} B_{i_{n+1} \ldots i_{n+m}} + \ (-1)^P \text{permutations} \tag{26.6}$$

or, more compactly,

$$A \wedge B = \frac{1}{(n+m)!} A_{i_1 \ldots i_n} B_{i_{n+1} \ldots i_{n+m}} dx^1 \wedge \cdots \wedge dx^{n+m}. \tag{26.7}$$

In this language the exterior derivative can be written as $d \wedge H$ or simply dH, where d is thought of as a one-form with components $d_i = \partial_i$.

It is important to practise with this notation, and some exercises are provided at the end of the chapter. One should check that

$$d^2 H = 0. \tag{26.8}$$

It is instructive to write electrodynamics in the language of forms. One should verify that the field strength tensor is a two-form, which can be written as

$$F = dA. \tag{26.9}$$

The homogeneous Maxwell's equations (the Bianchi identities for the field strength) follow from $d^2 = 0$:

$$dF = 0. \tag{26.10}$$

Apart from multiplication and differentiation, there is another important operation, denoted by $*$ and called the Hodge star. In d dimensions, this takes a p-form to a $(d-p)$-form:

$$(*H)_{i_1 \ldots i_{d-p}} = \frac{1}{p!} \epsilon_{i_1 \ldots i_{d-p}}{}^{i_{d-p+1} \ldots i_d} H_{i_{d-p+1} \ldots i_d}. \tag{26.11}$$

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

A particularly interesting object is $*d$. For example, $*d \wedge d$ is a d-form. But the components of a d-form are necessarily proportional to $\epsilon_{i_1 \dots i_d}$. With a little work, one can show that

$$*(*d \wedge d) = \partial^2. \tag{26.12}$$

Using the $*$ operation, we can write the action for a p-form field as

$$S = \frac{1}{2(p+1)!} \int *F \wedge F \tag{26.13}$$

with $F = dA$. This is clearly gauge invariant. It is easy to check that this reproduces the standard action for electrodynamics.

For physics, we are particularly interested in the zero modes of A, i.e. field configurations that satisfy $dA = 0$ but which are not simply gauge transformations; they cannot everywhere be written as

$$A = d\Lambda. \tag{26.14}$$

A simple example of what is at issue is provided by a gauge field on a circle, $0 \leq y \leq 2\pi R$. The one-form gauge field,

$$A_y = \partial_y \Lambda, \quad \Lambda = cy \tag{26.15}$$

is not a sensible gauge transformation unless $c = n/R$, since a fermion of unit charge will not transform into itself. In electrodynamics, for example, this corresponds to the fact that the Wilson line,

$$U = \exp\left(i \int_0^{2\pi R} dy\, A_y\right) \tag{26.16}$$

is gauge invariant and non-trivial, again, unless $c = n/R$.

This suggests that we want to consider closed p-forms α which satisfy

$$d\alpha = 0, \tag{26.17}$$

but that we are not interested in *exact* forms

$$\alpha = d\beta. \tag{26.18}$$

More generally, we want to define an equivalence class known as the cohomology class of α. We will view α and α' as equivalent if

$$\alpha' = \alpha + d\beta, \tag{26.19}$$

where β is well defined everywhere on the manifold.

In general, for field configurations on a manifold M the number of linearly independent zero modes is known as the *Betti number*, b_p. This number is related to the number of (basis) p-dimensional submanifolds which are not boundaries of $(p+1)$-dimensional surfaces. We will not prove this but will at least make it plausible. Consider the integration of a p-form, α, over a p-dimensional submanifold Σ:

$$\int_\Sigma \alpha_{i_1 \dots i_p}\, d\Sigma^{i_1 i_p}. \tag{26.20}$$

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

By Stokes' theorem, the integral of the exterior derivative of a $(p-1)$-form β over Σ is related to the integral of β over the boundary of Σ:

$$\int_{\Sigma} d\beta = \int_{\partial\Sigma} \beta. \tag{26.21}$$

If Σ is compact, it has no boundary so the integral of $d\beta = 0$.

Two p-forms are in the same cohomology class if

$$\int_{\Sigma} (\alpha - \alpha') = \int_{\Sigma} d\beta = \int_{\partial\Sigma} \beta = 0. \tag{26.22}$$

Note that, as before, it is important in this expression that β is defined throughout the manifold.

If we consider the structure of a massless chiral multiplet, we note that there are two scalars and a chiral fermion. In compactifications preserving $N = 1$ supersymmetry, modes of antisymmetric tensor fields which are annihilated by d will correspond to massless scalars; supersymmetry guarantees that the other elements of the multiplet are also present. The suggested readings at the end of the chapter contain more detailed discussions of these issues, but it is not too hard to understand how the various states arise in terms of the forms annihilated by d. The other massless *scalar* arises because one can also choose the form in such a way the Laplacian vanishes. The Dirac operator is closely related to differential forms on manifolds. This can be shown using the creation–annihilation operator construction of the Dirac matrices that we used in our discussion of orthogonal groups. One can exhibit in this way the required pairing.

With this machinery we can define an important set of topological invariants of manifolds: characteristic classes. Consider a gauge field F, where $F = dA$. Note that F is closed: $dF = 0$. The gauge field F is said to be an element of $H_1(M, R)$, the second cohomology group of the manifold M with real coefficients. The cohomology class of such two-forms is known as the first Chern class.

When the manifold is topologically non-trivial, if we consider a gauge field then it may not be possible to describe the field everywhere by a single non-singular potential. This problem is familiar to us from the case of the Dirac monopole. Instead, in different regions α and β we have to use different potentials, $A_{(\alpha)}$, $A_{(\beta)}$. In regions where α and β overlap (transition regions), $A_{(\alpha)}$ and $A_{(\beta)}$ will be gauge transforms of one another:

$$A_{(\alpha)} = A_{(\beta)} + \phi_{(\alpha\beta)}. \tag{26.23}$$

Another set of gauge fields is said to be in the same topological class if

$$\tilde{A}_{(\alpha)} = \tilde{A}_{(\beta)} + \phi_{(\alpha\beta)} \tag{26.24}$$

with the *same* transition function ϕ. Now, since the functions A and $\tilde{A}$ are not uniquely defined everywhere, on the one hand $F = dA$ and $\tilde{F} = d\tilde{A}$ are not in the trivial cohomology class in general. On the other hand, $F - \tilde{F}$ is in this class, since the difference $A - \tilde{A} = B$ is well defined. So $F - \tilde{F} = dB$ and F and $\tilde{F}$ are in the same cohomology class. Thus the cohomology class of F, the first Chern class, is a topological invariant.

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

There is a theorem which states that if the first Chern class is non-zero then one can always find a two-dimensional surface Σ with the property

$$I(\Sigma) = \frac{1}{2\pi} \int_{\Sigma} F \neq 0. \tag{26.25}$$

Note that this is a kind of magnetic flux. By Dirac's argument (see Chapter 7), $I(\Sigma)$ is an integer. The first Chern class plays an important role in the theory of Calabi–Yau spaces.

These ideas can be generalized to complex spaces. Here we define, as we did for the orbifold, complex coordinates z_i and $\bar{z}_i$. We then define a (p, q)-form ψ to be an object with p z_i-type indices and q $\bar{z}_i$-type indices. Note that ψ is totally antisymmetric in both types of indices. We can define two types of exterior derivatives, ∂ and $\bar{\partial}$, in an obvious way:

$$\partial \psi_{a_1 \dots a_{p+1}, \bar{a}_1 \dots \bar{a}_q} = \frac{1}{p+1} \partial_{a_1} \psi_{a_2 \dots a_{p+2} \bar{a}_1 \dots \bar{a}_q} + (-1)^P \text{ permutations.} \tag{26.26}$$

Note that $\partial^2 = 0$; $\bar{\partial}$ is defined similarly. In terms of these definitions,

$$d = \partial + \bar{\partial}. \tag{26.27}$$

These are known as the Dolbeault operators. We can then consider differential forms annihilated by these operators. The numbers of independent forms annihilated by the ∂ and $\bar{\partial}$ operators are known as the Hodge numbers, $h^{p,q}$. Then, for example, one has the Hodge decomposition

$$b_n = \sum_{p+q=n} h^{p,q}. \tag{26.28}$$

Again, is is possible to choose these forms so that they are annihilated by the Laplacian.

26.2 Calabi–Yau spaces: constructions

We have already constructed a rather rich set of four-dimensional string theories. But they are only a small subset of what appears to be a vast set of possibilities. We saw, for example, that the orbifold compactifications give rise to moduli which describe states which are *not* orbifolds. A rich set of compactifications of string theory, of which the orbifolds we studied in the last chapter are special cases, are provided by the Calabi–Yau spaces. In this section, we introduce these.

Our strategy to construct solutions is to look for solutions of the ten-dimensional field equations. One can ask: why is this sensible? There are two answers. First, if we consider spaces in which the massless ten-dimensional fields are slowly varying, it should be appropriate to integrate out the massive string modes and study the low-energy equations. A more serious question is: why is it that we can simply look at the low-order equations? Even at the classical level, integrating out the massive states will lead to terms with arbitrary numbers of derivatives. This question is far more serious. If we solve the equations, say, involving two derivatives then we can try to find solutions of the terms in up to four derivatives perturbatively. To do this we expand the fields in modes of the

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

lowest-order theory (e.g. eigenfunctions of the Laplace operator on the complex space). These are precisely the Kaluza–Klein modes. Calling these ϕ_n and substituting our lowest-order solution into the next-order terms, we will obtain equations of the form

$$\left(\nabla^2 + m_n^2\right)\phi_n = \frac{\alpha'}{R^2}\Gamma_n. \tag{26.29}$$

For $m_n \neq 0$, i.e. for the massive Kaluza–Klein modes, we simply obtain a small shift. But the massless modes are problematic. In the case of Calabi–Yau compactifications it is supersymmetry which will come to our rescue. We will see that, for the massless modes, the tadpoles (Γ_ns) vanish.

We begin with the Type II theory. Rather than examine the equations of motion we look at the supersymmetry variations. In flat-space four-dimensional theories, we are familiar with the idea that we can find minima of the potential by setting the auxiliary fields to zero. We can phrase this in a different, seemingly more obscure, way: we can find static solutions of the classical equations by requiring that the supersymmetry variations of all the fields vanish. That is, we require

$$\delta\psi = \epsilon F = 0, \quad \delta\lambda = \epsilon D = 0. \tag{26.30}$$

We will try the same strategy. In Chapter 17 we introduced the essential elements required to understand spinors in a gravitational background (the reader may want to reread Section 17.6). To make things simple, we will look for solutions where the antisymmetric tensor vanishes and the dilaton is constant, so only the metric is spatially varying. Then the condition that there should be a conserved supersymmetry becomes

$$\delta\psi_M = D_M\eta = 0. \tag{26.31}$$

So η is covariantly constant. This means that, under parallel transport around any closed curve, η returns to itself. As in gauge theories the effect of parallel transport can be described in terms of Wilson lines, where now the Wilson line is written in terms of the spin connection, ω:

$$U = P\exp\left(i\oint \omega\, dx\right). \tag{26.32}$$

The fact that η is unchanged under any such transformation greatly restricts the form of ω. To see how this works, consider that in the ten-dimensional Lorentz group, there is an $O(6)$ subgroup which acts on the compactified coordinates, as well as the four-dimensional Lorentz group acting on the Minkowski coordinates. The 16-component spinor in ten dimensions decomposes under these groups as

$$\eta = (4, 2) + (\bar{4}, 2^*). \tag{26.33}$$

By local Lorentz transformations, we can take the $(4, 2)$ representation to have the form (suppressing the four-dimensional spinor index)

$$\eta = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \eta_0 \end{pmatrix}. \tag{26.34}$$

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

In order that this be invariant, we require that the spin connection lie in an $SU(3)$ subgroup of $O(6)$. The space is said to be a space of $SU(3)$ holonomy.

In general ω is an $O(6)$ matrix. Restriction to $SU(3)$ is a strong constraint. Already $U(3)$ holonomy requires that the manifold be complex. We encountered this in the orbifold case, where we introduced three complex coordinates and their conjugates. There is no unique way to introduce the complex coordinates. The continuous set of choices will lead to a set of moduli of our solutions, known as the complex structure moduli. In addition, a manifold of $U(3)$ holonomy is Kahler. This means that the metric can be derived from a function $K(x^i, x^{\bar{i}})$, the Kahler potential, through

$$g_{i\bar{j}} = \partial_i \partial_{\bar{j}} K. \tag{26.35}$$

While proving that a manifold of $U(3)$ holonomy must be Kahler is challenging, it is not hard to check that a Kahler manifold has $U(3)$ holonomy. Some aspects of these manifolds are discussed in the exercises.

The Christoffel symbols (affine connection) and curvature for a Kahler manifold can be written in quite compact forms. (Verification of these formulas is left for the exercises.) The components of the Christoffel symbols are given by

$$\Gamma^a_{bc} = g^{a\bar{d}} \partial_b g_{c\bar{d}}, \quad \Gamma^{\bar{a}}_{\bar{b}\bar{c}} = g^{\bar{a}d} \partial_{\bar{b}} \partial \bar{b} g_{\bar{c}d}. \tag{26.36}$$

As a result, the non-zero components of the Riemann tensor are

$$R^{\bar{a}}_{\bar{b}c\bar{d}} = \partial_c \Gamma^{\bar{a}}_{\bar{b}\bar{d}} \tag{26.37}$$

and the Ricci tensor is

$$R_{\bar{b}c} = -\partial_c \Gamma^{\bar{a}}_{\bar{b}\bar{a}}. \tag{26.38}$$

Using

$$\Gamma^{\bar{a}}_{\bar{b}\bar{a}} = \partial_{\bar{b}} \ln \det g, \tag{26.39}$$

this can be further simplified:

$$R_{\bar{b}c} = -\partial_{\bar{b}} \partial_c \ln \det g. \tag{26.40}$$

Note that our result, Eq. (24.19), for the curvature of a two-dimensional Riemann surface is a special case of this.

The requirement that the metric have $SU(3)$ holonomy has a dramatic consequence for the curvature: the Ricci tensor vanishes. This follows from our discussion of the spin connection as a gauge field for local Lorentz transformations. On a six(real)-dimensional Kahler manifold we have seen that the spin connection is not an $O(6)$ field but, rather, a $U(3)$ field (in four dimensions it is a $U(2)$ field, etc.). The $U(1)$ part of the Riemann tensor is the trace over the Lorentz indices – the group indices, thinking of the Riemann tensor as a non-Abelian field strength. But this object is the Ricci tensor, so $SU(3)$ holonomy requires that the Ricci tensor itself vanish everywhere on the manifold. For such a configuration the lowest-order Einstein equation is automatically satisfied, $R_{i\bar{j}} = 0$. The question which we would like to address is: given a Kahler manifold, is it possible to deform the Kahler

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

potential in such a way that the Ricci tensor vanishes? Clearly a necessary condition for this is that the integral

$$c_1 = \frac{1}{2\pi} \int \mathrm{Tr}\, R \tag{26.41}$$

vanish. This quantity is the first Chern class, the topological invariant which we discussed earlier. It was Calabi who conjectured that the vanishing of the first Chern class for a manifold was a necessary and sufficient condition that the manifold admit a unique metric of $SU(3)$ holonomy. Yau later proved this conjecture. The spaces constructed in this way are the famous Calabi–Yau spaces. In general, while one can prove that such metrics exist, actually constructing them is a difficult numerical problem. Fortunately, many properties relevant to the low-energy behavior of string theory on these manifolds can be obtained from more limited, topological, information.

It is worthwhile comparing this with our orbifold constructions. The orbifolds are everywhere flat. But the existence of a deficit angle associated with the fixed points means that there is actually a δ-function curvature; this gives precisely the holonomy of these manifolds. If we decompose the spinors as before then, as we transport them about the fixed points, the i-components pick up a phase, $e^{\frac{2\pi i}{3}}$, while the 0-components are invariant. Correspondingly, we find one unbroken supersymmetry.

When we discuss the heterotic theory on a Calabi–Yau space, we will have to choose values for the gauge fields as well. It will not be possible to simply set the gauge fields to zero. From the point of view of four dimensions, gauge fields with indices in the extra dimensions are like scalars, so this will result in the breaking of some or all the gauge symmetry. As we will see in Section 26.6.1, there are many possible choices for these fields, with distinct consequences for the structure of the low-energy theory. In an interesting subclass, some features of the heterotic theory are closely related to those of Type II on Calabi–Yau spaces.

26.3 The spectrum of Calabi–Yau compactifications

In both the Type II and heterotic cases, many features of the low-energy spectrum follow from general topological features of the manifold and do not depend on details of the metric. In the heterotic case the number of generations (minus the number of antigenerations) is a topological invariant. Suppose that we have some number of generations for some choice of metric. If we now make smooth, continuous, changes in the metric then the massless spectrum can change, as generations and antigenerations pair to gain mass or become massless. In other words, a mass term in an effective action can pass through zero but the net number of generations cannot change. In some cases, other features of the spectrum are similarly invariant. So, while it is difficult to write down explicit metrics for manifolds having $SU(3)$ holonomy, it is possible to determine many important features of the low-energy theory from basic topological features of the manifold.

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

In the Type II theory the numbers of hypermultiplets and vector multiplets are separately topological. They do not pair up as one moves about on the moduli space; the $N=2$ supersymmetry ensures that if a field is massless at one point in the moduli space then it is massless at all points. Even more dramatic is that the massless states found in the lowest order of the α' expansion are in fact massless to all orders α' and in string perturbation theory. So it is enough to study the lowest-order supergravity equations of motion in order to count the massless particles.

The important non-zero Hodge numbers are $h^{2,1}$ and $h^{1,1}$. In the IIA theory there are $h^{1,1}$ vector multiplets and $h^{2,1}$ hypermultiplets. In the IIB theory this is reversed. In the heterotic case, the $(2,1)$-forms will correspond effectively to generations and the $(1,1)$-forms to antigenerations.

The counting of massless fields is not difficult to understand. Since we have taken the antisymmetric tensor fields and fermions to vanish in the background, the equations for these fields are particularly simple. Consider the antisymmetric tensor $B_{\mu\nu}$. On a complex manifold, as we explained earlier, there are $h^{1,1}$ $(1,1)$-forms $b^{(a)}_{i,\bar{j}}$ and $h^{2,1}$ $(2,1)$-forms annihilated by the operators ∂ and $\bar{\partial}$. Since the corresponding three-index field strengths $H = dB$ vanish, there is no energy cost to giving a constant expectation value to the associated four-dimensional fields; they correspond to massless scalars in four dimensions. The fields connected to the $(1,1)$-forms $b_{i,\bar{i}}$, are easy to describe. In addition to the antisymmetric tensor there is also a massless perturbation of the metric:

$$ig_{\bar{i}i}(x,y) = \phi(x)b_{i,\bar{i}}(y). \tag{26.42}$$

Here x refers to the ordinary four-dimensional Minkowski coordinates and y refers to the compactified coordinates. Similarly, in the IIA theory one can find a massless gauge field rounding out the bosonic components of the vector multiplet. This comes from the three-index Ramond field,

$$C_{\mu i,\bar{i}}(x,y) = A_\mu(x)b_{i,\bar{i}}(y). \tag{26.43}$$

We will leave to the reader the problem of working out the structure of the hypermultiplets in terms of the (2, 1)-forms and also of determining the pairings in the IIB case.

A $(1,1)$-form which is always present is the Kahler form,

$$b^K_{i,\bar{i}} = ig_{i,\bar{i}}, \quad b_{\bar{i},i} = -ig_{i,\bar{i}}. \tag{26.44}$$

This satisfies

$$\partial b^K = \bar{\partial} b^K = 0 \tag{26.45}$$

because $g_{i\bar{i}} = \partial_i \partial_{\bar{i}} K$. The real scalar which sits in the multiplet with b^K is just the metric itself. The corresponding massless field is the radius of the compact space:

$$g_{i,\bar{i}}(x^\mu, z^i) = R^2(x^\mu)g_{i,\bar{i}}(z), \quad B_{i,\bar{i}}(x^\mu, z^i) = b(x^\mu)b_{i,\bar{i}}(z). \tag{26.46}$$

That the field is massless is no surprise; the condition $R_{i\bar{i}} = 0$ is not changed under an overall rescaling of the metric, so the vev is undetermined.

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

26.4 World-sheet description of Calabi–Yau compactification

Thus far we have described the compactification of string theory in terms of ten-dimensional space–time. This analysis makes sense if the radius of the compactified space is large compared with the string length, ℓ_s. We can also formulate these questions in world-sheet terms. This provides a complementary way to understand many features of the compactified theory and is useful for at least two reasons. First, it provides tools to ask what happens when the compactification radius is of order the string scale or smaller. Second, there are some features of the spectrum and interactions which are more readily accessible in this framework.

In the Type II theory the non-linear sigma model which describes compactification on a Calabi–Yau space has some striking features. First, in the absence of background antisymmetric tensor fields it is left–right symmetric. Second, there are two left-moving and two right-moving supersymmetries on the world sheet as opposed to the one left-moving and one right-moving supersymmetry of a general configuration. This can be usefully understood in a number of ways. In the light cone gauge, one can work with the covariantly constant spinor η and its conjugate $\bar{\eta}$ to construct two left-moving and two right-moving supersymmetry generators, both in the sense of the world sheet and in space–time. We have already seen this in the case of orbifold constructions. There, in the light cone gauge, we have eight left-moving and eight right-moving supersymmetry generators, before the orbifold projection. We can organize these in terms of their transformation properties under the $SU(3) \times U(1)$ holonomy group. For both the left and right movers there are triplets Q_i, antitriplets $\bar{Q}_{\bar{i}}$ and singlets, Q_0 and $\bar{Q}_0$. The triplets and antitriplets are charged under the $U(1)$ symmetry; the singlets are not. The orbifold projection eliminates the triplets. The two singlets survive.

In a purely world-sheet description, non-linear sigma models described by a Kahler metric automatically have two left-moving and two right-moving supersymmetries. To describe these, we can introduce a superspace with four Grassmann coordinates, of which two are left movers and two are right movers: θ_+^A and θ_-^A. This superspace can be thought of as the truncation of $N = 1$ supersymmetry in four dimensions. As in four dimensions we can define, operators D_α and $\bar{D}_\alpha$ and left- and right-moving chiral fields annihilated by the $\bar{D}$s. Correspondingly, we can define chiral left- and right-moving fields

$$X_+^i(z,\theta) = x^i(z) + \theta_+^A \psi_A^i(z) + \text{auxiliary field} \tag{26.47}$$

and similarly for X_-^i. In terms of these fields we can write the action of the conformal field theory as

$$\int d^2\sigma \int d^2\theta_+ d^2\theta_- K(X, \bar{X}). \tag{26.48}$$

Integrating over the θs, the bosonic terms are just $\int d^2\sigma g_{i,\bar{i}}\, \partial_\alpha x^i \partial^\alpha x^{\bar{i}}$, with $g_{i,\bar{i}}$ the Kahler metric.

The superconformal algebra, in these backgrounds, is enlarged to what is referred to as the $N = 2$ superconformal algebra (one such algebra for the left movers, one for the right

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

movers). In addition to the stress tensor and the two supercurrents, this algebra contains a $U(1)$ current. The supersymmetry generators can be constructed by the Noether procedure. They can also be guessed by taking the generators in a flat background and making the expressions covariant:

$$G^+ = g_{i,\bar{i}} DX^i \psi^{\bar{i}}, \quad G^- = g_{i,\bar{i}} DX^{\bar{i}} \psi^i. \tag{26.49}$$

These have opposite charge under the $U(1)$ current (an R current) constructed from the fermions,

$$j(z) = \psi^{\bar{i}}(z)\psi^i(z), \tag{26.50}$$

with a similar current for the left movers. The full algebra is

$$\begin{aligned} T(z)G^\pm(0) &\approx \frac{3}{2z^2}G^\pm(0) + \frac{1}{z}\partial G^\pm(0), \\ T(z)j(0) &\approx \frac{1}{z^2}j(0) + \frac{1}{z}\partial j(0), \\ j(z)G^\pm(0) &\approx \pm\frac{1}{z}G^\pm(0). \end{aligned} \tag{26.51}$$

These equations say that G has dimension $3/2$ while j has dimension one, and $G^\pm$ have $U(1)$ charges plus and minus one. The central charge appears in the relations

$$\begin{aligned} G^+(z)G^-(0) &\approx \frac{2c}{3z^3} + \frac{2}{z^2}j(0) + \frac{2}{b}T(0) + \frac{1}{z}\partial j(0), \\ G^+(z)G^+(0) &\approx 0, \\ j(z)j(0) &\approx \frac{c}{3z^2}. \end{aligned} \tag{26.52}$$

The non-linear sigma models appropriate to *heterotic* compactifications on Calabi–Yau spaces have a number of interesting features. We will see that, for a particular choice of gauge fields, the world-sheet theory which describes the heterotic compactification is identical to that of the Type II theory. Thus again they have two left-moving and two right-moving supersymmetries ((2, 2) supersymmetry). The fact that the world-sheet theories of the two different string theories are the same allows us to argue, as we will below, that Calabi–Yau spaces are solutions of the full, non-perturbative, string equations of motion. But this observation also tells us about interesting features of the spectrum.

To understand the spectrum, it is helpful to ask, first, what is a vertex operator from the perspective of the two-dimensional conformal field theory? The answer is that a vertex operator is a *marginal deformation* of the theory, a perturbation of dimension 2 ((1, 1) in terms of the left- and right-moving Virasoro algebras). The standard way to compute the dimensions of operators is to treat them as perturbations and calculate, for example, the beta function of the perturbation. For marginal operators the beta function vanishes to first order. The moduli correspond to "exactly marginal deformations" of the theory. For these the beta functions vanish to all orders in the perturbation (and non-perturbatively), corresponding to the fact that the theory, even for a finite perturbation, is conformal.

The existence of moduli means that there is a multiparameter set of conformal field theories. Varying the action with respect to the parameters yields operators which are

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

exactly marginal. In this way, we have the two-dimensional version of the correspondence between moduli and massless fields.

An example of a modulus is the radius of the complex space. The lowest-order equation for the metric is invariant under an overall scaling of lengths. But this is not obviously true of the higher-order corrections. For Type II theories the space–time supersymmetry guarantees that there is no potential for the moduli, so the sigma model is a good conformal field theory, suitable for heterotic string compactification. On the heterotic side we can also give a more direct world-sheet argument. Here R^{-2} is the coupling constant of the sigma model. In other words, writing the metric as R^2 times a reference metric of order the string scale, R^2 appears in front of the Lagrangian. We know that the lowest-order beta function equation is the same as the field theory equation. It is trivially independent of R^2, since it is a one-loop effect. For higher orders there is a non-renormalization theorem. This follows from a combined world-sheet and space–time argument. The superpartner of fluctuations in the radius is the fluctuation of the antisymmetric tensor field, $b_{i,\bar{i}} = ig_{i,\bar{i}}$. The associated vertex operator term in the action is a total derivative on the world sheet at zero momentum. It is perhaps easiest to see this by writing the vertex operator at zero momentum in the form

$$\begin{aligned} V_b &= b_{MN}\epsilon^{\alpha\beta}\, \partial_\alpha X^M \partial_\beta X^N \\ &= \partial_M \partial_N K\, \epsilon^{\alpha\beta}\, \partial_\alpha X^M \partial_\beta X^N \\ &= \partial_\alpha (\epsilon^{\alpha\beta}\, \partial_\beta X^M \partial_M K). \end{aligned} \tag{26.53}$$

So b decouples at zero momentum. Because b is in a supermultiplet with R^2 this means that the superpotential, which is a holomorphic function of the superfields, is independent of R^2.

Actually, this statement is not precisely correct because K is not single-valued. In perturbation theory it is true since one is not sensitive to the global structure of the manifold (in perturbation theory, all fluctuations are small). Non-perturbatively, one can encounter instantons in the world-sheet theory. A more detailed analysis is required to determine whether there are corrections to the superpotential. In left–right symmetric compactifications of the heterotic string, i.e. those with two left-moving and two right-moving supersymmetries ((2, 2) models), a study of fermion zero modes in the presence of the instanton shows that no superpotential for the moduli is generated; this is consistent with one's expectations from the Type II theory. For compactifications with two right-moving but no left-moving supersymmetries ((2, 0) models), corrections can be generated though in some cases intricate cancelations still prevent the appearance of a potential for the moduli. These two classes of models are phenomenologically quite distinct, as we will see shortly.

26.5 An example: the quintic in CP^4

It is helpful to have a concrete example of a Kahler manifold with $c_1 = 0$, on which we know that one can construct a metric of $SU(3)$ holonomy. We have previously encountered

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

the complex projective spaces in N dimensions, CP^N. These are defined as spaces with $N+1$ complex coordinates Z_a and with the identification $Z_a \to \lambda Z_a$ for any complex number λ. We have written down a Kahler potential on this space:

$$K = \ln\left(1 + \sum_{a=1}^{N} Z_a \bar{Z}_a\right). \tag{26.54}$$

Any complex submanifold of a Kahler manifold is also a Kahler manifold; one can simply take the Kahler potential to be the Kahler potential of the full manifold, evaluated on the submanifold. To obtain a manifold with three complex dimensions we can start with CP^4 and write down an equation for the vanishing of a polynomial $P(Z)$. The polynomial should be homogeneous in order that it has a sensible action in CP^N. It turns out that it should also satisfy other conditions. Its gradient should at most vanish, at the origin (which is not a point in CP^N). In order that the first Chern class should vanish, it should be quintic. We will give an argument for this shortly.

The simplest (most symmetric) possibility is

$$P = Z_1^5 + Z_2^5 + Z_3^5 + Z_4^5 + Z_5^5 = 0, \tag{26.55}$$

but there are obviously many more. We can deform this polynomial by adding other quintic polynomials. These correspond to varying the complex structure. Since each deformation produces another solution of the string equations, each deformation corresponds to a modulus, one of the complex structure moduli. Associated with each deformation is a form of type $(2,1)$, which we will not attempt to construct here.

Before listing the deformations, we note that not every deformation corresponds to a change in the physical situation – and thus to a massless particle. Holomorphic changes of the coordinates which are non-singular and invertible do not change the complex structure. The transformation

$$Z_i \to Z_i + \epsilon^{ij} Z_j \tag{26.56}$$

is well defined in CP^4. As a consequence, deformations such as $Z_1^4 Z_2$ are not physical. So we can list the possible deformations:

$$Z_1^3 Z_2^2 \ldots, \quad Z_1^3 Z_2 Z_3 \ldots, \quad Z_1^2 Z_2^2 Z_3, \ldots, \quad Z_1^2 Z_2 Z_3 Z_4, \ldots, \quad Z_1 Z_2 Z_3 Z_4 Z_5. \tag{26.57}$$

All together there are 101 possible deformations of the polynomial, corresponding to $h_{2,1} = 101$. In this example, there is only one Kahler modulus, the overall radius of the compact space.

We can understand heuristically why the first Chern class vanishes, in a way which will help us to understand other features of these manifolds. A characteristic feature of the Calabi–Yau spaces is the existence of a covariantly constant three-form, ω_{ijk}. The existence of this form follows from the existence of a covariantly constant spinor η:

$$\omega_{ijk} = \bar{\eta}\Gamma_{[ijk]}\eta. \tag{26.58}$$

Working in terms of the creation–annihilation operator basis for the Γs, one sees that ω is holomorphic. The Γ_is can be defined in such a way that the $\Gamma_{\bar{i}}$ matrices annihilate η. Then,

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

because of the complete antisymmetrization, only components of ω with indices $1, 2, 3$ are non-vanishing. In the space defined by the vanishing of a quintic polynomial in CP^4, we can show that there exists a holomorphic three-form which is everywhere non-vanishing. Setting $x^i = Z_i/Z_5$, $i = 1, \ldots, 4$,

$$\omega = dx^1 \wedge dx^2 \wedge dx^3 \left(\frac{\partial P}{\partial x^4}\right)^{-1}. \tag{26.59}$$

One can show that this expression does not depend on singling out a particular coordinate and that it is not singular at the points where the derivative vanishes provided that the polynomial P is quintic and that the gradient of P vanishes only at the origin. The existence of such a form can be shown to be equivalent to the vanishing of the first Chern class.

26.6 Calabi–Yau compactification of the heterotic string at weak coupling

Much effort has been devoted to the study of compactifications of the weakly coupled heterotic string on Calabi–Yau spaces. These theories have many features of the Standard Model. They also allow one to consider many questions of Beyond the Standard Model physics. Before beginning an analysis of these models it is worth listing some points that we can address in this framework.

1. *Low-energy supersymmetry* Solutions of the classical equations of the heterotic string theory on Calabi–Yau spaces exist. They have $N = 1$ supersymmetry. Supersymmetry, as in field theory, is unbroken to all orders of perturbation theory but may be broken non-perturbatively.
2. *Low-energy gauge groups* The simplest constructions have gauge group $E_8 \times E_6$, broken perhaps by Wilson lines, which preserve the rank of the gauge group. But many models have a moduli space in which the gauge group is broken to precisely that of the Standard Model.
3. *Generations* The number of generations is typically determined in terms of topological features of the underlying manifold.
4. *Massless particles, not protected by symmetries* Various massless states arise which are not protected by chiral symmetries. This is precisely what we want in order to understand the presence of light Higgs fields in supersymmetric theories. We know that if such fields are present in the low-energy field theory, they are protected from gaining large masses by non-renormalization theorems. In field theory the vanishing of such mass terms appears mysterious; in these string constructions, it is automatic. Such states could play the role of Higgs fields in supersymmetric models. In other words, *the Huggs five-turning problem of ordinary supersymmetric field theories is readily solved in this framework.*
5. *Unification of couplings* The string theories that we are studying are not grand unified theories in the conventional sense. There is no energy scale at which these

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

compactifications appear as four-dimensional theories with a single unbroken gauge group. Yet, generically, the couplings are unified. These two features, which we will see are easy to understand in terms of the microscopic structure of string theory, are quite surprising from a low-energy point of view. They have sometimes been referred to as "string miracles".

6. *Continuous and discrete symmetries* It is easy to prove that for these compactifications (and for weak-coupling heterotic models in general) there are no continuous global symmetries; all continuous symmetries must be gauge symmetries. Discrete symmetries, however, hand, proliferate and might play the role of R-parity or lead to other interesting phenomena. These discrete symmetries are typically gauge symmetries, in the sense that they are residual symmetries left over after the breaking of continuous gauge symmetries.

We will also see that there are a number of problems with these models, which illustrate some of the basic difficulties in developing a string phenomenology, as follows.

1. *There are too many models* While there are many with three generations, there are also some with hundreds of generations, with non-standard gauge groups and the like.
2. *The problem of moduli* Non-perturbatively, moduli can acquire potentials but they typically vanish in various asymptotic regimes. Simple general arguments indicate that stable supersymmetry-breaking minima, if they exist, must be in regions which are inherently strongly coupled in the sense that *no* weak coupling approximation is available.
3. *The problem of the cosmological constant* This is closely related to the previous one. In many instances moduli potentials can be calculated. For any given value of the moduli the size of these potentials is scaled, as one would expect, by the scale of supersymmetry breaking. As a result, even if strongly coupled stable minima exist it is not clear why the cosmological constant should be small at these points.

We will not offer a solution to these problems in this chapter but will explore at least one proposed answer, known as the "landscape," in Chapter 30.

26.6.1 Features of Calabi–Yau compactifications of the heterotic string

In the previous section we asserted that, in suitable backgrounds, the world-sheet conformal field theory which describes the heterotic string is the same as that which describes the Type II theory. Here, we describe compactifications of the heterotic string theory in more detail.

To construct solutions, we still look for these which preserve a space–time supersymmetry. Again we require the supersymmetry variation of the gravitino to vanish, giving $D_\mu \eta = 0$, so once more we need a covariantly constant spinor. There is now an equation for the variation of the ten-dimensional gaugino, as well:

$$\delta\lambda \propto \Gamma^{ij} F_{ij} \eta. \tag{26.60}$$

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

One strategy, then, to find solutions which preserve $N = 1$ supersymmetry is to require that $F_{ij}\Gamma^{ij}$ is an $SU(3)$ matrix. There is a simple ansatz which achieves this. Both E_8 and $O(32)$ have $SU(3)$ subgroups:

$$SU(3) \times E_6 \times E_8 \subset E_8 \times E_8, \quad SU(3) \times O(26) \subset O(32). \tag{26.61}$$

On the Calabi–Yau space the spin connection is an $SU(3)$-valued field, so we take the gauge field to be a field in one of these $SU(3)$ subgroups. Then, for gauge generators not in $SU(3)$, expression (26.60) is automatically satisfied. For those in $SU(3)$ the condition is mathematically identical to that for the gravitinos and is again satisfied.

This ansatz satisfies another condition. We have set the antisymmetric tensor field B to zero but, because of the Chern–Simons terms, this does not by itself guarantee that the field strength H is zero. With this ansatz, however, the Chern–Simons terms for the gauge and gravitational fields are identical. As a quick check, note that

$$dH = \mathrm{Tr}\,(R \wedge R) - \mathrm{Tr}\,(F \wedge F), \tag{26.62}$$

and the two terms in this expression clearly cancel. This establishes that here we have a solution of the equations of motion to lowest order in the α' expansion. But there is another way to see this, which will allow us to establish, as we did for the Type II theory, that this is an *exact* solution, perturbatively and non-perturbatively. If we write down the non-linear sigma model which describes the heterotic string in this background, it is identical to that for the Type II theory. To see this, as in the orbifold case, we divide the left-moving gauge fermions into three sets. First, there are the fermions λ^A, $A = 1, \ldots, 16$, in the second E_8 group, which are not affected by the background gauge field and remain free.. In the first E_8, ten fermions, λ^a, $a = 1, \ldots, 10$ (transforming as a vector in the $O(10)$ subgroup of E_6), are also free. The remaining six interacting fermions can be grouped, like the left-moving coordinates, into three complex fermions, λ^i and $\lambda^{\bar{i}}$. These fermions interact in precisely the same way as the left-moving fermions in the Type II theory. This can be seen by writing the action of the Type II fermions in terms of the vierbein and spin structure rather than the metric and the Christoffel connection.

We see from this that the moduli of the Type II theory are also moduli of the heterotic theory. Actually, we knew this had to be so since we know that each of these conformal field theories, on the Type II side, is a good conformal field theory for the heterotic theory. But we can also see this pairing more directly in the language of vertex operators. Here it is somewhat more convenient to work in the RNS picture. The vertex operators correspond to small deformations of the background in the directions of the moduli. In the Type II theory they are built from right-moving fields, ∂X^i and ψ^i, and left-moving fields, $\bar{\partial} X^i$ and $\tilde{\psi}^i$. In the heterotic case we can trade $\tilde{\psi}^i$ with λ^i. Since the action for the λ^is is the same as that for the $\tilde{\psi}^i$s, the dimensions of the vertex operators are exactly the same. This does not preclude the existence of additional moduli on the heterotic side, and we will see that typically there *are* additional moduli in these compactifications.

While all moduli of the Type II theory are moduli of the heterotic theory, not all heterotic moduli correspond to states of the Type II theory. Vertex operators for moduli which preserve only two right-moving supersymmetries ($(2, 0)$) are not suitable vertex operators

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

for the Type II theory. The moduli we are considering here are distinguished because they preserve the two left-moving world-sheet supersymmetries, and we will refer to these as Type II moduli. Perhaps more interesting, though, than the pairing of moduli is a pairing of the Type II moduli with matter fields. The moduli associated with $(2,1)$-forms are paired with 27s of E_6 and $(1,1)$ moduli with $\overline{27}$s. This is most readily seen in the language of vertex operators, using the world-sheet superconformal symmetry. The vertex operators for the Type II theory are the highest components of the corresponding superconformal multiplets with respect to both left- and right-moving supersymmetries. In superspace they are the $\theta_+^2\theta_-^2$ components of operators of the form

$$f(X^i, \bar{X}^i). \tag{26.63}$$

The $\theta_+\theta_-^2$ component has dimension $(1/2, 1)$. We can form an operator of dimension $(1,1)$ by multiplying by λ^a, one of the free fermions. This operator does not have the highest weight with respect to the left-moving $N = 2$ algebra, but this is not a problem; this symmetry is not a gauge symmetry on the world sheet but simply an accident of our choice of background field. It is highest-weight with respect to the left-moving Virasoro algebra, which is all that matters.

We have already observed this pairing in the Z_3 orbifold model, which is a special case of the Calabi–Yau construction. In the untwisted sector the vertex operators for the moduli took the form, for the left-movers,

$$\bar{\partial}X^i, \tag{26.64}$$

while for the 27s they took the form

$$\lambda^a\lambda^i. \tag{26.65}$$

The supersymmetry transformation of the latter operator changes λ^i to $\bar{\partial}X^i$.

The distinction between 27s and $\overline{27}$s is readily understood. In the Type II case we can distinguish two types of moduli, depending on their charges under the $U(1)$ symmetry within the left-moving $N = 2$ algebra. In the orbifold context some vertex operators involve $\bar{\partial}X^i$ and some $\bar{\partial}X^{\bar{i}}$. In the heterotic case, the world-sheet $U(1)$ symmetry corresponds to the $U(1)$ subgroup of E_6 in the decomposition $O(10) \times U(1) \subset E_6$. This $U(1)$ charge is precisely what distinguishes the 10s, for example, in the 27 and $\overline{27}$. In the Type II case this distinction corresponds to the distinction between $(2,1)$ and $(1,1)$ moduli, so we obtain precisely the pairing we described above (note that what one calls a 27 and a $\overline{27}$ is a matter of convention; if one adopts an opposite convention, the identification is reversed).

This result holds everywhere in the moduli space; since the number of moduli of each type does not change as one moves in the moduli space, the number of 27s and $\overline{27}$s does not change. This is a surprising result. One might have thought that, in a complicated construction such as this, 27s and $\overline{27}$s would, whenever possible, pair to gain mass. But this is not the case. This is precisely the sort of phenomena one needs to understand light Higgs particles in supersymmetric theories. We will see shortly how this works in a more detailed model.

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

26.6.2 Gauge groups: symmetry breaking

The heterotic models we have been considering have group $E_8 \times E_6$. If we are to describe the Standard Model we need to be able to break this symmetry. We have seen in the case of toroidal compactifications that gauge symmetries can be broken by the expectation values of gauge fields with indices in compactified dimensions. Stated in a more gauge-invariant fashion, these are non-trivial expectation values for Wilson lines. In the Calabi–Yau case the same is possible.

We will consider a specific example: the quintic in CP^4, with the vanishing of the polynomial:

$$Z_1^5 + Z_2^5 + Z_3^5 + Z_4^5 + Z_5^5 = 0. \tag{26.66}$$

The corresponding Calabi–Yau manifold, as we saw, has 101 27s and one $\overline{27}$. This polynomial has a variety of symmetries. As in the case of the torus, we can use these to project out states and simplify the spectrum. Consider, for example, the symmetry

$$Z_i \to \alpha^i Z_i, \quad \alpha = e^{2\pi i/5}. \tag{26.67}$$

This is a symmetry of the polynomial. It is somewhat different from the orbifold symmetries we have discussed since, as the reader can check, it *acts without fixed points.* Mathematicians refer to such a symmetry as “freely acting”. For the physics it means that if we mod out the Calabi–Yau by this symmetry then, while it is still necessary to include twisted sectors, the twisted strings have mass of order R, the Calabi–Yau radius, and there are no light states in these sectors if R is large.

We can readily classify the states that are invariant under this symmetry. Among the moduli, there are 21 $h_{2,1}$ fields, associated with polynomials such as $Z_1^3 Z_3 Z_4$ and $Z_1 Z_2 Z_3 Z_4 Z_5$. The Kahler modulus (i.e. the overall radius) is also invariant under this transformation, and so survives the projection. The corresponding Euler number is 40, one fifth of the Euler number of the covering space. There are also 21 27s of E_6 and one $\overline{27}$. Further symmetries can be used to reduce the number of generations to as few as four.

But what interests us here is obtaining smaller gauge groups. We can define the Z_5 group to include a transformation in E_6. This is equivalent to the presence of a Wilson line on the manifold. An interesting way to do this is to consider a somewhat different decomposition of E_6 from what we have considered up to now:

$$SU(3) \times SU(3) \times SU(3) \subset E_6. \tag{26.68}$$

An example of a Wilson line in this product of $SU(3)$s is

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha^3 \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha^3 \end{pmatrix}. \tag{26.69}$$

This breaks E_6 to $SU(3) \times SU(2) \times SU(2) \times U(1)^2$.

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

26.6.3 Massless Higgs fields, or the μ problem

When we mod out in such a way as to reduce the gauge symmetry, we also alter the spectrum. We have seen that this greatly reduces the number of moduli and the number of generations. The presence of the Wilson lines also disrupts the left–right symmetry of the model. As a result, the pairing of moduli and matter fields is no longer quite so simple.

In the presence of the Wilson line one still obtains 20 complete E_6 generations. If one thinks, loosely, of some of the massless fields "gaining" mass then elements of the 27 and $\overline{27}$s must pair up to gain mass. More precisely, in this modding out procedure states disappear, but they must disappear in pairs. But one also obtains some incomplete multiplets, where paired states do not disappear. Consider the $\overline{27}$. This is invariant under the original Z_5s, so any state which survives must be invariant under the Wilson line. Using the decomposition of the 27 under $SU(3)^3$, we obtain

$$27 = (3, 1, \bar{3}) + (\bar{3}, 3, 1) + (1, \bar{3}, 3). \tag{26.70}$$

So we obtain Z_5 singlets from only the third multiplet. These form a $(1, 2, 2)$ under $SU(3) \times SU(2) \times SU(2)$, as well as a singlet. There is a corresponding pair of states from the 27s. This is the sort of multiplet we need to help understand the presence of light Higgs particles in supersymmetric models: massless states at tree level which arise, from a low-energy point of view, more or less by accident.

26.6.4 Continuous global symmetries

In the heterotic string theory, there are no continuous global symmetries. We will not give a formal proof here but the basic argument is not hard to understand. If there is a global symmetry, it should be a symmetry of the world-sheet theory. In this way we are guaranteed that vertex operators can be chosen to have well-defined transformation properties and that the S-matrix will transform properly. The global symmetry will be associated with a world-sheet current. This current can be decomposed into left- and right-moving pieces. But, from any left-moving current we can build a gauge boson vertex operator, so the symmetry is necessarily a gauge symmetry. Right-moving currents will not commute with the world sheet supersymmetry generators and will not have a well-defined action on states (in BRST language they do not commute with the BRST operator). So they are not symmetries in space–time.

There are subtleties needed to complete the proof. First, as we have already seen, string theories typically possess, in perturbation theory, symmetries under which a scalar field undergoes a constant shift. These symmetries, as we will discuss further, are only broken non-perturbatively. The space–time version of such symmetries is not a conventional selection rule but rather a statement that scattering amplitudes vanish in the limit that the momenta of certain particles tend to zero. Second are the selection rules associated with the Poincare group. These clearly have a different status. On the one hand, in some sense, these symmetries are connected to the gauge symmetries of general relativity. On the other hand, their world-sheet implementation is different. For example, translations would appear to be non-linearly realized symmetries from a world-sheet point of view, but

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

momentum is still conserved as a consequence of the Mermin–Wagner–Coleman theorem. In any case, these subtleties are readily isolated and resolved.

This argument also indicates that there are no global symmetries in the Type II theories. This is in accord with our expectation that global symmetries are unlikely to arise in a theory of quantum gravity.

26.6.5 Discrete symmetries

When we studied orbifold models, we found that discrete symmetries existed in a subset of vacua on the full moduli space. This is also the case for the Calabi–Yau manifold constructed from the vanishing of a quintic polynomial in CP^4. Such symmetries turn out to be quite common.

The quintic polynomial $P = \sum Z_i^5$ exhibits a set of Z_5 symmetries:

$$Z_i \to \alpha^{k_i} Z_i, \quad \alpha = e^{2\pi i/5}. \tag{26.71}$$

An overall phase rotation of all the Z_is has no effect in CP^4, so the symmetry here is Z_5^4. There is also a permutation symmetry, S_5. This symmetry group is a subgroup of the $O(6)$ symmetry which would act on six non-compact flat dimensions. We can thus think of these symmetries as discrete gauge transformations. So their existence in a theory of gravity is not surprising.

We would like to know whether these symmetries are R symmetries. We can address this by considering their effect on the covariantly constant spinor η. This is more challenging to do than in the orbifold context, since we do not have quite such explicit expressions. It is simplest to look at the covariantly constant three-form. We have already given a construction,

$$\omega = dx^1 \wedge dx^2 \wedge dx^3 \left(\frac{\partial P}{\partial x^4} \right)^{-1}, \tag{26.72}$$

with $x^i = Z_i/Z_5$. This construction treats the coordinates asymmetrically but, as we explained, ω is symmetric among the coordinates. Note that ω transforms essentially like η^2, i.e. like θ^2. So symmetries under which ω transforms non-trivially are R symmetries, and W transforms like ω.

Consider first the Z_5 transformations of the separate Z_is. We can read off immediately how ω transforms under transformations of the first three; the other two follow by symmetry. So, we have

$$\omega \to \alpha^{\sum k_i}. \tag{26.73}$$

Similarly, under those S_5 transformations which permute Z_1, Z_2, Z_3 we can see how ω transforms. If the permutation is odd, ω changes sign. So again the general S_5 transformation is an R symmetry.

Turning on the various complex structure moduli typically breaks some of or all this symmetry. For example, if we turn on the modulus associated with the polynomial

$$z_1 z_2 z_3 z_4 z_5 \tag{26.74}$$

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

then we break the Z_5^4 symmetry down to a subgroup satisfying $\sum k_i = 0 \bmod 5$. This group is Z_5^3 but it is a non-R-symmetry, in light of the transformation law of ω. An expectation value for this field clearly preserves the permutation symmetry.

Similarly, turning on say $aZ_1^3 Z_2 + bZ_1^2 Z_2^3$ breaks the symmetries acting on Z_1 and Z_2 as well as some of the permutation symmetry. Turning on enough fields breaks all the symmetry.

One might ask why one should be interested in points or surfaces in the moduli space which preserve a discrete symmetry, when in the bulk of the space there is no symmetry. This question is closely related to the question: what sorts of dynamics might determine the values of the moduli? This is a subject with which we will deal extensively later but for which we can provide no definitive resolution. But, even without understanding this dynamics, there is a simple reason to suspect that points in the moduli space with symmetries might be singled out by the dynamics. Imagine that we somehow manage to compute an effective potential for the moduli, arising, perhaps, due to some non-perturbative string effects. Symmetry points are necessarily *stationary points* of this effective potential. There is, of course, no guarantee that they are *minima* of the potential but they are certainly of interest as candidates for string ground states.

There are, as we have seen, certain facts of nature which suggest that discrete symmetries might play some role in extensions of the Standard Model, including proton decay and dark matter.

26.6.6 Further symmetry breaking: the Standard Model gauge group

The Wilson line mechanism, as we have described it, provides a path to reduce the gauge symmetry from $E_6 \times E_8$ but leaves the rank untouched.[1] We can hope to reduce the gauge symmetry further by giving expectation values to some matter fields. Ideally, these expectation values would be large. The presence of other gauge groups (as well as unwanted matter multiplets) can spoil the prediction of coupling unification and can lead to severe difficulties with proton decay and other rare processes. We are led, then, to ask whether we can consider more general states, in which the spin connection is not equal to the gauge connection and the rank is reduced.

This is a complex subject, which has been only partially explored. At lowest order in the α' expansion there are such flat directions. They are not left–right symmetric and, while in order that they exhibit space–time supersymmetry they have two right-moving supersymmetries, they have no left-moving supersymmetry. So they are not suitable backgrounds for Type II theories and one cannot argue as easily as for the standard embedding that these $(0, 2)$ configurations are solutions of exact classical string equations. They are still subject to perturbative non-renormalization theorems in α'. But a detailed study of instanton amplitudes is required to determine whether these flat directions are lifted non-perturbatively, i.e. by corrections of the form $e^{-R^2\alpha'}$.

[1] Non-Abelian discrete symmetries offer possibilities for reducing the rank but we will not explore these here.

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

There is, however, a class of vacua with the Standard Model gauge group which can be found by symmetry arguments, much as we found additional flat directions in the Z_3 orbifold model. Consider, again, the quintic in CP^5, with the symmetric polynomial. We can find flat directions of the D terms by taking $27 = \overline{27}$. More precisely, starting with the E_6 decomposition into $O(10)$ representations,

$$27 = 10_1 + 1_{-2} + 16_{-1/2}, \tag{26.75}$$

we can give expectation values to the singlets in the $\overline{27}$ and one of the 27s. These are also flat directions of the F terms. For example, consider the 27 corresponding to the polynomial $Z_1 Z_2 Z_3 Z_4 Z_5$. The product $\overline{27}\ 27$ is invariant under all the discrete R-symmetries; no terms of the form $(\overline{27}\ 27)^n$ can appear in the superpotential. So this direction is exactly flat (terms of the form 27^3, $\overline{27}^3$ cannot lift these directions either). In combination with Wilson lines these flat directions readily break to the $SU(3) \times SU(2) \times U(1)$ group of the Standard Model.

26.6.7 Gauge coupling unification

One of the striking successes of low-energy supersymmetry is its prediction of unification. Within the context of grand unification – where the gauge group of the Standard Model is unified in a simple group at a scale M_{GUT} – the fact that the couplings unify is readily understood. In the context of the compactifications considered here it is not immediately obvious why this should be the case. In the case of symmetry breaking by Wilson lines, for example, the compactification scale and the scale of the symmetry breaking are of the same order. So there is no energy scale where one has a unified, four-dimensional effective theory.

In the weakly coupled heterotic string, however, the couplings *do unify* under rather broad conditions. In the case of Wilson line breaking this can be understood immediately in field-theoretic terms. The effect of the Wilson line is to eliminate states from the E_6 unified theory, but at tree level no couplings are altered. So the couplings of all groups emerging from E_6 are the same. Perhaps more surprising is the fact that the E_6 and E_8 couplings are the same. This can be seen by considering the vertex operators for the gauge bosons in each group. In both cases the vertex operators are constructed in terms of free two-dimensional fields, which obey the same algebra (in the unbroken subgroup) as in the flat-space theory. So, for example, the operator product expansions of these gauge boson vertex operators are unaltered. There are constructions where unification does not hold. They involve replacing the two-dimensional fermions with a current algebra having a different central extension.

In the (2, 1) flat directions considered above we can give an argument, based on the low-energy field theory, that the couplings remain unified as one moves out along the flat direction. A change in the coupling requires that there be a coupling of this modulus to the gauge fields. But, at the classical level, we know that there are no such couplings because any such coupling would violate the axion shift symmetry. This symmetry is unaffected by the expectation value of these moduli.

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

When we come to consider strongly coupled strings, the problem of coupling unification will be more complicated. It will be less clear in what sense unification is generic. Whether this is a problem for the theory, or a clue to a way forward, is a question for the student to ponder.

26.6.8 Calculating the parameters of the low-energy Lagrangian

As we have explained, on the one hand string theory is a theory without fundamental dimensionless parameters. On the other hand, the structure of the low-energy theory, as we now see, depends on discrete choices: which Calabi–Yau, orbifold etc.?; in how many dimensions?; with how much supersymmetry?; with which Wilson lines and continuous dynamical quantities, the moduli? For any given choice, at least classically it should be a straightforward problem to calculate the parameters of the low-energy theory.

It is easy to calculate the four-dimensional gauge couplings in terms of the ten-dimensional dilaton and the radius. We have already seen how this works for simple compactifications, and this carries over directly to the Calabi–Yau case since the vertex operators for the gauge fields are constructed in terms of two-dimensional fields, as in the orbifold or toroidal case.

The cosmological constant is another interesting quantity in the low-energy theory. At the classical level in the Calabi–Yau compactifications, it vanishes. This can be understood in a variety of ways. First, if we examine the solution of the ten-dimensional equations of motion, we see that since the Ricci tensor vanishes; there is no cosmological term. Second, in the two-dimensional conformal field theory the cosmological constant would give rise to a tadpole for the dilaton but this is forbidden by conformal invariance. Ultimately, the absence of a cosmological constant is inherent in the form of the solution: we have assumed that four dimensions are flat. We will see later that this is not necessary: string theory admits anti-de Sitter (AdS) spaces as well as Minkowski spaces as classical solutions.

From the perspective of trying to understand the Standard Model, a particularly important set of parameters is the set of Yukawa couplings. These can certainly be computed in string theory. In principle we should construct the vertex operators for the appropriate matter fields and then construct the required OPE coefficients or suitable scattering matrices. In practice this can often be short-circuited. In the orbifold models, for example, in the untwisted sectors we can read off the Yukawa couplings by dimensional reduction of the ten-dimensional Lagrangian. The scalar fields are components of the original ten-dimensional gauge fields A_i. Similarly, the fermions are components of the ten-dimensional gauginos. In the orbifold theory, alternatively it is not difficult to construct the vertex operators and to compute the required OPE coefficients.

In the Calabi–Yau case we have seen that, in the α' expansion, the superpotential is independent of R. So one can work at very large radius and pick out the leading contribution. To actually do the computation one can construct the zero modes of the scalar and spinor fields and substitute into the Lagrangian. A priori one might expect that this would be quite difficult, given that one does not have an explicit formula for the metric. But it turns out that the Yukawa couplings have a topological significance, and their values

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

can be inferred by general reasoning. We will not have a particular use for explicit formulas here, but it is important to be aware of their existence.

26.6.9 Other perturbative heterotic string constructions

The quintic is just one of a large class of Calabi–Yau models which can be constructed. The exact number is not actually known. It is not even known, with certainty, whether the number of Calabi–Yau vacua is finite or infinite.

So, while we will not assess here the size of this space, there is clearly a large class of string solutions with gauge group identical to that of the Standard Model. These theories have varying numbers of generations, including both orbifold (or free-fermion) models and Calabi–Yau constructions with three generations. There are many models with groups, numbers of generations, and other features radically different from those of the Standard Model. Still, it is remarkable how easily we have obtained models which accord with some of our speculations for Beyond the Standard Model physics. We have found low-energy supersymmetry, coupling unification, light Higgs particles and discrete symmetries which can potentially suppress proton decay and give rise to a stable dark matter candidate, all in a framework where we can imagine that real calculations are possible.

In subsequent chapters we will turn to the problems of actually turning these observations and discoveries into a real theory which we can confront with experiment.

Suggested reading

Volume 2 of Green *et al.* (1987) provides a comprehensive introduction to Calabi–Yau compactification, and I have borrowed heavily from their presentation. Weakly coupled string models with three generations have been constructed in the context of Calabi–Yau compactification; their phenomenology is considered by Greene *et al.* (1987). Models based on free fermions were been constructed by Faraggi (1999). We will encounter non-perturbative constructions in Chapter 28. At special points in their moduli spaces, some Calabi–Yau spaces can be described in terms of solvable conformal field theories. This program was initiated by Gepner (1987) and is described at some length by Polchinski (1998). A very accessible description, including computations of physically interesting couplings, appears in Distler and Greene (1988).

Exercises

(1) Write down the field strength of electrodynamics as a two-form and express its gauge invariance in the language of forms. Verify that $dF = 0$ is equivalent to the Bianchi identity (the homogeneous Maxwell equations).

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press

(2) Show that, for a Kahler manifold, the non-vanishing components of the affine connection (Christoffel symbols) are given by Eq. (26.36). Then show that the non-zero components of the Riemann tensor are given by Eq. (26.37) and verify Eq. (26.38). Derive Eq. (26.40) by noting that

$$\Gamma^{\bar{a}}_{\bar{b}\bar{a}} = \partial_{\bar{b}} \ln \det g. \tag{26.76}$$

Show that our result for the two-dimensional curvature of a Riemann surface is a special case of this.

(3) For a flat two-dimensional torus, introduce complex coordinates and verify that the bosonic and fermionic terms are just those of the free string action. You can take $K = X^\dagger X$ for this case.

(4) Write out in some detail the action of the heterotic string propagating in the Calabi–Yau background with spin connection equal to the gauge connection. Determine the form of the vertex operators for the 27 and $\overline{27}$ fields, in the RNS formulation (you can limit yourself to the NS–NS sector).

(5) Exhibit a combination of Wilson lines and $SU(5)$ singlet expectation values which breaks the gauge group to that of the Standard Model in the case of the quintic in CP^4.

https://doi.org/10.1017/9781009290883.033 Published online by Cambridge University Press