Searching for the standard model in the string landscape: SUSY GUTs

Stuart Raby

Department of Physics, The Ohio State University, 191 W. Woodruff Ave., Columbus, OH 43210, USA
E-mail: raby@pacific.mps.ohio-state.edu

Received 23 September 2010, in final form 6 December 2010
Published 28 February 2011
Online at stacks.iop.org/RoPP/74/036901

Abstract
The standard model is the theory describing all observational data from the highest energies to the largest distances. (There is, however, one caveat: additional forms of energy, not part of the standard model, known as dark matter and dark energy must be included in order to describe the Universe at galactic scales and larger.) High energies refers to physics at the highest energy particle accelerators, including CERN’s LEP II (which ceased operation in 2000 to begin construction of the Large Hadron Collider now in operation) and Fermilab’s Tevatron, as well as to the energies obtained in particle jets created in so-called active galactic nuclei scattered throughout the visible Universe. Some of these extra-galactic particles bombard our own Earth in the form of cosmic rays, or super-energetic protons which scatter off nucleri in the upper atmosphere.

String theory is, on the other hand, an unfinished theoretical construct which attempts to describe all matter and their interactions in terms of the harmonic oscillations of open and/or closed strings. It is regarded as unfinished since at present it is a collection of ideas, tied together by powerful consistency conditions, called dualities, with the ultimate goal of finding the completed string theory. At the moment we only have descriptions which are valid in different mutually exclusive limits with names such as type I, IIA, IIB, heterotic, M and F theory. The string landscape has been described in the pages of many scholarly and popular works. It is perhaps best understood as the collection of possible solutions to the string equations; albeit these solutions look totally different in the different limiting descriptions. What do we know about the string landscape? We know that there are such a large number of possible solutions that the only way to represent this number is as $10^{500}$ or a 1 followed by 500 zeros. Note that this is not a precise value since the uncertainty is given by a number just as large. Moreover, we know that most of these string states look nothing like the standard model. They have the wrong matter and wrong forces. Moreover, they are not off by a small amount, they are totally wrong. So the question becomes, does string theory really describe our observable world? In order to address this question, one must find at least one string state that resembles it. One possibility is that our observable world is in fact a unique string state. If this is the case, then the problem becomes one of finding the proverbial needle in the largest possible haystack! On the other hand, there may be many states which are sufficiently close to the observable world, and we need only to understand why we are in this finite subspace of the string landscape. And perhaps there are good reasons why this subspace is preferred over 99.999999999...% of the myriad of non-standard-model-like string states. Perhaps, just by confining our attention to this subspace we can learn something about our observable world which we cannot learn otherwise. Thus the goal of this work is to understand what it takes to find the standard model in the string landscape.

(Some figures in this article are in colour only in the electronic version)

This article was invited by S F King.

1 OHSTPY-HEP-T-10-004.
1. Introduction

1.1. Prologue

In this paper I will discuss several recent attempts to find the standard model in the string landscape. I will argue that the most efficient path is paved with two new proposed symmetries of nature, called grand unification and supersymmetry. Together they make the journey exciting and inspiring. In order to bring this adventure to you, the reader, I should first define what I mean when I talk of the standard model, then describe these two new symmetries and finally discuss how to find the standard model in the string landscape. This is the goal of this paper.

1.2. The standard model

The standard model is defined by the collection of four fundamental forces—strong, weak, electromagnetic and gravitational and the matter particles which interact via these forces. The last ingredient of the standard model, yet to be observed directly in an experiment, is a fifth force, described by the Higgs interaction. The name standard model is archaic and no longer appropriate. It was coined in the early 1970s when the standard model was the best proposal for a new theory describing all known phenomena and thus the theory to be tested in every possible way. After some 30 odd years and millions of experimental tests later, the standard model is now the accepted theory of all observed properties of nature. However, with this major correction the standard model is truly an accepted theory\(^2\).

The standard model includes three families of quarks and leptons. The quarks come in six different flavors: up, down, charm, strange, top and bottom, with each flavor coming in three colors. Quarks feel the strong force, exchanging gluons which themselves come in eight colors. Leptons are, however, color singlets. The three-fold symmetry of the strong interactions is described by the mathematical group called SU\(^3\)color. Prior to obtaining mass via the Higgs VEV, massless quarks separate into left-handed and right-handed components. So what is handedness? Quarks and leptons are spin 1/2 particles, called fermions. If a fermion’s spin points in the direction opposite to its motion, it is called left-handed and if its spin points in the same direction as its motion, it is called right-handed. This may seem like a trivial distinction, but in fact the weak doublets are only left-handed. So it is necessary to distinguish left-handed and right-handed fermion fields, because nature does. To describe an up quark we then need two fields, \(u\) and \(\bar u\). The first up field annihilates a left-handed up quark and creates a right-handed anti-up quark. The second annihilates a left-handed anti-up quark (hence the bar) and creates a right-handed up quark. Then the weak boson \(W^+\) takes \(d\) into \(u\), and \(W^-\) does the reverse. We describe this doublet (under the weak SU\((2)_L\) group) as a single field \(q\) with two pieces, i.e. \(q = (u, \bar u)\).

Note that in the standard model, all particles are described by relativistic quantum fields\(^3\). It is a property of relativity\(^2\) that said, the standard model has two major deficiencies with regards to cosmology and astrophysics. It lacks an explanation for the observed dark matter and dark energy of the Universe. These topics are typically reserved for a discussion of the physics beyond the standard model.

\(^2\) Again, one caveat with respect to gravity and the rest of the standard model. A self-consistent quantum mechanical treatment of gravity is problematic. This concerns the description of black holes and the so-called ‘information loss paradox’. It is this paradox which calls for a better description of gravity and the standard model.
that each particle has an anti-particle with identical mass but opposite charges under all three forces. Moreover, when a particle and its own anti-particle meet, they annihilate into anything which can be produced while conserving energy. This may sound like an esoteric, Star Wars like, concept; it is reality. In fact, a positron emission tomography (PET) scan that you may have seen in your local hospital is a manifestation of this fact. A PET scan is a process whereby a radioactive isotope is administered to the patient. This radioactive isotope emits positrons (or anti-electrons). The isotope travels to a spot depending on the chemical properties of the isotope. For example, $^{58}$Ga may be used for a PET scan of the brain. When the positrons are emitted they quickly annihilate with local electrons. Two photons are created which travel to the detector moving out in opposite directions. Each photon, sometimes called a gamma ray, $\gamma$, has energy equal to the mass of the electron which disappeared. The energy of the photon satisfies the famous Einstein relation, $E_\gamma = mc^2$.

So the first family of quarks and leptons is obtained from the fermion fields:

\[
\begin{aligned}
q &= \left( \begin{array}{c} u \\ d \end{array} \right), \quad \bar{u}, \bar{d}, \quad l = \left( \begin{array}{c} \nu \\ e \end{array} \right), \quad \bar{\nu}, \bar{e},
\end{aligned}
\]  

Note that the quark fields $u$, $d$ are color triplets, $\bar{u}$, $\bar{d}$ are anti-triplets and leptons are color singlets, i.e. they do not have color charge. The fields $q$, $l$ are weak $SU(2)_L$ doublets and under weak hypercharge, $U(1)_Y$, these fields have charge

\[
Y = \{1/3, -4/3, 2/3, -1, 2, 0\}.
\]  

Note that electric charge, $Q$, of these fields is given by the simple formula

\[
Q = T_3 + \frac{Y}{2}.
\]  

For example, in the lepton doublet we have $T_3 = +1/2$ for the neutrino, $\nu$, and $T_3 = -1/2$ for the electron, $e$. Hence the neutrino (electron) has electric charge 0 ($-1$), respectively. The anti-electron field, $\bar{e}$, has $T_3 = 0$, since it is not part of a weak doublet (we say it is a weak singlet) and thus has electric charge $+1$. It is then a simple exercise to work out the electric charge of all the fields in one family. We find that quarks have fractional charge, $2/3$ for $u$ and $-1/3$ for $d$. Finally, $\bar{\nu}$ is neutral under all three standard model (SM) symmetries, $SU(3)_{color} \times SU(2)_L \times U(1)_Y$. It is a so-called ‘sterile’ neutrino whose sole purpose, as we shall discuss later, is to give neutrinos mass. Note that there are inequivalent ways of introducing neutrino masses in the literature. In the later discussion of neutrino masses, I will only consider the simplest possibility known as the type I see-saw mechanism. In this case only SM singlet states, such as the ‘sterile’ neutrinos or SM singlet scalars, are added to the original version of the SM, where neutrinos were, in fact, massless.

Yet the only particles we observe in nature have integral charge. This is because a proton is a bound state of two up quarks and one down quark, while a neutron is a bound state of one up quark and two down quarks. Protons and neutrons have electric charge +1 and 0, respectively. The strong force is extremely strong and is so strong that an isolated free quark cannot exist. Only color singlet combinations of a red up, a blue up and green down is allowed. However, the residual strong force between color singlet combinations of quarks in protons and neutrons holds these particles tightly bound in the nuclei of atoms. The electromagnetic force is not as strong but then binds electrons and protons together to form electrically neutral atoms.

In fact, the electromagnetic force is many times stronger than gravity. Gravity holds us to the Earth, but the electromagnetic force between atoms in matter keep us from falling through the floor! The electromagnetic force is so very strong that no person on Earth is strong enough to hold two electrons together, assuming these were the only two charged particles around. So if atoms made of electrons and protons were not electrically neutral, we could not possibly have structure in the Universe. Tables, chairs, stars and galaxies would be pulled apart by the strong electromagnetic repulsion. So it is a wonderful fact that charge is quantized and protons and electrons have equal and opposite charges, so when they are put together they make electrically neutral atoms. This property of the standard model is known as charge quantization.

There are three families of quarks and leptons. All the states of the first family, with all their charges under the strong and electroweak forces, are duplicated with a second family and then a third family. The only difference is that the second family particles are heavier than the first and the third family is even heavier than the second family. It is a complete mystery why nature would require three almost identical families of quarks and leptons.

The standard model describes all known phenomena both on experiments performed on Earth as well as a description of cosmology and astrophysics starting as way back in time as the first 3 min of the Universe. All of this can be described with just 28 fundamental parameters (see table 1). These are the three low energy gauge coupling constants, nine parameters associated with quark and charged lepton masses, four quark weak mixing angles contained in the Cabibbo–Kobayashi–Maskawa (CKM) matrix and nine parameters associated with neutrino masses and mixing.

Table 1. The 28 parameters of the standard model. The values for all of these parameters are set by experiment. Note that the charges of the quarks and leptons under $SU(3)_{color} \times SU(2)_L \times U(1)_Y$ are typically not included in this list. They are nevertheless important ingredients of the theory, whose origin is a great mystery.

| Sector                | # Parameters |
|-----------------------|--------------|
| Gauge                 | 3 $\alpha_1, \alpha_2, \alpha_3$ |
| Quark masses          | 6 $m_u, m_d, m_s, m_c, m_t, m_b$ |
| Quark weak mixing angles | $V_{ub}, V_{cb}, V_{td}, \delta$ |
| Charged lepton masses | 3 $m_e, m_\mu, m_\tau$ |
| Neutrino masses       | 3 $m_\nu_1, m_\nu_2, m_\nu_3$ |
| Lepton weak mixing angles | $\theta_{13}, \theta_{12}, \theta_{23}, \delta_i, i = 1, 2, 3$ |
| Electroweak scales    | 2 $M_Z, m_{Higgs}$ |
| Strong CP angle       | 1 $\theta$ |
1.3. Supersymmetric grand unified theories (SUSY GUTs)

SUSY GUTs provide a framework for solving many of the problems of the standard model. But let us first describe these two symmetries one-by-one.

1.3.1. Grand unification. Grand unified theories unify the three fundamental interactions of the standard model. These are the strong forces, represented by a three-fold symmetry called color; the weak forces, responsible for one form of radioactive nuclear decay which powers the engine for stellar burning, including our own Sun; and the electromagnetic force, associated with the first relativistic field theory unifying the theories of electric and magnetic phenomenon into one. The latter two forces are themselves intertwined in what is called the electroweak theory, represented by a two-fold symmetry called weak isospin and and a phase symmetry called weak hypercharge. It is an amazing fact of nature that the electroweak symmetry is a symmetry of equations of motion of the theory, but not a symmetry of the vacuum. In fact, this very symmetry breaking is due to the vacuum expectation value of the Higgs field. And this non-zero expectation value is responsible for the mass of the weak bosons, \( W^\pm \) and \( Z^0 \), and all quarks and leptons.

Grand unified theories provide an explanation of these amazing facts. By unifying the three symmetries, \( SU(3)_{\text{color}}, SU(2)_L, U(1)_Y \), one also unifies the fermions of the standard model into multiplets transforming under the new grand unified symmetry group. In fact the simplest extension of the standard model, the symmetry group \( SO(10) \) [1], puts all the fermions of one family into one single representation, see table 2. This is a phenomenal fact and a significant point in favor of grand unification.

\( SO(10) \) contains the SM as subgroup. Note that the spinor representation of \( SO(10) \) is represented as a direct product of five spin 1/2 states with spin in the third direction labeled \( \pm (1/2) \), \( SU(3) \) acts on the first three spin states by raising (or lowering) one spin and leaving (or raising) another. Hence the states represented by \( - + + ; + - + ; + + - \) are an \( SU(3) \) triplet and \( + + + ; - - + ; - - - \) are singlets. \( SU(2) \) acts on the last two spin states, with \(-; +; +; -; -; +; +\) representing an \( SU(2) \) doublet and \(+; +; +; -; -; -\) are singlets. Finally, electroweak hypercharge \( Y \) is given by \( Y = \frac{2}{3} \sum_{i=1}^{3} S_i - \sum_{i=1}^{3} \bar{S}_i \), with \( S_i = \pm \frac{1}{2} \). \( SO(10) \) also has two very interesting subgroups, Pati–Salam \( [2] \), \( SO(6) \otimes SO(4) = SU(4)_C \otimes SU(2)_L \otimes SU(2)_R \) and Georgi–Glashow \( [3] \). \( SU(5) \) multiplets can be obtained by raising one spin and lowering another spin out of the full tensor product state. As such, one sees that the spinor representation of \( SO(10) \) breaks up into three irreducible representations of \( SU(5) \) given by the multiplets with zero, two or four spin down states, corresponding to an \( SU(5) \) \( 1, 10 \) or \( \bar{5} \). In particular, note that the \( 10 = [u, d, \ell] \), \( \bar{5} = [d, l] \) and \( 1 = \bar{\nu} \). The full \( SO(10) \) operations then allow for simultaneously raising or lowering two different spins. Similarly, \( SU(4)_C \) is an extension of \( SU(3)_C \) with the additional operations of raising or lowering two different color spins. Hence the spinor representation of \( SO(10) \) decomposes to two irreducible multiplets of Pati–Salam, with \( Q = [q, l] = (4, 2, 1) \) and \( \bar{Q} = [\bar{q}, \bar{l}] = (\bar{4}, 1, 2) \) with \( SU(2)_R \) represented by the operation of raising or lowering two different weak indices.

The three fundamental forces, defined by the symmetries \( SU(3)_{\text{color}}, SU(2)_L, U(1)_Y \), couple to matter proportional to the three fine-structure constants, \( \alpha_3(\mu), \alpha_2(\mu), \alpha_1(\mu) \) defined at some energy scale \( \mu \). A grand unified gauge theory, such as \( SO(10) \), on the other hand has only one coupling constant, \( \alpha_{\text{GUT}} \). Note that the values of the observed coupling constants depend on the energy scale at which they are measured. Indeed we are able to calculate the energy dependence of the fine-structure constants, if one knows the mass spectrum of particles which carry strong and electroweak charge. If one assumes that only particles observed to date and the theoretically predicted Higgs boson contribute to the so-called renormalization group running of the standard model couplings one finds the result given in figure 1 (left). Although the three fine-structure constants approach to a common value, they do NOT actually meet. This is a significant problem with gauge coupling unification. However, this calculation explicitly assumes no new particle states above the weak scale, since only particle states with mass less than the relevant energy scale affect the running of the coupling. This assumption will shortly be challenged.

Assuming grand unification at a scale \( M_{\text{GUT}} \sim 10^{15} \text{ GeV} \), introduces a fundamental problem. Why is the weak scale, \( M_{\text{weak}} \sim 100 \text{ GeV} \), so much smaller than the GUT scale? This is particularly problematic since the weak scale is set by the Higgs boson mass. Note that boson masses are sensitive to physics at the highest scales. This is because quantum mechanical corrections to any boson mass are proportional to some effective coupling \( \alpha \) times the largest scale in the theory. In this case we would expect the Higgs mass to be given at leading order by \( m_{\text{H}} \approx m_0^2 + \alpha M_{\text{GUT}}^2 + \cdots \) where \( m_0 \) is the Higgs mass prior to quantum corrections. Why the
Figure 1. Gauge coupling unification in non-SUSY GUTs on the left versus SUSY GUTs on the right using the LEP data as of 1991. Note that the difference in the running for SUSY is the inclusion of supersymmetric partners of standard model particles at scales of the order of a TeV. Given the present accurate measurements of the three low energy couplings, in particular $\alpha_s(M_Z)$, GUT scale threshold corrections are now needed to precisely fit the low energy data. The dark blob in the plot on the right represents these model-dependent corrections.

Figure 2. This figure represents the anti-symmetric wave function for two identical fermions at positions $x_1$, $x_2$. Note that due to the anti-symmetry of the wave function, it is not possible to have two fermions at the same position in space, at the same time.

1.3.2. Supersymmetry. Supersymmetry is the largest possible symmetry of space-time, i.e. it is an extension of Einstein’s special theory of relativity and ordinary space-time translation invariance which adds a new quantum direction of space. This quantum direction of space, $\theta$, is, in fact, infinitesimal since $\theta^2 = 0$. Moreover, there is at least two such directions, $\theta_1$, $\theta_2$ with the property that $\theta_1 \theta_2 = -\theta_2 \theta_1$. Ordinary coordinates of space and time are pure numbers, where given two different directions $x$ and $y$, we have $xy = yx$. One obtains the same number no matter which order these two numbers are multiplied. The difference in these multiplication rules is identical to the difference between the quantum statistics of identical fermions and identical bosons. Supersymmetry is a rotation of ordinary space into superspace where the $\theta$-like coordinates rotate into the ordinary space-time coordinates.

Let us take a small digression which we will quickly relate to the above discussion. Matter particles, such as electrons, protons and neutrons, are spin 1/2 particles, known as fermions. They satisfy the Pauli exclusion principle. This is the property postulated by Wolfgang Pauli which says that no two identical fermions can be put in the same position at the same time. It is mathematically expressed by the form of the quantum mechanical wave function for two identical fermions. The square of the wave function gives the probability for finding the fermions at points $x_1$ and $x_2$ in space, at the same time (see figure 2). According to the Pauli exclusion principle, this wave function must be anti-symmetric under interchanging the two fermion positions. As a result the probability for finding two identical fermions at the same point in space, at the same time, vanishes. This property is in agreement with the intuitive property of matter.

ratio $m_H/M_{GUT} \sim 10^{-13}$ is known as the ‘gauge hierarchy problem’. One possible solution to the gauge hierarchy problem is to postulate a new symmetry of nature, so-called ‘supersymmetry’. Supersymmetry is, as it sounds, a pretty amazing symmetry.

All the force particles of nature have integral spin. Photons, the weak particles, $W^\pm$, $Z^0$, and strong interaction gluons are all spin 1. Gravitons are spin 2 and the postulated Higgs particle has spin 0. According to the physicists Bose and Einstein, all identical integral spin particles, called bosons, satisfy Bose–Einstein statistics. The quantum mechanical wave function describing two identical bosons must be symmetric under interchange (see figure 3). Hence it is most probable to find two identical bosons at the same position in space at the same time. Moreover, it is more probable to have any number of bosons sitting right on top of each other. In fact, with millions and millions of photons sitting in the same quantum state one obtains a macroscopic electric and magnetic field. It is no wonder that all force particles are bosons.

The bosonic wave function behaves as ordinary spatial coordinates. We say that they commute with each other. On the other hand, interchanging two fermions changes the sign of the wave function. We say that they anti-commute, just like the new anti-commuting coordinates of supersymmetry. In fact, supersymmetry is a symmetry which rotates bosonic...
particles into fermionic ones and vice versa. Supersymmetric rotations change the spin and statistics of the particles, but not their other properties. For example, if supersymmetry were an exact symmetry, then for every spin 1/2 electron there would necessarily exist a spin 0 bosonic electron, a scalar electron or (selectron). For every spin 1 photon, there would necessarily exist a spin 1/2 fermionic photon or (photino). In the minimal supersymmetric extension of the standard model or (MSSM), every particle has its supersymmetric partner. So why would we want to postulate doubling the entire particle spectrum? In addition to the fact that supersymmetry is the unique extension of special relativity, i.e. the largest possible symmetry of nature in four dimensions, supersymmetric theories, as developed by Wess and Zumino, and Salam and Strathdee, are much better behaved than ordinary relativistic field theories. As a result supersymmetry can solve the gauge hierarchy problem [4]. Although an ordinary boson gets quantum mass corrections which scale like the largest mass in the theory, fermions on the other hand are protected from such large corrections. In fact it was shown long ago that due to the handedness (or chiral) symmetry of fermions their mass terms receive quantum corrections which are only logarithmically sensitive to the highest mass scale in the theory. For example, the electron mass at leading order receives a quantum correction of the form, $m_e \approx m_0 (1 + \log M_{GUT}/m_0)$. In supersymmetric theories bosons are protected by the chiral symmetry of their fermionic partners. In the MSSM all quarks and leptons have scalar partners, called squarks and sleptons, and all gauge bosons have fermionic partners, called gauginos. Moreover, the Higgs doublets have fermionic partners, called Higgsinos. As long as supersymmetry is unbroken, then the SM particles and their superpartners are degenerate. Hence the Higgs scalars are protected from large radiative corrections. However, once SUSY is spontaneously broken then scalars will naturally obtain mass of the order of the SUSY breaking scale, due to radiative corrections. However, for gauginos to obtain mass, SUSY as well as the chiral symmetry for gauginos, known as $R$ symmetry, must be broken.

Including the supersymmetric particle spectrum (at a scale of the order of 1 TeV, as required to solve the gauge hierarchy problem) in the renormalization group running of the three low energy fine structure constants results in the graph of figure 1 (right). Miraculously the three gauge coupling constants now unify at a GUT scale of the order of $3 \times 10^{16}$ GeV [5]. As we shall argue there are three experimental pillars of SUSY GUTs:

(i) gauge coupling unification,
(ii) low energy supersymmetry,
(iii) proton decay.

Hence the first experimental pillar of SUSY GUTs stands firm due to the LEP data from 1991. The CDF and DZero experiments at the Tevatron accelerator at Fermi National Accelerator Laboratory in Batavia, Illinois, are searching for the SUSY particles. And now the LHC at CERN will begin the search for supersymmetry. The discovery of super particles at the Fermilab Tevatron or by the CMS or ATLAS detectors at the LHC will verify the second pillar of SUSY GUTs. Supersymmetry is clearly the new standard model of the new millennium. It is a new symmetry of nature which most elementary particle theorists would expect to show up at the TeV energies accessible to the LHC. The third pillar, which may take some time, is the observation of proton decay! Experiments searching for proton decay are on-going at Super-Kamiokande in Kamioka, Japan. However, we may have to wait for the next generation of proton decay experiments, such as the large water Čerenkov detector at the planned Deep Underground Science and Engineering Laboratory (DUSEL) in South Dakota.

1.4. Virtues of SUSY GUTs

The framework provided by SUSY GUTs can be used to address the following open problems of the standard model. In particular,

(i) it can ‘naturally’ explain why the weak scale $M_Z \ll M_{GUT}$;
(ii) it explains charge quantization;
(iii) it predicts gauge coupling unification (!);
(iv) it predicts SUSY particles at the LHC, and
(v) it predicts proton decay.

In addition, it is a natural framework for understanding quark and lepton masses and it has non-trivial consequences for astrophysics and cosmology.

(i) It predicts Yukawa coupling unification,
(ii) and with family symmetry can explain the family mass hierarchy.
(iii) It can accommodate neutrino masses and mixing via the see-saw mechanism, and
(iv) it can generate a cosmological asymmetry in the number of baryons minus anti-baryons, i.e. baryogenesis via leptogenesis, via the decay of the heavy right-handed neutrinos.

(v) In the MSSM with $R$ parity, the lightest supersymmetric particle (LSP) is a dark matter candidate.

For all of these reasons we might suspect that SUSY GUTs are a fundamental component of any realistic string vacuum. And thus if one is searching for the MSSM in the mostly barren string landscape, one should incorporate SUSY GUTs at the first step. So what do string theories have to offer?
1.5. String theory

As I explained earlier, the term the standard model is a severe understatement and misnomer, since the standard model has now been verified in millions of ways at multiple laboratories on Earth and in the Heavens. Similarly the term string theory is a misnomer and a major overstatement, since there is no overall theoretical construct which describes in one formalism all the properties attributed to the many faces of string theory.

String theory evolves by studying the several different limiting forms of the theory, called type I, type IIA, type IIB, $E_8 \otimes E_8$ and $SO(32)$ heterotic, M and F theory. These limiting forms can all be studied in their respective perturbative limits. Then they are all tied together via so-called duality transformations which relate the solutions of the different theories.

That said, perturbative string theory starts by studying the oscillations of a one-dimensional vibrating string moving along some trajectory in space with time. Type I describes open and closed strings, while type IIF and heterotic theories describe the motion of closed strings. The space-time in which they move is 10-dimensional (10D). Finally, M theory can all be studied in their respective perturbative limits. Then they are all tied together via so-called duality transformations which relate the solutions of the different theories.

The normal oscillation modes of the string are set by the string tension (or string scale—$M_S$) which is typically taken to be a scale of order the 4D Planck scale—$M_{Pl}$. At energies below $M_S$ only massless string modes can be excited. String theories also have the amazing property that they are finite theories, i.e. all field theoretic divergences are naturally cut-off at $M_S$. The effective theory of the massless excitations of the string is typically described by supergravity field theory (supersymmetric field theory, plus the supersymmetric version of Einstein gravity) including all the gauge interactions and matter multiplets described by the massless modes of the string. In fact, the amazing property of strings is that a massless spin 2 graviton (the particle of Einstein gravity) naturally appears in the spectrum. This is why Scherk and Schwarz first proposed that the string scale should satisfy $M_S \sim M_{Pl}$, since $M_S$ sets the scale for the graviton to couple to matter (recall Newton’s constant of gravity $G_N = M_S^{-2}$). String theory excitations also naturally contain massless spin one excitations with properties like the standard model gauge particles, such as the photon, and even matter multiplets. However, the particular massless spectrum of the string depends in detail on the choice of a string vacuum. The problem is that there does not appear to be a unique choice of string vacuum. Moreover, in the supersymmetric limit, there are a continuous infinity of possible string vacua. And even after supersymmetry is spontaneously broken, there are still estimated to be of order $10^{500}$ string vacua.

So how would one ever expect to be able to find the standard model in this huge landscape of string vacua? The analog of finding the standard model in the string landscape may be like finding a golf ball on the surface of the Earth. The surface of the Earth is huge with many mountains, valleys, deserts and oceans and any random search over this huge landscape will have almost zero probability of finding a golf ball. However, if the search was not random and one first found all the golf courses on the Earth, then the probability of finding a golf ball would increase dramatically! I propose that the analog of the golf course for our problem is SUSY GUTs. When one first looks for SUSY GUTs in the string landscape, the probability of finding the standard model jumps by many, many orders of magnitude.

1.6. Preview

In the introduction I have tried to provide the basic motivation and phenomena associated with supersymmetric grand unified theories and super strings. In the rest of this article I will discuss more details associated with SUSY GUTs in four dimensions. I will then introduce the concept of GUTs in extra dimensions, so-called orbifold GUTs, which will then take us to the goal of string GUTs. Or finding the MSSM in the string landscape. The bottom line of this discussion is the assertion that in order to find the MSSM one must first find regions of the string landscape containing SUSY GUTs. In these regions the MSSM becomes natural.

2. Four-dimensional SUSY GUTs and gauge coupling unification

Now that we have identified the states of one family in $SU(5)$, let us exhibit the fermion Lagrangian (with gauge interactions). We have

$$L_{\text{fermion}} = \frac{\bar{\psi}}{2}(i\gamma_{\mu}D^\mu)\psi + 10\epsilon_{\alpha\beta}(i\bar{\psi}D^\mu)\psi \frac{\sqrt{10}}{\sqrt{5}}$$

where

$$D^\mu = \partial^\mu + ig A^A_{\mu}$$

and $A_{\mu}$ is in the 5 or 10 representation. We see that since there is only one gauge coupling constant at the GUT scale we have

$$g_3 = g_2 = g_1 = g$$

where, after weak scale threshold corrections are included, we have

$$g_3 \to g_3, \quad g_2 \to g, \quad g_1 \to \frac{\sqrt{5}}{\sqrt{3}} g.$$  \hspace{1cm} (7)

At the GUT scale we have the relation

$$\sin^2 \theta_W = \frac{(g')^2}{g^2 + (g')^2} = 3/8.$$  \hspace{1cm} (8)

But these are tree-level relations which do not take into account threshold corrections at either the GUT or the weak scales nor renormalization group (RG) running from the GUT scale to the weak scale. Consider first RG running. The one-loop RG equations are given by

$$\frac{d\alpha_i}{dt} = -\frac{b_i}{2\pi} \alpha_i^2$$  \hspace{1cm} (9)

where $\alpha_i = \frac{g_i^2}{4\pi}$, $i = 1, 2, 3$ and

$$b_i = \frac{4}{3} C_2(G_i) - \frac{2}{3} T_R N_F - \frac{1}{3} T_R N_S.$$  \hspace{1cm} (10)
Note that $t = -\ln\frac{M_{\text{GUT}}}{M_\text{TH}}$, $\sum_i (T^2_i) = C_2(G_i)$ with $T_A$ in the adjoint representation defines the quadratic Casimir for the group $G_i$, with $C_2(SU(N)) = N$ and $C_2(U(1)) = 0$. \( Tr(T_A T_B) = T_R \delta_{AB} \) for $T_A$ and $T_B$ in the representation $R$ for $U(1)_Y$, $T_R = \frac{1}{2} \, Tr(T^2)$ and $N_F(N_G)$ is the number of Weyl fermions (complex scalars) in representation $R$. For $N = 1$ supersymmetric theories, equation (10) can be made more compact. We have

$$b_i = 3 C_2(G_i) - T_R \, N_f$$

(11)

where the first term takes into account the vector multiplets and $N_f$ is the number of left-handed chiral multiplets in the representation $R$ [5,6]. The solution to the one-loop RG equations is given by

$$\alpha_i(M_Z)^{-1} = \alpha_i^{-1} - b_i \ln \left( \frac{M_G}{M_Z} \right).$$

(12)

For the SM we find

$$b_{\text{SM}} = (b_1, b_2, b_3) = (-\frac{2}{3}N_{\text{fam}} - \frac{22}{9}N_{H}, \frac{5}{7}, \frac{3}{7}N_{\text{fam}}$$

(13)

where $N_{\text{fam}}$ is the number of families (Higgs doublets). For SUSY we have

$$b_{\text{SUSY}} = (-2N_{\text{fam}} - \frac{2}{3}N_{H} + \frac{2}{3}N_{H}, 6 - 2N_{\text{fam}}$$

(14)

where $N_{H}$ is the number of pairs of Higgs doublets. Thus for the MSSM we have

$$b_{\text{MSSM}} = (-33/5, -1, 3).$$

(15)

The one-loop equations can be solved for the value of the GUT scale $M_G$ and $\alpha_G$ in terms of the values of $\alpha_{\text{EM}}(M_Z)$ and $\sin^2 \theta_W(M_Z)$. We have (without including weak scale threshold corrections)

$$\alpha_2(M_Z) = \frac{\alpha_{\text{EM}}(M_Z)}{\sin^2 \theta_W(M_Z)}$$

and we find

$$= \left( \frac{3}{5} \frac{8}{5} \sin^2 \theta_W(M_Z) \right) \alpha_{\text{EM}}(M_Z)^{-1}$$

$$= \left( \frac{b_{\text{MSSM}}^{\text{MSSM}} - b_1^{\text{MSSM}}}{2\pi} \right) \ln \left( \frac{M_G}{M_Z} \right)$$

(17)

which we use to solve for $M_G$. Then we use

$$\alpha_G^{-1} = \sin^2 \theta_W(M_Z) \alpha_{\text{EM}}(M_Z)^{-1} + \frac{b_{\text{MSSM}}^{\text{MSSM}}}{2\pi} \ln \left( \frac{M_G}{M_Z} \right)$$

(18)

to solve for $\alpha_G$. We can then predict the value for the strong coupling using

$$\alpha_3(M_Z)^{-1} = \alpha_G^{-1} - \frac{b_1^{\text{MSSM}}}{2\pi} \ln \left( \frac{M_G}{M_Z} \right).$$

(19)

Given the experimental values $\sin^2 \theta_W(M_Z) \approx 0.23$ and $\alpha_{\text{EM}}(M_Z)^{-1} \approx 128$ we find $M_G \approx 1.3 \times 10^{13}$ GeV with $N_H = 1$ and $\alpha_G^1 \approx 42$ for the SM with the one-loop prediction for $\alpha_3(M_Z) \approx 0.07$. On the other hand, for SUSY we find $M_G \approx 2.7 \times 10^{16}$ GeV, $\alpha_G^1 \approx 24$ and the predicted strong coupling $\alpha_3(M_Z) \approx 0.12$. How well does this agree with the data? According to the PDG the average value of $\alpha_3(M_Z) = 0.1176 \pm 0.002$ [7]. So at one loop the MSSM is quite good, while non-SUSY GUTs are clearly excluded.

At the present date, the MSSM is compared with the data using 2 loop RG running from the weak to the GUT scale with one-loop threshold corrections included at the weak scale. These latter corrections have small contributions from integrating out the $W$, $Z$, and top quark. But the major contribution comes from integrating out the presumed SUSY spectrum. With a ‘typical’ SUSY spectrum and assuming no threshold corrections at the GUT scale, one finds a value for $\alpha_3(M_Z) \geq 0.127$ which is too large [8]. It is easy to see where this comes from using the approximate analytic formula

$$\alpha_3^{-1}(M_Z) = \alpha_G^{-1} - \frac{b_1^{\text{MSSM}}}{2\pi} \ln \left( \frac{M_G}{M_Z} \right) + \delta_1$$

(20)

where

$$\delta_1 = \delta_i^0 + \delta_i^2 + \delta_i^3.$$ (21)

The constants $\delta_i^0$, $\delta_i^2$, $\delta_i^3$ represent the two loop running effects [6], the weak scale threshold corrections and the GUT scale threshold corrections, respectively. We have

$$\delta_i^0 = -\frac{1}{\pi} \sum_{j=1}^{3} b_{ij}^{\text{MSSM}} \log \left[ 1 - b_{ij}^{\text{MSSM}} \left( 3 - 8 \sin^2 \theta_W \right) \right]$$

(22)

where the matrix $b_{ij}^{\text{MSSM}}$ is given by [6]

$$b_{ij}^{\text{MSSM}} = \left( \begin{array}{ccc}
9 & 27 & 22 \\
19 & 20 & 25 \\
11 & 20 & 7
\end{array} \right).$$

(23)

The lightest thresholds are given by

$$\delta_i^2 = \frac{1}{\pi} \sum_j b_i^j \log \left( \frac{m_j}{M_Z} \right)$$

(24)

where the sum runs over all states at the weak scale including the top, $W$, Higgs and the supersymmetric spectrum. Finally the GUT scale threshold correction is given by

$$b_i^j = -\frac{1}{2\pi} \sum_c b_i^c \log \left( \frac{M_c}{M_G} \right).$$

(25)

In general the prediction for $\alpha_3(M_Z)$ is given by

$$\alpha_3^{-1}(M_Z) = \left( b_3 - b_1 \right) \alpha_2^{-1}(M_Z) - \left( \frac{b_3}{b_2} - b_1 \right) \alpha_1^{-1}(M_Z)$$

$$+ \left( \frac{b_3}{b_2} - b_1 \right) \delta_1 - \left( \frac{b_3}{b_2} - b_1 \right) \delta_2 + \delta_3$$

$$= \frac{12}{7} \alpha_2^{-1}(M_Z) - \frac{5}{7} \alpha_1^{-1}(M_Z) + \frac{1}{7} \left( 5\delta_1 - 12\delta_2 + 7\delta_3 \right)$$

$$\equiv (\alpha_3^{\text{LO}})^{-1} + \delta_1$$

(26)
where $b_i \equiv b_i^\text{MSSM}$, $\alpha_i^\text{LO}(M_Z)$ is the leading order one-loop result and $\delta_i \equiv \frac{1}{2}(\delta g_i - 12 \delta s_i + 7 \delta b_i)$. We find $\delta_i^2 \approx -0.82$ [9] and $\delta_i' \approx -0.04 + \frac{1}{2 \pi} \ln(\frac{\sqrt{s}}{M_{\text{GUT}}})$ where the first term takes into account the contribution of the $W_3$ top and the correction from switching from the $MS$ to $DR$ RG schemes and following [10])

$$T_{\text{SUSY}} = m_{\tilde{H}} \left( \frac{m_{\tilde{g}}}{m_{\tilde{g}}} \right)^{28/19} \left( \frac{m_{\tilde{H}}}{m_{\tilde{g}}} \right)^{3/19} \left( \frac{m_{\tilde{g}}}{m_{\tilde{H}}} \right)^{4/19}.$$

(27)

For a Higgsino mass $m_{\tilde{H}} = 400$ GeV, a Wino mass $m_{\tilde{g}} = 300$ GeV, a gluino mass $m_{\tilde{g}} = 900$ GeV and all other mass ratios of order 1, we find $\delta_i^2 \approx 0$. If we assume $\delta_i = 0$, we find the predicted value of $\alpha_i^\text{LO}(M_Z) = 0.0135$. In order to obtain a reasonable value of $\alpha_i^\text{LO}(M_Z)$ with only weak scale threshold corrections, we need $\delta_i^2 + \delta_i' = 0$ corresponding to a value of $T_{\text{SUSY}} \approx 5$ TeV. But this is very difficult considering the weak dependence $T_{\text{SUSY}}$ (equation (27)) has on squark and slepton masses. Thus in order to have $\delta_i \approx 0$ we need a GUT scale threshold correction

$$\delta_i^h \approx +0.94.$$

(28)

At the GUT scale we have

$$\alpha_i^{-1}(M_G) = \alpha_i^{-1} + \delta_i^h.$$

(29)

Define

$$\alpha_i^{-1} = \frac{1}{12} \alpha_i^{-1}(M_G) - 5 \alpha_i^{-1}(M_Z)$$

(30)

(or if the GUT scale is defined at the point where $\alpha_1$ and $\alpha_2$ intersect, then $\alpha_i^{-1} = \alpha_i^{-1}(M_G) = \alpha_i^{-1}(M_G)$). Hence, in order to fit the data, we need a GUT threshold correction

$$\epsilon_3 \equiv \frac{\alpha_3(M_G) - \alpha_3}{\alpha_3} = -\alpha_3 \delta_3^h \approx -4\%.$$

(31)

Note that this result depends implicitly on the assumption of universal soft SUSY breaking masses at the GUT scale, which directly affect the spectrum of SUSY particles at the weak scale. For example, if gaugino masses were not unified at $M_G$, and, in particular, gluinos were lighter than winos at the weak scale, then it is possible that, due to weak scale threshold corrections, a much smaller or even slightly positive threshold correction at the GUT scale would be consistent with gauge coupling unification [11].

2.1. Nucleon decay

Baryon number is necessarily violated in any GUT [12]. In $SU(5)$ nucleons decay via the exchange of gauge bosons with GUT scale masses, resulting in baryon number violating operators suppressed by $(1/M_G^2)$. The nucleon lifetime is calculable and given by $\tau_N \propto M_G^2/(\alpha_i^2 m_i^2)$. The dominant decay mode of the proton (and the baryon violating decay mode of the neutron), via gauge exchange, is $p \rightarrow e^+ \pi^0$ ($n \rightarrow e^- \pi^-$). In any simple gauge symmetry, with one universal GUT coupling and scale ($\alpha_G$, $M_G$), the nucleon lifetime from gauge exchange is calculable. Hence, the GUT scale may be directly observed via the extremely rare decay of the nucleon. In SUSY GUTs, the GUT scale is of order $3 \times 10^{16}$ GeV, as compared with the GUT scale in non-SUSY GUTs which is of order $10^{15}$ GeV. Hence the dimension 6 baryon violating operators are significantly suppressed in SUSY GUTs [5] with $\tau_p \sim 10^{34}$–$10^{38}$ yr.

However, in SUSY GUTs there are additional sources for baryon number violation—dimension 4 and 5 operators [13]. Although our notation does not change, when discussing SUSY GUTs all fields are implicitly bosonic superfields and the operators considered are the so-called $F$ terms which contain two fermionic components and the rest scalars or products of scalars. Within the context of $SU(5)$ the dimension 4 and 5 operators have the form $(10 \overline{5} \overline{5}) \supset (\bar{U} \bar{D} \bar{D}) + (Q L \bar{D}) + (\bar{E} \bar{L} \bar{L})$ and $(10 \overline{10} \overline{5}) \supset (Q Q L) + (\bar{U} \bar{U} \bar{D} \bar{E}) + B$ and $L$ conserving terms, respectively. The dimension 4 operators are renormalizable with dimensionless couplings; similar to Yukawa couplings. On the other hand, the dimension 5 operators have a dimensionful coupling of the order of $1/M_G$.

The dimension 4 operators violate baryon number or lepton number, respectively, but not both. The nucleon lifetime is extremely short if both types of dimension 4 operators are present in the low energy theory. However, both types can be eliminated by requiring $R$ parity. In $SU(5)$ the Higgs doublets reside in $5\overline{5}$, $\tilde{H}_U$ and $R$ parity distinguishes the $\tilde{H}_U$ (Higgs), $R$ parity [14] (or its cousin, family reflection symmetry (see Dimopoulos and Georgi [5] and DRW [15]) takes $F \rightarrow - F$, $H \rightarrow H$ with $F = \{10, \overline{5}\}$, $H = \{5\overline{5}, \overline{5}5\}$. This forbids the dimension 4 operator $(10 \overline{5} \overline{5})$, but allows the Yukawa couplings of the form $(10 \overline{5} \overline{5})$ and $(10 \overline{10} \overline{5})$. It also forbids the dimension 3, lepton number violating, operator $(\overline{5} \overline{5}5) \supset (L H_u)$ with a coefficient with dimensions of mass which, like the $\mu$ parameter, could be of order the weak scale and the dimension 5, baryon number violating, operator $(10 \overline{10} \overline{5} \overline{5}) \supset (Q Q H_u) + \cdots$.

Note that in the MSSM it is possible to retain $R$ parity violating operators at low energy as long as they violate either baryon number or lepton number only but not both. Such schemes are natural if one assumes a low energy symmetry, such as lepton number, baryon number or a baryon parity [16]. However, these symmetries cannot be embedded in a GUT. Thus, in a SUSY GUT, only $R$ parity can prevent unwanted dimension 4 operators. Hence, by naturalness arguments, $R$ parity must be a symmetry in the effective low energy theory of any SUSY GUT. This does not mean to say that $R$ parity is guaranteed to be satisfied in any GUT. A possible exception to this rule using constrained matter content which generates the effective $R$ parity violating operators in a GUT can be found in [17, 18] or for a review on $R$ parity violating interactions, see [19]. For example, in [18], the authors show how to obtain the effective $R$ parity violating operator $O^{1/2} = (\overline{5} \cdot \overline{5}) \overline{5} \overline{5} \cdot (10 \cdot \overline{10} \cdot \overline{10} \cdot \overline{10})$ where $\sum$ is an $SU(5)$ adjoint field and the subscripts $\overline{5}$, $\overline{10}$ indicate that the product of fields in parentheses have been projected into these $SU(5)$ directions. As a consequence the operator $O^{3/2} = O^{3/2}$, and out of $10 \overline{5} \overline{5}$ only the lepton number/R parity violating operator $QLD$ survives.
$R$ parity also distinguishes Higgs multiplets from ordinary families. In $SU(5)$, Higgs and quark/lepton multiplets have identical quantum numbers; while in $E(6)$, Higgs and families are unified within the fundamental 27 representation. Only in $SO(10)$ are Higgs and ordinary families distinguished by their gauge quantum numbers. Moreover, the $\mathbb{Z}_2$ center of $SO(10)$ distinguishes 10s from 16s and can be associated with $R$ parity [20].

Dimension 5 baryon number violating operators may be forbidden at tree level by symmetries in an $SU(5)$ model, etc. These symmetries are typically broken, however, by the VEVs responsible for the color-triplet Higgs masses. Consequently these dimension 5 operators are generically generated via color-triplet Higgsino exchange (figure 4). Hence, the color-triplet partners of Higgs doublets must necessarily obtain mass of order the GUT scale. (It is also important to note that Planck or string scale physics may independently generate dimension 5 operators, even without a GUT. These contributions must be suppressed by some underlying symmetry. See the discussion in section 4.4.8.)

The dominant decay modes from dimension 5 operators are $p \to K^+ \bar{\nu}$ ($n \to K^{\prime} \bar{\nu}$). This is due to a simple symmetry argument; the operators $(Q_i Q_j Q_k L_l)$, $(\bar{U}_i \bar{U}_j \bar{D}_k \bar{E}_l)$ (where $i$, $j$, $k$, $l = 1, 2, 3$ are family indices and color and weak indices are implicit) must be invariant under $SU(3)_C$ and $SU(2)_L$. As a result their color and weak doublet indices must be anti-symmetrized. However, since these operators are given by bosonic superfields, they must be totally symmetric under interchange of all indices. Thus the first operator vanishes for $i = j = k$ and the second vanishes for $i = j$. Hence a second or third generation particle must appear in the final state [15].

The dimension 5 operator contribution to proton decay requires a sparticle loop at the SUSY scale to reproduce an effective dimension 6 four Fermi operator for proton decay (see figure 5). The loop factor is of the form

$$(LF) \propto \frac{\lambda_t \lambda_t}{16\pi^2} \sqrt{\mu^2 + M_{1/2}^2} \frac{M_{16}^2}{m^2}$$

leading to a decay amplitude

$$A(p \to K^+ \bar{\nu}) \propto \frac{c_c}{M_{eff}^T} (LF),$$

where $M_{eff}^T$ is the effective Higgs color-triplet mass (see figure 4), $M_{1/2}$ is a universal gaugino mass and $m_{16}$ is a universal mass for squarks and sleptons. In any predictive SUSY GUT, the coefficients $c_c$ are 3 × 3 matrices related to (but not identical to) Yukawa matrices. Thus these tend to suppress the proton decay amplitude. However, this is typically not sufficient to be consistent with the experimental bounds on the proton lifetime. Thus it is also necessary to minimize the loop factor, $(LF)$. This can be accomplished by taking $\mu$, $M_{1/2}$ and $m_{16}$ large. Finally the effective Higgs color-triplet mass $M_{eff}^T$ must be maximized. With these caveats, it is possible to obtain rough theoretical bounds on the proton lifetime given by [21–23]

$$\tau_{p \to K^+ \bar{\nu}} \leq (\frac{1}{3} - 3) \times 10^{34} \text{ yr.}$$

2.1.1. Gauge coupling unification and proton decay. The dimension 5 operator (see figure 4) is given in terms of the matrices $c$ and an effective Higgs triplet mass by

$$\frac{1}{M_{eff}^T} \left[ Q_i \frac{1}{2} c q Q_j c q L + \bar{U} c d \bar{D} U c d E \right].$$

Note that $M_{eff}^T$ can be much greater than $M_G$ without fine-tuning and without having any particle with mass greater than the GUT scale. Consider a theory with two pairs of Higgs $S_i$ and $\tilde{S}_i$ with $i = 1, 2$ at the GUT scale with only $S_1$, $\tilde{S}_1$ coupling to quarks and leptons. Then we have

$$\frac{1}{M_{eff}^T} = (M_T^{-1})_{11}.$$  

If the Higgs color-triplet mass matrix is given by

$$M_T = \begin{pmatrix} 0 & M_G \\ M_G & X \end{pmatrix}$$

then we have

$$\frac{1}{M_{eff}^T} = \frac{X}{M_G^2}.$$  

Thus for $X \ll M_G$ we obtain $M_{eff}^T \gg M_G$.

We assume that the Higgs doublet mass matrix, on the other hand, is of the form

$$M_D = \begin{pmatrix} 0 & 0 \\ 0 & X \end{pmatrix}$$

with two light Higgs doublets. Note that this mechanism is natural in $SO(10)$ [24, 25] with a superpotential of the form

$$W \supset 10 45 10^I + X (10^I)^2$$

with only 10 coupling to quarks and leptons, $X$ is a gauge singlet and (45) = $(B - L) M_G$. 

Figure 4. The effective dimension 5 operator for proton decay.

Figure 5. The effective four Fermi operator for proton decay obtained by integrating out particles at the weak scale.
Recall $\epsilon_3 \equiv \frac{\alpha_3(M_3) - \alpha_3}{\alpha_3} \sim -4\%$. At one loop we find

$$\epsilon_3 = \epsilon_3^{\text{Higgs}} + \epsilon_3^{\text{GUT breaking}} + \ldots.$$  \hspace{1cm} (41)

Moreover

$$\epsilon_3^{\text{Higgs}} \equiv \frac{3\alpha_G}{5\pi} \ln \left( \frac{M_{\text{eff}}^2}{M_{\text{GUT}}^2} \right).$$  \hspace{1cm} (42)

See table 3 for the contribution to $\epsilon_3$ in a minimal SUSY SU(5), and in an SU(5) and SO(10) model with natural Higgs doublet–triplet splitting.

Recent Super-Kamiokande bounds on the proton lifetime severely constrain these dimension 6 and 5 operators with $t_{p-e^{-}\tau} > 1.0 \times 10^{34}$ yr (172.8 ktyr) at (90\% CL) [26], and $t_{p-\nu_{\tau}} > 2.3 \times 10^{34}$ yr (92 ktyr) at (90\% CL) based on the listed exposures [27]. These constraints are now sufficient to rule out minimal SUSY SU(5) [28]. The upper bound on the proton lifetime from these theories (particularly from dimension 5 operators) is approximately a factor of 5 above the experimental bounds. These bounds are also being pushed to their theoretical limits. Hence if SUSY GUTs are correct, then nucleon decay must be seen soon.

2.2. Yukawa coupling unification

2.2.1. Third generation, $b - \tau$ or $t - b - \tau$ unification. In SU(5), there are two independent renormalizable Yukawa interactions given by $\lambda_t \left( 10 10 5_{H} \right) + \lambda_t \left( 10 5 \bar{5}_{H} \right)$.

These contain the SM interactions $\lambda_t \left( Q \bar{u} \ H_u \right) + \lambda_b \left( Q \bar{d} H_d \right)$.

Hence, at the GUT scale we have the tree-level relation, $\lambda_b \approx \lambda_t \approx \lambda$. [30]. In SO(10) (or Pati–Salam) there is only one independent renormalizable Yukawa interaction given by $\lambda \left( 16 16 10_{H} \right)$ which gives the tree-level relation, $\lambda_t = \lambda_b = \lambda$. [31–33]. Note that in the discussion above we assume the minimal Higgs content with Higgs in $5, \bar{5}$ for SU(5) and 10 for SO(10). With Higgs in higher dimensional representations there are more possible Yukawa couplings [34–36].

In order to make contact with the data, one now renormalizes the top, bottom and $\tau$ Yukawa couplings, using two loop RG equations, from $M_G$ to $M_Z$. One then obtains the running quark masses $m_t(M_Z) = \lambda_t(M_Z) v_t$, $m_b(M_Z) = \lambda_b(M_Z) v_d$ and $m_{\tau}(M_Z) = \lambda_{\tau}(M_Z) v_d$ where $\langle H_u^0 \rangle = v_u = \sin \beta \; v / \sqrt{2}$, $\langle H_d^0 \rangle = v_d = \cos \beta \; v / \sqrt{2}$, $v_u/v_d = \tan \beta$ and $v \sim 246$ GeV is fixed by the Fermi constant, $G_F$.

Including one-loop threshold corrections at $M_Z$ and additional RG running, one finds the top, bottom and $\tau$ pole masses. In SUSY, $b - \tau$ unification has two possible solutions with $\tan \beta \sim 1$ or $40–50$. The small $\tan \beta$ solution is now disfavored by the LEP limit, $\tan \beta > 2.4$ [37]. However, this bound disappears if one takes $M_{\text{SUSY}} = 2 \text{TeV}$ and $m_t = 180 \text{GeV}$ [38]. The large $\tan \beta$ limit overlaps the SO(10) symmetry relation.

When $\tan \beta$ is large there are significant weak scale threshold corrections to down quark and charged lepton masses from either gluino and/or chargino loops [39]. Yukawa unification (consistent with low energy data) is only possible in a restricted region of SUSY parameter space with important consequences for SUSY searches [40].

Consider a minimal SO(10) SUSY model [MSO10SM] [40]. Quarks and leptons of one family reside in the 16-dimensional representation, while the two Higgs doublets of the MSSM reside in one 10-dimensional representation. For the third generation we assume the minimal Yukawa coupling term given by $\lambda \left( 16 16 10_{H} \right)$. On the other hand, for the first two generations and for their mixing with the third, we assume a hierarchical mass matrix structure due to effective higher dimensional operators. Hence the third generation Yukawa couplings satisfy $\lambda_t = \lambda_b = \lambda_\tau = \lambda_{\nu_s} = \lambda$. Note that these relations are only approximate once the off-diagonal elements of the Yukawa matrices are taken into account. If the Yukawa matrices are hierarchical then the corrections are typically $\lesssim 10\%$.

Soft SUSY breaking parameters are also consistent with SO(10) (with $1$) a universal gaugino mass $M_{1/2}$ (2) a universal squark and slepton mass $m_{16}$, (SO(10) does not require all sfermions to have the same mass, but it may be enforced by non-abelian family symmetries or possibly by the SUSY breaking mechanism) (3) a universal scalar Higgs mass $m_H$ and (4) a universal A parameter $A_0$. In addition we have the supersymmetric (soft SUSY breaking) Higgs mass parameters $\mu$ (B$_{12}$); $B_{12}$ may, as in the CMSSM, be exchanged for $\tan \beta$. Note that not all of these parameters are independent. Indeed, in order to fit the low energy electroweak data, including the third generation fermion masses, it has been shown that $A_0$, $m_{16}$, $m_{16}$ must satisfy the constraints $[40]$

$$A_0 \approx -2 m_{16}, \quad m_{10} \approx \sqrt{2} m_{16},$$  \hspace{1cm} (43)

$$m_{16} > 1.2 \text{ TeV}, \quad \mu, M_{1/2} \ll m_{16}.$$  \hspace{1cm} (44)

with $\tan \beta \approx 50$.  \hspace{1cm} (45)

This result has been confirmed by several independent analyses [41–43]. Note that different regions of parameter space consistent with Yukawa unification have also been discussed in [41,42,44]. Although the conditions (equations (43), (44)) are not obvious, it is, however, easy to see that equation (45) is simply a consequence of third generation Yukawa unification, since $m_t(m_t)/m_{16}(m_t) \sim \tan \beta$.

Finally, as a bonus, these same values of soft SUSY breaking parameters, with $m_{16} \gg \text{TeV}$, result in two very

\begin{table} [h]
\centering
\caption{Contribution to $\epsilon_3$ in three different GUT models.}
\begin{tabular}{|c|c|c|c|}
\hline
Model & $\epsilon_3^{\text{SU(5)}}$ & $\epsilon_3^{\text{SU(5)}}$ & $\epsilon_3^{\text{SO(10)}}$ \[22\] & $\epsilon_3^{\text{SO(10)}}$ \[23\] \\
\hline
$\epsilon_3^{\text{GUT breaking}}$ & 0 & 0 & 0 & 0 \\
$\epsilon_3^{\text{Higgs}}$ & $-7.7\%$ & $-7.7\%$ & $-4\%$ & $-4\%$ \\
$M_{10}^2$ (GeV) & $2 \times 10^{14}$ & $3 \times 10^{18}$ & $6 \times 10^{19}$ & $6 \times 10^{19}$ \\
\hline
\end{tabular}
\end{table}
interesting consequences. Firstly, it ‘naturally’ produces an inverted scalar mass hierarchy (ISMH) [45]. With an ISMH, squarks and sleptons of the first two generations obtain mass of order \(m_{16}\) at \(M_Z\). The stop, sbottom, and stau, on the other hand, have mass less than (or of the order of) a TeV. An ISMH has two virtues: (1) it preserves ‘naturalness’ (for values of \(m_{16}\) which are not too large), since only the third generation squarks and sleptons couple strongly to the Higgs, and (2) it ameliorates the SUSY CP and flavor problems, since these constraints on CP violating angles or flavor violating squark and slepton masses are strongest for the first two generations, yet they are suppressed as \(1/m^3_{16}\). For \(m_{16} > a\) few TeV, these constraints are weakened [46]. Secondly, Super-Kamiokande bounds on Dirac neutrinos with mass \(m_{\nu}\) are strongest for the first two generations, yet they are on CP violating angles or flavor violating squark and slepton masses are strongest for the first two generations, yet they are suppressed as \(1/m^3_{16}\). For \(m_{16} > a\) few TeV, these constraints are weakened [46].

2.2. Three families. Simple Yukawa unification is not possible for the first two generations of quarks and leptons. Consider the SU(5) GUT scale relation \(\lambda_3 = \lambda_\nu\). If extended to the first two generations one would have \(\lambda_3 = \lambda_\nu\), \(\lambda_d = \lambda_e\) which gives \(\lambda_3/\lambda_d = \lambda_\nu/\lambda_e\). The last relation is a renormalization group invariant and is thus satisfied at any scale. In particular, at the weak scale one obtains \(m_\nu/m_d = m_{\nu}/m_e\) which is in serious disagreement with the data with \(m_\nu/m_d \approx 20\) and \(m_{\nu}/m_e \approx 200\). An elegant solution to this problem was given by Georgi and Jarlskog [47]. Of course, a three-family model must also give the observed CKM mixing in the quark sector. Note that although there are typically many more parameters in the GUT theory above \(M_G\), it is possible to obtain effective low energy theories with many fewer parameters making strong predictions for quark and lepton masses.

It is important to note that grand unification alone is not sufficient to obtain predictive theories of fermion masses and mixing angles. Other ingredients are needed. In one approach additional global family symmetries are introduced (non-abelian family symmetries can significantly reduce the number of arbitrary parameters in the Yukawa matrices). These family symmetries constrain the set of effective higher dimensional fermion mass operators. In addition, sequential breaking of the family symmetry is correlated with the hierarchy of fermion masses. Three-family models exist which fit all the data, including neutrino masses and mixing [48]. In a completely separate approach for SO(10) models, the standard model Higgs bosons are contained in the higher dimensional Higgs representations including the 10, \(\overline{16}\) and/or \(\mathbf{126}\). Such theories have been shown to make predictions for neutrino masses and mixing angles [34–36]. Some simple patterns of fermion masses (see table 4) must be incorporated into any successful model.

### 2.3. Neutrino masses

Atmospheric and solar neutrino oscillations require neutrino masses. Using the three ‘sterile’ neutrinos \(\bar{\nu}\) with the Yukawa coupling \(\lambda_\nu (\bar{\nu} L H_u)\), one easily obtains three massive Dirac neutrinos with mass \(m_\nu = \lambda_\nu v_a\). Note that these

| Table 4. Patterns of masses and mixing. |
|----------------------------------------|
| \(\lambda_3 = \lambda_\nu = \lambda_e\) | \(SO(10)\) at \(M_G\) |
| \(\lambda_3 \approx 1/\lambda_\nu, \lambda_d \approx 3\lambda_\nu\) | \(SU(5)\) at \(M_G\) |
| \(m_\nu \approx 4 \cdot m_\mu, m_{\nu} \approx 3 \cdot m_\mu\) | \(M_H\) |
| \(\lambda_\nu \approx 0.5 (\lambda_\mu, \lambda_\tau)\) | \(SU(5)\) at \(M_G\) |
| \(\Delta_3^{\nu} \lambda \approx \Delta_3^{\nu} \lambda_3\) | \(SU(5)\) at \(M_G\) |
| \(V_{\nu} = \sqrt{m_{\nu}/m_\tau - 1} V_{\mu}/m_\mu\) | \(\overline{10}\) |
| \(V_{\nu} = \sqrt{m_{\nu}/m_\mu} \approx \sqrt{m_{\nu}/m_\tau}\) | \(\overline{10}\) |

‘sterile’ neutrinos are quite naturally identified with the right-handed neutrinos necessarily contained in complete families of \(SO(10)\) or Pati–Salam. However, in order to obtain a tau neutrino with mass of order 0.1 eV, one needs \(\lambda_\tau/\lambda_\nu \lesssim 10^{-10}\). The see-saw mechanism, on the other hand, can naturally explain such small neutrino masses [52, 53]. Since \(\bar{\nu}\) has no SU(3)color \(\times SU(2)_L \times U(1)_Y\) quantum numbers, there is no symmetry (other than global lepton number) which prevents the mass term \(\frac{1}{2} \bar{\nu} \lambda_\nu m_\nu\). Moreover, one might expect \(M \approx M_G\). Heavy ‘sterile’ neutrinos can be integrated out of the theory, defining an effective low energy theory with only light active Majorana neutrinos with the effective dimension 5 operator \(\frac{1}{2} (L H_u) \lambda_3 M^{-1} \lambda_\nu (L H_u)\). This then leads to a \(3 \times 3\) Majorana neutrino mass matrix \(m = m_\nu M^{-1} m_\nu\).

Atmospheric neutrino oscillations require neutrino masses with \(\Delta m^2_{\nu} \approx 3 \times 10^{-3}\) eV\(^2\) with maximal mixing, in the simplest two neutrino scenario. With hierarchical neutrino masses \(m_{\nu_t} = \sqrt{\Delta m^2_{\nu}} \approx 0.055\) eV. Moreover, via the see-saw mechanism \(m_{\nu_{\tau}} = m_{\nu_t} r^2/(3 M_H)\). Hence one finds \(M \approx 2 \times 10^{14}\) GeV; remarkably close to the GUT scale. Note that we have related the neutrino Yukawa coupling to the top quark Yukawa coupling \(\lambda_\nu = \lambda_t\) at \(M_G\) as given in \(SO(10)\) or \(SU(4) \times SU(2)_L \times SU(2)_R\). However, at low energies they are no longer equal and we have estimated this RG effect by \(\lambda_\nu (M_Z) = \lambda_t (M_Z)\).

2.4. \(SO(10)\) GUT with \([D_3 \times U(1)]\) family symmetry

A complete model for fermion masses was given in [54, 55]. Using a global \(\chi^2\) analysis, it has been shown that the model fits all fermion masses and mixing angles, including neutrinos, and a minimal set of precision electroweak observables. The model is consistent with lepton flavor violation and lepton electric dipole moment bounds. In two papers, [56, 57], the model was also tested by flavor violating processes in the B system.

The model is an \(SO(10)\) SUSY GUT with an additional \(D_3 \times [U(1) \times Z_2 \times Z_3]\) family symmetry. The symmetry group fixes the following structure for the superpotential:

\[ W = W_f + W_\nu \]

with

\[ W_f = 16_3 \times 10 \times 16_e \times 10 \times N_\nu \]

\[ + \bar{\nu}_a \left( M_X \bar{X}_{\nu} + 45 \frac{\phi_a}{M} 16_3 + 45 \frac{\phi_a}{M} 16_\nu + A 16_\alpha \right) \]

\[ W_\nu = \overline{10}(\lambda_2 N_\nu 16_e + \lambda_3 N_\nu 16_\nu) + \frac{1}{2} (S_3 N_\nu N_\nu + S_3 N_3 N_3) \]

\( (47) \)

\( (48) \)
The first two families of quarks and leptons are contained in the superfield $16_a$, $a = 1, 2$, which transforms under $SO(10) \times D_3$ as $(16,2_\lambda)$, whereas the third family in $16_3$ transforms as $(16,1_{1\bar{3}})$. The two MSSM Higgs doublets $H_u$ and $H_d$ are contained in a $10$. As can be seen from the first term on the right-hand side of (47), Yukawa unification $\lambda_a = \lambda_b = \lambda_c = \lambda_{1\bar{3}}$, at $M_G$ is obtained only for the third generation, which is directly coupled to the Higgs $10$ representation. This immediately implies large $\tan \beta \approx 50$ at low energies and constrains soft SUSY breaking parameters.

The effective Yukawa couplings of the first and second generation fermions are generated hierarchically via the Froggatt–Nielsen mechanism [58] as follows. Additional fields are introduced, i.e. the 45 which is an adjoint of $SO(10)$, the $SO(10)$ singlet flavon fields $\phi^i, \tilde{\phi}^i, A$ and the Froggatt–Nielsen states $\chi_a, \tilde{\chi}_a$. The latter transform as a $(10,2_\lambda)$ and a $(16,2_\lambda)$, respectively, and receive masses of $O(M_G)$ as $M_F$ acquires an $SO(10)$ breaking VEV. Once they are integrated out, they give rise to effective mass operators which, together with the VEVs of the flavon fields, create the Yukawa couplings for the first two generations. This mechanism breaks systematically the full flavor symmetry and produces the right mass hierarchies among the fermions.

Upon integrating out the FN states one obtains Yukawa matrices for up quarks, down quarks, charged leptons and neutrinos given by

$$Y_u = \begin{pmatrix}
0 & \epsilon' \rho & -\epsilon \xi \\
-\epsilon' \rho & \tilde{\epsilon} \rho & -\epsilon \\
\epsilon \xi & \epsilon & 1
\end{pmatrix} \lambda,$$

$$Y_d = \begin{pmatrix}
0 & \epsilon' & -\epsilon \xi \sigma \\
-\epsilon' \tilde{\epsilon} & -\epsilon \xi & \tilde{\epsilon} \\
\epsilon \xi & \epsilon & 1
\end{pmatrix} \lambda,$$

$$Y_e = \begin{pmatrix}
0 & -\epsilon' & 3 \epsilon \xi \\
0 & 3 \tilde{\epsilon} \xi & 3 \epsilon \\
-3 \epsilon \xi \sigma & -3 \epsilon \xi & 1
\end{pmatrix} \lambda,$$

$$Y_\nu = \begin{pmatrix}
0 & -\epsilon' \omega & \frac{1}{2} \epsilon \xi \xi \\
0 & 3 \tilde{\epsilon} \omega & \frac{1}{2} \epsilon \xi \xi \\
-3 \epsilon \xi \sigma & -3 \epsilon \xi & 1
\end{pmatrix} \lambda. \tag{49}$$

From equation (49) one can see that the flavor hierarchies in the Yukawa couplings are encoded in terms of the four complex parameters $\rho, \sigma, \tilde{\epsilon}, \xi$ and the additional real ones $\epsilon, \epsilon', \lambda$.

For neutrino masses one invokes the see-saw mechanism [52,53]. In particular, three $SO(10)$ singlet Majorana fermion fields $N_a, N_3 (a = 1, 2)$ are introduced via the contribution of $\frac{1}{2} (S_u N_a N_0 + S_3 N_3 N_0)$ to the superpotential (equation (48)). The mass term $\frac{1}{2} N M_N N$ is produced when the flavon fields acquire VEVs $(\tilde{S}_a) = M_{N_a}$ and $(S_3) = M_{N_3}$. Together with a $16$ Higgs one is allowed to introduce the interaction terms $16 [(\lambda_2 N_a, 16_a, \lambda_3 N_3, 16_3)$ (equation (48)]. This then generates a mixing matrix $V$ between the right-handed neutrinos and the additional singlets $(\nu^c V N)$, when the $16$ acquires an $SO(10)$ breaking VEV $(16)_{\nu^c} = v_{16}$. The resulting effective right-handed neutrino mass terms are given by

$$W_N = \bar{\nu} V N + \frac{1}{2} N M_N N \tag{50}$$

$$V = v_{16} \begin{pmatrix}
0 & \lambda_2 & 0 \\
\lambda_2 & 0 & 0 \\
0 & 0 & \lambda_3
\end{pmatrix}$$

$$M_N = \text{diag}(M_{N_1}, M_{N_2}, M_{N_3}). \tag{51}$$

Diagonalization leads to the effective right-handed neutrino Majorana mass

$$M_\nu = -V M_N V^T \equiv -\text{diag}(M_{R_1}, M_{R_2}, M_{R_3}). \tag{52}$$

By integrating out the EW singlets $\nu^c$ and $N$, which both receive GUT scale masses, one ends up with the light neutrino mass matrix at the EW scale given by the usual see-saw formula

$$\mathcal{M} = m_\nu M_R^{-1} m_\nu^T. \tag{53}$$

The model has a total of 24 arbitrary parameters, with all except $\tan \beta$ defined at the GUT scale (see table 5). Using a two loop RG analysis the theory is redefined at the weak scale. Then a $\chi^2$ function is constructed with low energy observables. In [55] fermion masses and mixing angles, a minimal set of precision electroweak observables and the branching ratio $BR(b \to s \gamma)$ were included in the $\chi^2$ function. Then predictions for lepton flavor violation, lepton electric dipole moments, Higgs mass and spurion masses were obtained. The $\chi^2$ fit was quite good. The light Higgs mass was always around 120 GeV. In the recent paper, [56], precision B physics observables were added. See tables 6 and 7 for the 28 low energy observables and table 8 for the 4 experimental bounds included in their analysis. The fits were not as good as before with a minimum $\chi^2 \sim 25$ obtained for large values of $m_{16} = 10$ TeV.

The dominant problem was due to constraints from the processes $B \to X_s \gamma, B \to X_s \ell^+ \ell^-$. The former process constrains the magnitude of the Wilson coefficient $C_7$ for the operator

$$O_7 = m_b \bar{s}_L \Sigma_{\mu \nu} b_R F_{\mu \nu} \tag{54}$$

with $C_7 \sim |C_7^{\text{SM}}|$, while the latter process is also sensitive to the sign of $C_7$. Note that the charged and neutral Higgs contributions to $BR(B \to X_s \ell^+ \ell^-)$ are strictly positive. While the sign of the chargino contribution, relative to the SM, is ruled by the following relation

$$C_7^2 \propto + \mu A_\tau \times \text{sign}(C_7^{\text{SM}}), \tag{55}$$

with a positive proportionality factor, so it is opposite to that of the SM one for $\mu > 0$ and $A_\tau < 0$. Hence it is possible for $C_7 \approx \pm |C_7^{\text{SM}}|$. Note that the experimental result for

| Sector | # Parameters |
|--------|--------------|
| Gauge  | 3 $a_G, M_{1L}, e$, |
| SUSY (GUT scale) | 5 $m_{16}, M_{1/2}, A_0, m_{H_d}, m_{H_u}$ |
| Textures | 11 $e, e', \lambda, \rho, \sigma, \epsilon, \xi$ |
| Neutrino | 3 $M_{R_1}, M_{R_2}, M_{R_3}$ |
| SUSY (EW scale) | 2 $\tan \beta, \mu$ |

Table 5. The 24 parameters defined at the GUT scale which are used to minimize $\chi^2$. |
Table 6. Flavor conserving observables used in the fit.
Dimensionful quantities are expressed in GeV, unless otherwise specified [56].

| Observable | Value($\sigma_{exp}$) | Observable | Value($\sigma_{exp}$) |
|------------|-----------------|------------|-----------------|
| $M_{H^\pm}$ | 80.403(29) | $M_t$ | 1.777(0) |
| $M_{H^0}$ | 91.1876(21) | $M_\chi$ | 0.10566(0) |
| $10^3 G_A$ | 1.16637(1) | $10^3 M_t$ | 5.111(0) |
| $\alpha_s(M_Z)$ | 137.036 | $\alpha_s(M_t)$ | 0.2258(14) |
| $\alpha_s(M_Z)$ | 0.1176(20) | $10^3 |V_{ud}|$ | 4.10(4) |
| $M_t$ | 170.9(1.8) | $10^3 |V_{ts}|$ | 4.16(7) |
| $m_b(m_t)$ | 4.20(7) | sin $\beta$ | 0.675(26) |
| $m_c(m_t)$ | 1.25(9) | $10^3 \Delta m^2_{21}$ [eV$^2$] | 2.6(0.2) |
| $m_{\ell}(2\text{ GeV})$ | 0.095(25) | $10^3 \Delta m^2_{31}$ [eV$^2$] | 7.90(0.28) |
| $m_d(2\text{ GeV})$ | 0.0005(2) | sin$^2$ $\theta_{12}$ | 0.852(32) |
| $m_s(2\text{ GeV})$ | 0.00225(75) | sin$^2$ $\theta_{13}$ | 0.996(18) |

Table 7. FC observables used in the fit [56].

| Observable | Value($\sigma_{exp}$) | Observable | Value($\sigma_{exp}$) |
|------------|-----------------|------------|-----------------|
| $10^3 \tau_F$ | 2.2299(10)(252) | $\Delta M_f/\Delta M_d$ | 350.0(4)(3.6) |
| $10^4 \text{ BR}(B \rightarrow X_s \gamma)$ | 3.55(26)(46) | $10^4 \text{ BR}(B \rightarrow X_{sT} \ell^- \ell^-)$ | 1.60(51)(40) |
| $10^4 \text{ BR}(B \rightarrow \tau^+ \nu \ell^-)$ | 1.31(48)(9) | BR($B_s \rightarrow \mu^+ \mu^-$) | <1.0 $\times$ 10$^{-7}$ |

Table 8. Mass bounds used in the fit [56].

| Observable | Lower bound (GeV) |
|------------|-----------------|
| $M_{h^0}$ | 114.4 |
| $m_t$ | 60 |
| $m_{\mu^+}$ | 104 |
| $m_{\tilde{g}}$ | 195 |

In a recent analysis [57], it was shown that better $\chi^2$ can be obtained by allowing for a 20% correction to Yukawa unification. Note that this analysis only included Yukawa terms necessary for GUT symmetry breaking and Higgs doublet–triplet splitting [60]. Let there be a single adjoint field, $A$, and two pairs of spinors, $C + \tilde{C}$ and $C + \tilde{C}$. The complete Higgs superpotential is assumed to have the form

$$W = W_A + W_C + W_{ACC} + (T_1 A T_2 + S T_2^2).$$

(56)

The precise forms of $W_A$ and $W_C$ do not matter, as long as $W_A$ gives $(A)$ the Dimopoulos–Wilczek form, and $W_C$ makes the VEVs of $C$ and $\tilde{C}$ point in the $SU(5)$-singlet direction. For specificity we will take $W_A = \frac{1}{16} \text{Tr} A^3 + \frac{1}{3} P_A (\text{Tr} A^2 + M_A^2) + f(P_A)$, where $P_A$ is a singlet, $f$ is an arbitrary polynomial and $M \sim M_G$. (It would be possible, also, to have simply $m \text{ Tr} A^2$, instead of the two terms containing $P_A$, however, derived from a more fundamental theory, such as string theory. In order to resolve these difficulties, it becomes natural to discuss grand unified theories in higher spatial dimensions.)

Finally, an analysis of dark matter for this model has been performed with good fits to WMAP data [59]. The authors of [43] also analyze dark matter in the context of the minimal $SO(10)$ model with Yukawa unification. They have difficulty in fitting WMAP data. We believe this is because they do not adjust the CP odd Higgs mass to allow for dark matter annihilation on the resonance.

2.5. Problems of 4D GUTs

There are two aesthetic (perhaps more fundamental) problems concerning 4D GUTs. They have to do with the complicated sectors necessary for GUT symmetry breaking and Higgs doublet–triplet splitting. These sectors are sufficiently complicated that it is difficult to imagine that they may be
explicit mass terms for adjoint fields may be difficult to obtain in string theory.) We take $W_C = X(\bar{C}C - P_C)$, where $X$ and $P_C$ are singlets, and $(P_C) \sim M_G$. The crucial term that couples the spinor and adjoint sectors together has the form

$$W_{ACC} = \bar{C} \left( \left( \frac{P}{M_P} \right) A + Z \right) C + \bar{C} \left( \left( \frac{\bar{P}}{M_P} \right) A + \bar{Z} \right) C'$$

(57)

where $Z, \bar{Z}, P$ and $\bar{P}$ are singlets. $(P)$ and $(\bar{P})$ are assumed to be of order $M_G$. The critical point is that the VEVs of the primed spinor fields will vanish, and therefore the terms in equation (3) will not make a destabilizing contribution to $-\kappa A^t = \partial W/\partial A$. This is the essence of the mechanism.

$W$ contains several singlets $(P_C, P \bar{\mathcal{P}}$ and $S)$ that are supposed to acquire VEVs of order $M_G$, but which are left undetermined at tree level by the terms so far written down. These VEVs may arise radiatively when SUSY breaks, or may be fixed at tree level by additional terms in $W$.

In $SU(5)$ the construction which gives natural Higgs doublet–triplet splitting requires the $SU(5)$ representations 75, 50, 50 and a superpotential of the form [22, 61]

$$W \supset 75^3 + M 75^2 + 5 75 50 + 5 H 75 50 + 50 50 50.$$  

(58)

The 50, 50 contain Higgs triplets but no Higgs doublets. Thus when the 75 obtains an $SU(5)$ breaking VEV, the color triplets obtain mass but the Higgs doublets remain massless.

### 3. Orbifold GUTs

#### 3.1. GUTs on a circle

As the first example of an orbifold GUT consider a pure $SO(3)$ gauge theory in five dimensions [62]. The gauge field is

$$A_M = A_M^a T^a \quad a = 1, 2, 3 \quad M, N = [0, 1, 2, 3, 5].$$

(59)

The gauge field strength is given by

$$F_{MN} = F_{M}^{\mu N} = \partial_M A_N - \partial_N A_M + i [A_M, A_N],$$

(60) where $T^a$ are $SO(3)$ generators. The Lagrangian is

$$L_5 = -\frac{1}{4g_5^2} \mathcal{F}(F_{MN}F^{MN})$$

(61) and we have $\mathcal{F}(T^a T^b) \equiv k\delta^{ab}$. The inverse gauge coupling squared has mass dimensions 1.

Let us first compactify the theory on $M_4 \times S^1$ with coordinates $(x^\mu, y)$ and $y = [0, 2\pi R]$. The theory is invariant under the local gauge transformation

$$A_M(x^\mu, y) \rightarrow U A_M(x^\mu, y) U^\dagger - iU \partial_M U^\dagger$$

(62)

Consider the possibility $\delta S A_M \equiv 0$. We have

$$F_{\mu 5} = \partial_\mu A_5 + i [A_\mu, A_5] \equiv D_\mu A_5.$$  

(63)

We can then define

$$\Phi \equiv A_5 \sqrt{2\pi R/g_5} = A_5/g$$

(64) where $g_5 = \sqrt{2\pi R/g}$ and $g$ is the dimensionless 4D gauge coupling. The 5D Lagrangian reduces to the Lagrangian for a 4D $SO(3)$ gauge theory with massless scalar matter in the adjoint representation, i.e.

$$L_5 = \frac{1}{2 \pi R} \left[ -\frac{1}{4g_5^2k} \mathcal{F}(F_{\mu\nu}F^{\mu\nu}) + \frac{1}{2k} \mathcal{F}(D_\mu \Phi D^\mu \Phi) \right].$$

(65)

In general we have the mode expansion

$$A_M(x_\mu, y) = \sum_n \left[ a_n^M \cos n \frac{y}{R} + b_n^M \sin n \frac{y}{R} \right]$$

(66) where only the cosine modes with $n = 0$ have zero mass. Otherwise the 5D Laplacian $\partial_M \partial^M = \partial_\mu \partial^\mu + \partial_5 \partial^5$ leads to Kaluza–Klein (KK) modes with effective 4D mass

$$m_n^2 = n^2 \frac{1}{R^2}.$$  

(67)

#### 3.2. Fermions in 5D

The Dirac algebra in 5D is given in terms of the 4 $\gamma$ matrices $\gamma_\mu$. $\gamma_M = 0, 1, 2, 3, 5$ satisfying $\{\gamma_M, \gamma_N\} = 2g_{MN}$. A four component massless Dirac spinor $\psi(x^\mu, y)$ satisfies the Dirac equation

$$i\gamma_M \partial^M \psi = 0 = i(\gamma_\mu \partial^\mu - \gamma_5 \partial_5) \psi$$

(68) with $\gamma_5 = i(\gamma_0 \gamma_1 \gamma_2 \gamma_3 \gamma_5)$. In 4D the four component Dirac spinor decomposes into two Weyl spinors with

$$\Psi = \begin{pmatrix} \psi_1 \\ i\sigma_2 \psi_2 \end{pmatrix} = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$$

(69)

where $\psi_{1,2}$ are two left-handed Weyl spinors. In general, we obtain the normal mode expansion for the fifth direction given by

$$\psi_{L,R} = \sum_n \left( a_n(x) \cos n \frac{y}{R} + b_n(x) \sin n \frac{y}{R} \right).$$

(70)

If we couple this 5D fermion to a local gauge theory, the theory is necessarily vector-like; coupling identically to both $\psi_{L,R}$.

We can obtain a chiral theory in 4D with the following parity operation

$$P : \psi(x_\mu, y) \rightarrow \psi(x_\mu, -y) = P \psi(x_\mu, y)$$

(71) with $P = i\gamma_5$ (figure 6). We then have

$$\psi_L \sim \cos n \frac{y}{R} \quad \psi_R \sim \sin n \frac{y}{R}.$$  

(72)

#### 3.3. GUTs on an orbi-circle

Let us briefly review the geometric picture of orbifold GUT models compactified on an orbi-circle $S^1/\mathbb{Z}_2$. The circle $S^1 \equiv \mathbb{R}^1/T$ where $T$ is the action of translations by $2\pi R$. [References: Rep. Prog. Phys. 74 (2011) 036901]
All fields $\Phi$ are thus periodic functions of $y$ (up to a finite gauge transformation), i.e.

$$T : \Phi(x_{\mu}, y) \rightarrow \Phi(x_{\mu}, y + 2\pi R) = T \Phi(x_{\mu}, y)$$  \hspace{1cm} (73)

where $T \in SO(3)$ satisfies $T^2 = 1$. This corresponds to the translation $T$ being realized non-trivially by a degree-2 Wilson line (i.e. background gauge field—$\langle A_k \rangle \neq 0$ with $T \equiv \exp(i (A_k) dy)$). Hence the space group of $S^1/\mathbb{Z}_2$ is composed of two actions: a translation, $T : y \rightarrow y + 2\pi R$, and a space reversal, $P : y \rightarrow -y$. There are two (conjugacy) classes of fixed points, $y = (2n)\pi R$ and $(2n + 1)\pi R$, where $n \in \mathbb{Z}$.

The space-group multiplication rules imply $TPT = P$, so we can replace the translation by a composite $\mathbb{Z}_2$ action $P' = PT : y \rightarrow -y + 2\pi R$. The orbic-circle $S^1/\mathbb{Z}_2$ is equivalent to an $\mathbb{R}/(\mathbb{Z}_2 \times \mathbb{Z}_2)$ orbifold, whose fundamental domain is the interval $[0, \pi R]$, and the two ends $y = 0$ and $y = \pi R$ are fixed points of the $\mathbb{Z}_2$ and $\mathbb{Z}_2$ actions, respectively.

A generic 5D field $\Phi$ has the following space-time properties under the $\mathbb{Z}_2$ and $\mathbb{Z}_2$ orbifoldings (the 4D space-time coordinates are suppressed):

$$\mathcal{P} : \Phi(y) \rightarrow \Phi(-y) = \mathcal{P}\Phi(y)$$

$$\mathcal{P}' : \Phi(y) \rightarrow \Phi(-y + 2\pi R) = \mathcal{P}'\Phi(y)$$  \hspace{1cm} (74)

where $\mathcal{P}, \mathcal{P}' = PT = \pm$ are orbifold parities acting on the field $\Phi$ in the appropriate group representation and where it is assumed that $[P, T] = 0$. The four combinations of orbifold parities give four types of states, with wave functions

$$\zeta_m(++) \sim \cos(my/R)$$

$$\zeta_m(+) \sim \cos[(2m + 1)y/2R]$$

$$\zeta_m(-+) \sim \sin[(2m + 1)y/2R]$$

$$\zeta_m(--) \sim \sin(my/R)$$  \hspace{1cm} (75)

where $m \in \mathbb{Z}$. The corresponding KK towers have mass

$$M_{KK} = \begin{cases} 
\frac{m}{R}, & \text{for } (P P') = (++) \\
\frac{(2m + 1)y}{2R}, & \text{for } (P P') = (+-) \text{ and } (-+) \\
\frac{m + 1}{R}, & \text{for } (P P') = (--) 
\end{cases}$$  \hspace{1cm} (76)

Note that only the $\Phi_{++}$ field possesses a massless zero mode.

For example, consider the Wilson line $T = \exp(i(\pi T^3)) = \text{diag}(-1, -1, 1)$. Let $A_3(y)$ ($A_3(y)$) have parities $P = (+)$, respectively. Then only $A_3^{++}$ has orbifold parity $++$ and $A_3^{+}$ has orbifold parity $-$. Note that $A_3^{-(y)} = -A_3^{++}(y) + \frac{i}{R}$. Define the fields

$$W^\pm = \frac{1}{\sqrt{2}}(A^1 \mp iA^2)$$  \hspace{1cm} (77)

with $T^\pm = \frac{1}{\sqrt{2}}(T^1 \pm iT^2)$ and $[T^3, T^\pm] = \pm T^\pm$. Then $W_{++}^{\pm}$ [W_{--}^{\pm}$] have orbifold parity $(+-)$ $([-+])$, respectively. Thus the $SO(3)$ gauge group is broken to $SO(2) \approx U(1)$ in 4D. The local gauge parameters preserve the $(P, T)$ parity/holonomy, i.e.

$$\theta^3(x_{\mu}, y) = \theta^3_m(x_{\mu})\zeta_m(++)$$

$$\theta^{1/2}(x_{\mu}, y) = \theta^{1/2}_m(x_{\mu})\zeta_m(+-)$$  \hspace{1cm} (78)

Therefore, $SO(3)$ is not the symmetry at $y = \pi R$.

### 3.4. A supersymmetric $SU(5)$ orbifold GUT

Consider the 5D orbifold GUT model of [63]. The model has an $SU(5)$ symmetry broken by orbifold parities to the SM gauge group in 4D. The compactification scale $M_c = R^{-1}$ is assumed to be much less than the cut-off scale.

The gauge field is a 5D vector multiplet $V = (A_{\mu}, \lambda, \lambda', \sigma)$, where $A_{\mu}$, $\sigma$ (and their fermionic partners $\lambda, \lambda'$) are in the adjoint representation (24) of $SU(5)$. This multiplet consists of one 4D $N = 1$ supersymmetric vector multiplet $V = (A_{\mu}, \lambda)$ and one 4D chiral multiplet $\Sigma = ((\sigma + iA_{\mu}/\sqrt{2}), \lambda')$. We also add two 5D hypermultiplets containing the Higgs doublets, $H_1 = (H_3, H^{-}_3)$, $H_2 = (\bar{H}_3, \bar{H}^{-}_3)$. The 5D gravitino $\Psi_M = (\psi_{m1}^I, \psi_{m2}^I, \psi_{m3}^I)$ decomposes into two 4D gravitini $\psi_{m1}^{\pm}$ and two dilatini $\psi_{m2}^{\pm}$, $\psi_{m3}^{\pm}$. To be consistent with the 5D supersymmetry transformations one can assign positive parities to $\psi_{m1}^{\pm}, \psi_{m2}^{\pm}$ and $\psi_{m3}^{\pm}$ and negative parities to $\psi_{m1}^{-}, \psi_{m2}^{-}, \psi_{m3}^{-}$; this assignment partially breaks $N = 2$ to $N = 1$ in 4D.

The orbifold parities for various states in the vector and hyper-multiplets are chosen as follows [63] (where we have decomposed all the fields into SM irreducible representations and under $SU(5)$ we have taken $P = (++++, P' = (- - - +))$.

| States | $P$ | $P'$ | States | $P$ | $P'$ |
|--------|-----|-----|--------|-----|-----|
| $V(8, 1, 0)$ | $+$ | $+$ | $\Sigma(8, 1, 0)$ | $-$ | $-$ |
| $V(1, 3, 0)$ | $+$ | $+$ | $\Sigma(1, 3, 0)$ | $-$ | $-$ |
| $V(1, 1, 0)$ | $+$ | $+$ | $\Sigma(1, 1, 0)$ | $-$ | $-$ |
| $V(\bar{3}, 2, 5/3)$ | $-$ | $+$ | $\Sigma(3, 2, -5/3)$ | $-$ | $-$ |
| $V(3, 2, -5/3)$ | $-$ | $+$ | $\Sigma(3, 2, 5/3)$ | $-$ | $-$ |
| $T(3, 1, -2/3)$ | $+$ | $+$ | $T^{\pm}(\bar{3}, 1, 2/3)$ | $-$ | $-$ |
| $H(1, 2, +1)$ | $+$ | $+$ | $H^{\pm}(1, 2, -1)$ | $-$ | $-$ |
| $\bar{T}(\bar{3}, 1, +2/3)$ | $-$ | $+$ | $\bar{T}^{\pm}(3, 1, -2/3)$ | $-$ | $-$ |
| $H(1, 2, -1)$ | $+$ | $+$ | $H^{\pm}(1, 2, +1)$ | $-$ | $-$ |

(79)

We see the fields supported at the orbifold fixed points $y = 0$ and $\pi R$ have parities $P = +$ and $P' = +$, respectively. They form complete representations under the $SU(5)$ and SM groups; the corresponding fixed points are called $SU(5)$ and SM ‘branes’. In a 4D effective theory one would integrate out all the massive states, leaving only massless modes of the $P = P' = +$ states. With the above choices of orbifold parities, the SM gauge fields and the $H$ and $\bar{H}$ chiral multiplet are the only surviving states in 4D. We thus have an $N = 1$ SUSY
in 4D. In addition, the $T + ar{T}$ and $T^c + ar{T}^c$ color-triplet states are projected out, solving the doublet–triplet splitting problem that plagues conventional 4D GUTs.

3.5. Gauge coupling unification

We follow the field theoretical analysis in [64] (see also [65, 66]). It has been shown there that the correction to a generic gauge coupling due to a tower of KK states with masses $M_{KK} = m/R$ is

$$\alpha^{-1}(\Lambda) = \alpha^{-1}(\mu_0) + \frac{b}{4\pi} \int_{\tau_{KK}^{-1}} \frac{dt}{t} \theta_3 \left( \frac{it}{\pi R^2} \right)$$

(80)

where the integration is over the Schwinger parameter $t$, $\mu_0$ and $\Lambda$ are the IR and UV cut-offs, and $r = \pi/4$ is a numerical factor. $\theta_3$ is the Jacobi theta function, $\theta_3(t) = \sum_{m=-\infty}^{\infty} e^{\pi m^2 r}$, representing the summation over KK states.

For our $S^1/Z_2$ orbifold there is one modification in the calculation. There are four sets of KK towers, with mass $M_{KK} = m/R$ (for $P = P^* = +$), $(m + 1)/R$ (for $P = P^* = -$) and $(m + 1/2)/R$ (for $P = +$, $P^* = -$ and $P = -$, $P^* = +$), where $m \geq 0$. The summations over KK states give, respectively, $\frac{1}{\pi} \theta_3(i/\pi R^2)$ for the first two cases and $\frac{1}{\pi} \theta_2(i/\pi R^2)$ for the last two (where $\theta_2(t) = \sum_{m=-\infty}^{\infty} e^{2\pi m^2 r}$), and we have separated out the zero modes in the $P = P^* = +$ case.

Tracing the renormalization group evolution from low energy scales, we are first in the realm of the MSSM, and the beta function coefficients are $\delta_{MSSM} = (-\frac{3}{33}, -1, 3)$. The next energy threshold is the compactification scale $M_c$. From this scale to the cut-off scale, $M_\star$, we have the four sets of KK states.

Collecting these facts, and using $\theta_2(i/\pi R^2) \simeq \frac{1}{\sqrt{\pi}} R$ for $t/R^2 \ll 1$, we find the RG equations

$$\alpha_i^{-1}(M_Z) = \alpha_i^{-1} - \frac{\delta_{MSSM}}{2\pi} \log \frac{M_Z}{M_c} + \frac{1}{4\pi} (b^+_{\tau i} + b^-_{\tau i}) \log \frac{M_s}{M_c}$$

$$- \frac{b^0}{2\pi} \left( \frac{M_c}{M_c} - 1 \right) + \delta_i^2 + \delta_i^\prime$$

(81)

for $i = 1, 2, 3$, where $\alpha_i^{-1} = \frac{\pi r a_i}{\delta_i}$ and we have taken the cut-off scales, $\mu_0 = M_c = \frac{1}{\pi}$ and $\Lambda = M_\star$. (Note that this 5D orbifold GUT is a non-renormalizable theory with a cut-off. In string theory, the cut-off will be replaced by the physical string scale, $M_{string}$.) $b^0 = \sum_{p = \pm} b^{0}_{p R}$, so in fact it is the beta function coefficient of the orbifold GUT gauge group, $G = SU(5)$. The beta function coefficients in the last two terms have an $N = 2$ nature, since the massive KK states enjoy a larger supersymmetry. In general we have $b^0 = 2C(2G) - 2N(hyper T_R)$. The first term (in equation (81)) on the right is the 5D gauge coupling defined at the cut-off scale, the second term accounts for the KK modes running in the MSSM from the weak scale to the cut-off, the third and fourth terms take into account the KK modes in loops above the compactification scale and the last two terms account for the corrections due to two loop RG running and weak scale threshold corrections.

Figure 7. The differences $\delta_i = 2\pi(1/\alpha_i - 1/\alpha)$ are plotted as a function of energy scale $\mu$. The threshold correction $\delta_\ast$ defined in the 4D GUT scale is used to fix the threshold correction in the 5D orbifold GUT.

$$\alpha_i^{-1}(4D) \leftrightarrow \alpha_i^{-1} - \frac{b^0}{2\pi} \left( \frac{M_c}{M_c} - 1 \right)$$

$$\delta_i^b(4D) \leftrightarrow \frac{1}{4\pi} (b^+_{\tau i} + b^-_{\tau i}) \log \frac{M_s}{M_c}$$

$$- \frac{b^0}{2\pi} \log \frac{M_c}{M_c}$$

(82)

Thus in 5D the GUT scale threshold corrections determine the ratio $M_c/M_c$ (note that the second term in equation (82) does not contribute to $\delta_i^\ast$). For $SU(5)$ we have $b^+_{\tau i} + b^-_{\tau i} = (-6/5, 2, 6)$ and given $\delta_i^b$ (equation (28)) we have

$$\delta_i^b = \frac{12}{28\pi} \log \frac{M_c}{M_c}$$

or

$$M_c/M_c \approx 0.94$$

(83)

$$\frac{M_c}{M_c} \approx 10^3.$$  

(84)

If the GUT scale is defined at the point where $\alpha_1 = \alpha_2$, then we have $\delta_1 = \delta_2 = \log \frac{M_c}{M_c} \approx 10^3$. In 5D orbifold GUTs, nothing in particular happens at the 4D GUT scale. However, since the gauge bosons affecting the dimension 6 operators for proton decay obtain their mass at the compactification scale, it is important to realize that the compactification scale is typically lower than the 4D GUT scale and the cut-off is higher (see figure 7).

3.6. Quarks and leptons in 5D orbifold GUTs

Quarks and lepton fields can be put on either of the orbifold branes or in the 5D bulk. If they are placed on the $SU(5)$ brane at $y = 0$, then they come in complete $SU(5)$ multiplets. As a consequence a coupling of the type

$$W \supset \int d^2 \theta \int dy \delta(y) H \bar{H} 10 5$$

(85)

will lead to bottom-tau Yukawa unification. This relation is good for the third generation and so it suggests that the third family should reside on the $SU(5)$ brane. Since this
relation does not work for the first two families, they might be placed in the bulk or on the SM brane at \( y = \pi R \). Without further discussion of quark and lepton masses (see \([9, 67-69]\)) for complete \( SU(5) \) or \( SO(10) \) orbifold GUT models, let us consider proton decay in orbifold GUTs.

3.7. Proton decay

3.7.1. Dimension 6 operators. The interactions contributing to proton decay are those between the so-called \( X \) gauge bosons \( A_{\mu}^{(n)} = Y_{\mu} \) in \( V(+) \) (where \( A_{\mu}^{(n)} = \eta \) is the 5D gauge boson with quantum numbers \((\bar{3}, 2, +5/3)\) under \( SU(3) \times SU(2) \times U(1) \), \( a \) and \( i \) are color and \( SU(2) \) indices, respectively) and the \( N = 1 \) chiral multiplets on the \( SU(5) \) brane at \( y = 0 \). Assuming all quarks and leptons reside on this brane we obtain the \( \Delta B \neq 0 \) interactions given by

\[
S_{\Delta B \neq 0} = -\frac{g_5}{\sqrt{2}} \int d^4x A_{\mu}^{(n+1)}(x, y) J_{\mu}^{(n)}(x, 0) + \text{h.c.}.
\]  

(86)

The currents \( J_{\mu}^{(n)} \) are given by

\[
J_{\mu}^{(n)} = 6 \epsilon_{abc} \epsilon_{ij}(\bar{u}_b \gamma^\mu u_c) \bar{d}_i \gamma^\mu d_j + q_{ai} \bar{\sigma}^\mu \bar{q}_i \bar{\sigma}^\mu q_j \bar{\sigma}^\mu \bar{d}_i - \bar{\bar{\sigma}}^\mu \bar{\sigma}^\mu (\bar{d}_i a_i).
\]

(87)

Upon integrating out the \( X \) gauge bosons we obtain the effective Lagrangian for proton decay

\[
\mathcal{L} = \frac{g_5^2}{2 M_X^2} \sum_{i,j} \left[ (q_{ai} \bar{\sigma}^\mu \bar{d}_i)(\bar{\bar{\sigma}}^\mu \bar{d}_j) + (q_{bi} \bar{\sigma}^\mu \bar{q}_i) (\bar{\bar{\sigma}}^\mu \bar{q}_j) \right].
\]

(88)

where all fermions are weak interaction eigenstates and \( i, j, k = 1, 2, 3 \) are family indices. The dimensionless quantity

\[
g_5 \equiv g_5\frac{1}{\sqrt{2\pi R}}
\]

is the 4D gauge coupling of the gauge bosons zero modes. The combination

\[
M_X = \frac{M_X}{\pi}
\]

(90)

proportional to the compactification scale

\[
M_c = \frac{1}{R}
\]

(91)

is an effective gauge vector boson mass arising from the sum over all the KK levels:

\[
\sum_{n=0}^{\infty} \frac{4}{(2n+1)^2 M_c^2} = \frac{1}{M_X^2}.
\]

(92)

Before one can evaluate the proton decay rate one must first rotate the quark and lepton fields to a mass eigenstate basis. This will bring in both left- and right-handed quark and lepton mixing angles. However, since the compactification scale is typically lower than the 4D GUT scale, it is clear that proton decay via dimension 6 operators is likely to be enhanced.

3.7.2. Dimension 5 operators. The dimension 5 operators for proton decay result from integrating out color-triplet Higgs fermions. However, in this simplest \( SU(5) \) 5D model the color-triplet mass is of the form \([70]\)

\[
W = \int d^7\theta d\gamma (T(-))^c \partial_\gamma T(+) + \bar{T}(-)T^c \partial_\gamma \bar{T}(+))
\]

(93)

where a sum over massive KK modes is understood. Since only \( T, \bar{T} \) couple directly to quarks and leptons, no dimension 5 operators are obtained when integrating out the color-triplet Higgs fermions.

3.7.3. Dimension 4 baryon and lepton violating operators. If the theory is constructed with an \( R \) parity or family reflection symmetry, then no such operators will be generated.

4. String theory

As mentioned in the introduction, there are several limiting forms of string theory, known as type I, type II A and B, \( E_8 \otimes E_8 \) and \( SO(32) \) (or more precisely \( spin(32)/\mathbb{Z}_2 \)) heterotic and F and M theories. The first five are perturbative string theories defined in terms of quantum superstrings propagating in \( 9 + 1 \) space-time dimensions\(^5\). Type I is a theory of open and closed strings with the open strings coupled to gauge fields taking values in \( spin(32)/\mathbb{Z}_2 \). Type II is a theory of closed strings with abelian gauge symmetry in IIA and no gauge symmetry in IIB. Heterotic strings are a combination of superstrings for right-moving excitations along the string and bosonic strings for left-moving excitations. The right-movers travel in a target space of 10 space-time dimensions, while the left-movers travel in 26 space-time dimensions. However, 16 spatial dimensions are compactified on a \( E_8 \otimes E_8 \) or \( spin(32)/\mathbb{Z}_2 \) lattice. This compactification leads to the two possible gauge groups of the heterotic string.

Type I and II strings also include non-perturbative excitations called \( D_p \)-branes. \( D_p \)-branes are hyper-surfaces in \( p \) spatial dimensions on which open strings can end. \( N \) \( D_p \)-branes, piled on top of each other, support massless \( U(n) \) gauge excitations. And when \( D_p \) branes intersect, massless chiral matter appears. The low energy effective field theories for the perturbative strings correspond to \( N = 1 \) supergravity in 10 space-time dimensions. In addition, type I and II low energy theories include massless gauge matter living on \( D \)-branes and chiral matter living at the intersection of two \( D \)-branes. M and F theories are only defined in terms of non-perturbative effective field theories in \( 10 + 2 \) and \( 10 + 1 \) space-time dimensions, respectively. F theory is essentially a type IIB string with the extra internal two dimensions corresponding to the complex type IIB string coupling. \( D_p \)-branes in the type IIB picture occur at the position of singularities in the string coupling in the F theory description. M theory has no fundamental string limit. It is defined as \( N = 1 \) supergravity in 11 dimensions which reduces to \( N = 8 \) supergravity in \( 5 \) A superstring theory denotes a string theory which has 2D worldsheet supersymmetry. Such theories can lead to space-time supersymmetry, depending on the vacuum.
four dimensions. This is the maximal supergravity theory! $D_0$-branes of Type IIB form a tower of excitations which is interpreted as the 11th dimension. In M theory, gauge fields live on 3D singular sub-manifolds and chiral matter lives on 0D singular points in the internal seven dimensions. There is also a purely quantum mechanical formulation of M theory in terms of a quantum theory of matrices, the so-called Matrix theory. However, this formulation is even more difficult to deal with.

All of these limiting forms are related by so-called $S$, $T$, or $U$ dualities [71–74]. The dilaton $S$ is a measure of the string coupling, thus $S \rightarrow \frac{1}{T}$ corresponds to strong–weak coupling duality. Which means to say, one limiting string theory calculated in the weak coupling limit is dual to a different limiting string theory, calculated in the strong coupling limit. Examples of $S$ dual theories are type I - $SO(32)$ heterotic, type IIB–type IIB, type IIA–E$_8 \otimes$ E$_8$ heterotic (via the intermediate step of M theory). The modulus $T$ is a measure of the volume associated with the extra six internal dimensions of a perturbative string. Thus $T$ duality is the equality of two different limiting theories with one compactified on a space of large volume and the other compactified on a space of small volume. Finally, if one theory compactified on a space of large (or small) volume is equivalent to another theory at strong (or weak) coupling, they are called $U$ dual. If the two theories are the same, then the two theories are said to be self-dual. $T$ duality, unlike $S$ or $U$ duality, can be understood perturbatively. Type IIA–type IIB and $E_8 \otimes E_8$ heterotic–$SO(32)$ heterotic are $T$ dual. Combining $S$ and $T$ dualities, one can show that type IIB–$E_8 \otimes E_8$ heterotic are dual. In addition, in this way one can also obtain a non-perturbative definition of type IIB, i.e. F theory.

String model building refers to choosing a particular limiting string theory and compactifying the extra (6 for type I, II and heterotic, 7 for M theory and 8 for F theory) dimensions. If the compactification manifold is flat, then the string spectrum and interactions can be calculated perturbatively (with the addition of open strings attached to $D_p$-brane states in type I and II theories). The $D_p$-branes are massive and will distort the flat background metric of the internal dimensions. In order to include the back reaction on the metric, one studies the effective low energy supergravity limit of the string. In this case one compactifies on an internal manifold which preserves at least one supersymmetry in four dimensions. This requires compactifying on a Calabi–Yau 3-fold for type I, II and heterotic strings, a $G_2$ manifold for M theory and a Calabi–Yau 4-fold for F theory. In order to complete the picture of string model building, it is necessary to also mention string theories defined solely in four dimensions in terms of tensor products of 2D conformal field theories (another term for strings). Such theories go under the name of free fermionic heterotic strings [75–77] or Gepner models [78]. Note that the free fermionic construction can be put into one-to-one correspondence with orbifold constructions compactified, at the first step, on toroidal manifolds with radii determined as rational ratios of the string scale. For arbitrary radii, as allowed in bosonic constructions, one requires additional Thirring interactions included in the fermionic construction. Hence, the free fermionic construction describes isolated points in the bosonic moduli space.

Now consider $E_6 \otimes E_8$ heterotic–F theory duality. It turns out that F theory compactified on $K_3$ is dual to the $E_8 \otimes E_8$ heterotic string compactified on a 2D torus, $T^2$. In addition F theory on a CY$_3$ is dual to the $E_8 \otimes E_8$ heterotic string compactified on $K_3$. Finally, there is a duality between F theory on CY$_1$ and the $E_8 \otimes E_8$ heterotic string on CY$_3$.

The string landscape refers to all possible supersymmetry breaking solutions of the string in four space-time dimensions. This, however, is a particularly ill-defined concept, since in string theory the coordinates of space-time are themselves dynamical variables and hence the space-time background geometry is quantum mechanical. In this sense it may be said that in string theory space-time itself is an emergent concept. Moreover, since energy and momentum are derived quantities in Lagrangian systems with space and time translation invariance, one must first define the Lagrangian for an effective field theory describing the modes of a string. String field theory exists describing all the string excitations. But this theory is quite difficult to handle. Thus in most situations, when one speaks of the string vacua or the string landscape one is most likely referring to the ground state for the massless degrees of freedom in terms of the supergravity limit of the string. As discussed in the introduction, the string landscape is immense. It has been estimated that the landscape of SUSY breaking string vacua numbers of order $10^{500}$ [79]. For a recent animated history of the subject, see [80].

4.1. Random searches for the MSSM in the string landscape

The literature is replete with searches over the string landscape. Among these have been random searches in the string landscape looking for features common to the MSSM. In particular, vacua with $N = 1$ supersymmetry (SUSY), the standard model gauge group and three families of quarks and leptons. These random searches have for the most part shown that the MSSM is an extremely rare point in the string landscape. For example, searches in Type II intersecting D-brane models [81] have found nothing looking like the MSSM in 10$^4$ tries. Searches in Gepner orientifolds have been a bit more successful finding one MSSM-like model for every 10 000 tries [82]. Even searches in the heterotic string, using the free fermionic construction, have shown that the MSSM is a very rare point in the string landscape [83]. The bottom line: if you want to find the MSSM, then a random search is not the way to go. In fact, MSSM-like models have been found in type II D-brane vacua, but by directed searches [84] and not random ones. For a recent discussion of random searches in the string landscape, see [85].

4.2. String model building

In recent years, some progress has been made in finding MSSM-like theories starting from different points in the string landscape [86–106], i.e. free fermionic, orbifold or smooth Calabi–Yau constructions of the heterotic string, intersecting (and local) D-brane constructions in the type II string, and $M$ or $F$ theory constructions. Much of this progress has benefited
from the requirement of an intermediate grand unified gauge
symmetry which naturally delivers the standard model particle
spectrum. I will discuss a few recent results.

String model building is typically a two stage process.
At the first stage one searches for an MSSM-like theory with
N = 1 supersymmetry. One is interested in the Yukawa and
gauge coupling constants which, at this stage, are functions
of moduli, i.e. SM singlet fields with no potential. These
moduli also determine the masses of any vector-like exotic
partners of ordinary particles have not been discovered. Hence
they must be heavier than their SM namesakes. We discuss
supersymmetry breaking and moduli stabilization in section 6.

4.3. Heterotic constructions on smooth manifolds

Bouchard et al [107] have obtained an SU(5) GUT model on a
CY3 with the following properties. They have three families
of quarks and leptons, and one or two pairs of Higgs doublets.
They accomplish GUT symmetry breaking and Higgs doublet–
triplet splitting via a Wilson line in the weak hypercharge
direction. The CY3 is defined by a double elliptic fibration, i.e.
two tori whose radii change as the tori move over the surface
of a sphere (see figure 8).

In addition, they obtained a non-trivial up Yukawa matrix
given by [108]

\[
\lambda_u = \begin{pmatrix}
a & b & c \\
b & d & e \\
c & e & 0
\end{pmatrix}.
\]

The parameters a, ..., e are functions of the moduli. The
down and charged lepton Yukawa matrices are, however, zero
and would require non-perturbative effects to change this.

Braun et al [86] obtained an SU(5) \( \otimes U(1)_{B-L} \) GUT
model on a CY3. GUT symmetry breaking and Higgs doublet–
triplet splitting is accomplished via a Wilson line in the weak
hypercharge direction. The low energy theory contains three
families of quarks and leptons, one pair of Higgs doublets and
the standard model gauge symmetry plus the additional
U(1)_{B-L}. The latter forbids R parity violating operators and
is only spontaneously broken near the weak scale via right-
handed sneutrino VEVs. Hence in this theory the tau neutrino
mixes with neutralinos and is Majorana. However, in the
simplest scenario, the electron and muon neutrinos are pure
Dirac. Moreover, they are light due to the assumption of very
small Yukawa couplings. The phenomenology of this model
has been discussed in [109].

4.4. Heterotic string orbifolds and orbifold GUTs

Early work on orbifold constructions of the heterotic string
was started over 20 years ago [110, 111, 113–115]. The first
complete MSSM model from the \( \mathbb{E}_8 \otimes \mathbb{E}_8 \) heterotic string
was obtained using the free fermionic construction by Faraggi
et al [87, 116]. The authors impose an intermediate \( SO(10) \)
SUSY GUT.

Progress has been made recently [100–102, 104, 106,
117–121]. In a ‘mini-landscape’ search of the \( E(8) \times E(8) \)
heterotic landscape [101] 223 models with three families,
Higgs doublets and only vector-like exotics were found out of a
total of order 30000 models or approximately 1 in 100 models
searched looked like the MSSM! We called this a ‘fertile patch’
in the heterotic landscape. Let me describe this focused search
in more detail.

4.4.1. Phenomenological guidelines. We use the following
guidelines when searching for ‘realistic’ string models
[100, 101]. We want to

(i) preserve gauge coupling unification;
(ii) keep low energy SUSY as a solution to the gauge hierarchy
problem, i.e. why is \( M_Z \ll M_G \);
(iii) put quarks and leptons in 16 of \( SO(10) \);
(iv) put Higgs in 10, thus quarks and leptons are distinguished
from Higgs by their \( SO(10) \) quantum numbers;
(v) preserve GUT relations for third family Yukawa
couplings;
(vi) use the fact that GUTs accommodate a ‘natural’ see-saw
scale \( \mathcal{O}(M_G) \);
(vii) use intuition derived from orbifold GUT constructions,
[120, 121] and
(viii) use local GUTs to enforce family structure [102, 104, 122].

It is the last two guidelines which are novel and characterize
our approach.

4.4.2. \( \mathbb{E}_8 \times \mathbb{E}_8 \) heterotic string compactified on \( \mathbb{Z}_3 \times \mathbb{Z}_2 \)
6D orbifold. There are many reviews and books on string
theory. I cannot go into great detail here, so I will confine
my discussion to some basic points. We start with the
10D heterotic string theory, consisting of a 26D left-moving
bosonic string and a 10D right-moving superstring. Modular
invariance requires the momenta of the internal left-moving
bosonic degrees of freedom (16 of them) to lie in a 16D
Euclidean even self-dual lattice, we choose to be the \( \mathbb{E}_8 \times \mathbb{E}_8 \)
root lattice. Note that for an orthonormal basis, the \( \mathbb{E}_8 \)
root lattice consists of the following vectors, \( (n_1, n_2, \ldots, n_8) \) and
\( (n_1 + \frac{1}{2}, n_2 + \frac{1}{2}, \ldots, n_8 + \frac{1}{2}) \), where \( n_1, n_2, \ldots n_8 \) are integers
and \( \sum_{i=1}^{8} n_i = 0 \mod 2 \).

Figure 8. Calabi–Yau 3-fold defined in terms of a double elliptic
fibration. The line connecting the tori to the 2-sphere represents the
fibration.
4.4.3. Heterotic string compactified on $T^6/\mathbb{Z}_6$. We first compactify the theory on a 6D torus defined by the space-group action of translations on $\mathbb{R}^6$ giving the torus $T^6$ in terms of a factorizable Lie algebra lattice $G_2 \oplus SU(3) \oplus SO(4)$ (see figure 9). Then we mod out by the $\mathbb{Z}_6$ action on the three complex compactified coordinates given by $Z' \mapsto e^{2\pi i n} Z'$, $n = 1, 2, 3$, where $v_6 = \frac{1}{6}(1, 2, -3)$ is the twist vector, and $r_1 = (1, 0, 0, 0)$, $r_2 = (0, 1, 0, 0)$, $r_3 = (0, 0, 1, 0)$. (Together with $r_4 = (0, 0, 0, 1)$, they form the set of positive weights of the $8_s$ representation of $SO(8)$, the little group in 10D. $\pm r_4$ represent the two uncompactified dimensions in the light-cone gauge. Their space-time fermionic partners have weights $r = (\pm \frac{1}{2}, \pm \frac{1}{2}, \pm \frac{1}{2}, \pm \frac{1}{2})$ with even numbers of positive signs; they are in the $8_s$ representation of $SO(8)$. In this notation, the fourth component of $v_6$ is zero.)

The $\mathbb{Z}_6$ orbifold is equivalent to a $Z_2 \times Z_3$ orbifold, where the two twist vectors are $v_2 = 3v_6 + \frac{1}{2}(1, 0, -1)$ and $v_3 = 2v_6 + \frac{1}{4}(1, -1, 0)$. The $Z_2$ and $Z_3$ sub-orbifold twists have the $SU(3)$ and $SO(4)$ planes as their fixed torii. In abelian symmetric orbifolds, gauge embeddings of the point group elements and lattice translations are realized by shifts of the momentum vectors, $P$, in the $E_8 \times E_8$ root lattice. [114, 115, 123–125], i.e. $P \mapsto P + kV + lW$, where $k, l$ are some integers, and $V$ and $W$ are known as the gauge twists and Wilson lines [113]. These embeddings are subject to modular invariance requirements [110–112]. (Note that the $E_8$ root lattice is given by the set of states $P = \{n_1, n_2, \ldots, n_8\}$, $\{n_1 + \frac{1}{2}, n_2 + \frac{1}{2}, \ldots, n_8 + \frac{1}{2}\}$ satisfying $n_i \in \mathbb{Z}$, $\sum_{i=1}^{8} n_i = 2Z_+$.) The Wilson lines are also required to be consistent with the action of the point group. In the $\mathbb{Z}_6$ model, there are at most three consistent Wilson lines [126], one of degree 3 ($W_3$), along the SU(3) lattice, and two of degree 2 ($W_2, W_2$), along the SO(4) lattice.

The $\mathbb{Z}_6$ model has three untwisted sectors ($U_i$, $i = 1, 2, 3$) and five twisted sectors ($T_i$, $i = 1, 2, 3, 4, 5$). (The $T_i$ and $T_{6-k}$ sectors are CFT conjugates of each other.) The twisted sectors split further into sub-sectors when discrete Wilson lines are present. In the SU(3) and SO(4) directions, we can label these sub-sectors by their winding numbers, $n_3 = 0, 1, 2$ and $n_2, n'_2 = 0, 1$, respectively. In the $G_2$ direction, where both the $Z_2$ and $Z_3$ sub-orbifold twists act, the situation is more complicated. There are four $Z_2$ fixed points in the $G_2$ plane. Not all of them are invariant under the $Z_2$ twist, in fact three of them are transformed into each other. Thus for the $T_3$ twisted-sector states one needs to find linear combinations of these fixed-point states such that they have definite eigenvalues, $\gamma = 1$ (with multiplicity 2), $e^{2\pi i /3}$, or $e^{4\pi i /3}$, under the orbifold twist [126, 127] (see figure 10). Similarly, for the $T_{5,4}$ twisted-sector states, $\gamma = 1$ (with multiplicity 2) and $-1$ (the fixed points of the $T_{2,4}$ twisted sectors in the $G_2$ torus are shown in figure 11). The $T_i$ twisted-sector states have only one fixed point in the $G_2$ plane, thus $\gamma = 1$ (see figure 12).

These eigenvalues $\gamma$ provide another piece of information to differentiate twisted sub-sectors.

Massless states in 4D string models consist of those momentum vectors $P$ and $r$ (are in the $SO(8)$ weight lattice) which satisfy the following mass-shell equations [110, 111, 114, 115, 123–125]:

$$\alpha' \frac{m_R^2}{2} = N_R^k + \frac{1}{2}[r + k\hat{v}]^2 + a_R^k = 0$$

$$\alpha' \frac{m_L^2}{2} = N_L^k + \frac{1}{2}[P + kX]^2 + a_L^k = 0$$

where $\alpha'$ is the Regge slope, $N_R^k$ and $N_L^k$ are (fractional) numbers of the right- and left-moving (bosonic) oscillators, $X = V + n_3 W_3 + n_2 W_2 + n'_2 W'_2$, and $a_R^k, a_L^k$ are the normal ordering constants

$$a_R^k = -\frac{1}{2} + \frac{1}{2} \sum_{i=1}^{3} [\tilde{k} v_i (1 - |\tilde{k} v_i|)]$$

$$a_L^k = -\frac{1}{2} + \frac{1}{2} \sum_{i=1}^{3} [\tilde{k} v_i (1 - |\tilde{k} v_i|)]$$

with $\tilde{k} v_i = \text{mod}(k v_i, 1)$.

These states are subject to a generalized Gliozzi–Scherk–Olive (GSO) projection $P = \frac{1}{2} \sum_{\alpha'} \Delta^\alpha [114, 115, 123–125]$. For the simple case of the $k$th twisted sector ($k = 0$ for the untwisted sectors) with no Wilson lines ($n_3 = n_2 = n'_2 = 0$) we have

$$\Delta = \gamma \phi \exp[i\pi [(2P + kX) \cdot X - (2r + k\hat{v}) \cdot v]]$$
where \( \phi \) are phases from bosonic oscillators. However, in the \( \mathbb{Z}_3 \) model, the GSO projector must be modified for the untwisted sector and \( T_{2,4} \), \( T_3 \) twisted-sector states in the presence of Wilson lines \cite{121}. The Wilson lines split each twisted sector into sub-sectors and there must be additional projections with respect to these sub-sectors. This modification in the projector gives the following projection conditions:

\[
P \cdot V - r_i \cdot v = \mathbb{Z} (i = 1, 2, 3)
\]

\[
P \cdot W_3 \quad P \cdot W_2 \quad P \cdot W'_2 = \mathbb{Z},
\]

(99)

for the untwisted-sector states, and

\[
T_{2,4} : P \cdot W_2 \quad P \cdot W'_2 = \mathbb{Z} \quad T_3 : P \cdot W_3 = \mathbb{Z}
\]

(100)

for the \( T_{2,3,4} \) sector states (since twists of these sectors have fixed torii). There is no additional condition for the \( T_1 \) sector states.

4.4.4. An orbifold GUT—heterotic string dictionary. We first implement the \( \mathbb{Z}_3 \) sub-orbifold twist, which acts only on the \( G_2 \) and \( SU(3) \) lattices. The resulting model is a 6D gauge theory with \( N = 2 \) hypermultiplet matter, from the untwisted and \( T_{2,4} \) twisted sectors. This 6D theory is our starting point to reproduce the orbifold GUT models. The next step is to implement the \( \mathbb{Z}_2 \) sub-orbifold twist. The geometry of the extra dimensions closely resembles that of 6D orbifold GUTs. The \( SO(4) \) lattice has four \( \mathbb{Z}_2 \) fixed points at \( 0, \pi R, \pi' R \) and \( \pi (R + R') \), where \( R \) and \( R' \) are on the \( e_5 \) and \( e_6 \) axes, respectively, of the lattice (see figures 10 and 12). When one varies the modulus parameter of the \( SO(4) \) lattice such that the length of one axis \( R \) is much larger than the other \( R' \) and the string length scale \( (\ell_s) \) the lattice effectively becomes the \( S^1 / \mathbb{Z}_2 \) orbicircle in the 5D orbifold GUT, and the two fixed points at 0 and \( \pi R \) have degree-2 degeneracies. Furthermore, one may identify the states in the intermediate \( \mathbb{Z}_3 \) model, i.e. those of the untwisted and \( T_{2,4} \) twisted sectors, as bulk states in the orbifold GUT.

Space-time supersymmetry and GUT breaking in string models work exactly as in the orbifold GUT models. First consider supersymmetry breaking. In the field theory, there are two gravitini in 4D, coming from the 5D (or 6D) gravitino. Only one linear combination is consistent with the space reversal, \( y \rightarrow -y \); this breaks the \( N^c = 2 \) supersymmetry to that of \( N = 1 \). In string theory, the space-time supersymmetry currents are represented by those half-integral \( SO(8) \) momenta. (Together with \( r_4 = (0, 0, 0, 1) \), they form the set of positive weights of the \( 8 \), representation of the \( SO(8) \), the little group in 10D. \( \pm r_4 \) represent the two uncompactified dimensions in the light-cone gauge. Their space-time fermionic partners have weights \( r = (\pm \frac{1}{2}, \pm \frac{1}{2}, \pm \frac{1}{2}, \pm \frac{1}{2}) \) with even numbers of positive signs; they are in the \( 8 \), representation of \( SO(8) \). In this notation, the fourth component of \( v_0 \) is zero.) The \( \mathbb{Z}_3 \) and \( \mathbb{Z}_2 \) projections remove all but two of them, \( r = \pm \frac{1}{2} (1, 1, 1, 1) \); this gives \( N = 1 \) supersymmetry in 4D.

Now consider GUT symmetry breaking. As usual, the \( \mathbb{Z}_2 \) orbifold twist and the translational symmetry of the \( SO(4) \) lattice are realized in the gauge degrees of freedom by degree-2 gauge twists and Wilson lines, respectively. To mimic the 5D orbifold GUT example, we impose only one degree 2 Wilson line, \( W_2 \), along the long direction of the \( SO(4) \) lattice, \( R \). (Wilson lines can be used to reduce the number of chiral families. In all our models, we find it is sufficient to get three-generation models with two Wilson lines, one of degree 2 and one of degree 3. Note, however, that with two Wilson lines in the \( SO(4) \) lattice we can break \( SO(10) \) directly to \( SU(3) \times SU(2) \times U(1) \times U(1) \) (see for example, \cite{128}).) The gauge embeddings generally break the 5D/6D (bulk) gauge group further down to its subgroups, and the symmetry breaking works exactly as in the orbifold GUT models. This can clearly be seen from the following string theoretical realizations of the orbifold parities:

\[
P = p e^{2 \pi i [P \cdot V_2 - r_2]}, \quad P' = p' e^{2 \pi i [P \cdot (V_2 + W_2) - r_2]}
\]

(101)

where \( V_2 = 3V_6 \), and \( p = \gamma \phi \) can be identified with intrinsic parities in the field theory language. (For gauge and untwisted-sector states, \( p \) are trivial. For non-oscillator states in the \( T_{2,4} \) twisted sectors, \( p = \gamma \) are the eigenvalues of the \( G_2 \)-plane fixed points under the \( \mathbb{Z}_2 \) twist. Note that \( p = + \) and \( - \) states have multiplicities 2 and 1, respectively since the corresponding fixed points in the \( G_2 \) plane are 2 and 1.) Since \( 2(\mathbf{W} \cdot V_2 - r \cdot v_2) = 2 \mathbf{W} \cdot W_2 = \mathbb{Z} \), by properties of the \( E_8 \times E_8 \) and \( SO(8) \) lattices, thus \( p^2 = P^2 = 1 \), and equation (101) provides a representation of the orbifold parities. From the string theory point of view, \( P = P' = + \) are nothing but the projection conditions, \( \Delta = 1 \), for the untwisted and \( T_{2,4} \) twisted-sector states (see equations (98), (99) and (100)).

To reaffirm this identification, we compare the masses of KK excitations derived from string theory with that of orbifold GUTs. The coordinates of the \( SO(4) \) lattice are untwisted under the \( \mathbb{Z}_3 \) action, so their mode expansions are the same as that of toroidal coordinates. Concentrating on the \( R \) direction, the bosonic coordinate is \( X_{1,R} = X_{1,R} + p_{1,R} (\tau \pm \sigma) + \) oscillator terms, with \( p_{1,L}, p_{1,R} \) given by

\[
p_{1,L} = \frac{m}{2R} + \left( 1 - \frac{1}{4} |W_2|^2 \right) \frac{n_2 R}{\ell_s^2} + \frac{P\cdot W_2}{2R}
\]

\[
p_{1,R} = \frac{m}{2R} - \frac{2n_2 R}{\ell_s^2} + \frac{P\cdot W_2}{2R}
\]

(102)

where \( m \) (or \( n_2 \)) are KK levels (winding numbers). The \( \mathbb{Z}_2 \) action maps \( m \) to \( -m \), \( n_2 \) to \( -n_2 \) and \( W_2 \) to \( -W_2 \), so physical states must contain linear combinations, \( \{ m, n_2 \} \pm \{ -m, -n_2 \} \); the eigenvalues \( \pm 1 \) correspond to the first \( \mathbb{Z}_2 \) parity, \( p \), of orbifold GUT models. The second orbifold parity, \( p' \), induces a non-trivial degree-2 Wilson line; it shifts the KK level by \( m \rightarrow m + P \cdot W_2 \). Since \( 2W_2 \) is a vector of the (integral) \( E_8 \times E_8 \) lattice, the shift must be an integer or half-integer. When \( R \gg \ell_s \), the winding modes and the KK modes in the smaller dimension of \( SO(4) \) decouple. Equation (102) then gives four types of KK excitations, reproducing the field theoretical mass formula in equation (76).
4.4.5. **MSSM with R parity.** In this section we discuss just one ‘benchmark’ model (model 1) obtained via a ‘mini-landscape’ search [100, 101, 105] of the $E_8 \times E_8$ heterotic string compactified on the $Z_6$ orbifold [101]. (For earlier work on MSSM models from $Z_6$ orbifolds of the heterotic string, see [102, 104]. For reviews, see [129, 130]).

The model is defined by the shifts and Wilson lines

\[
V = \left( \begin{array}{c}
\frac{1}{3}, -\frac{1}{2}, -\frac{1}{2}, 0, 0, 0, 0, 0 \\
\frac{1}{2}, -\frac{1}{6}, 1, 1, 1, -\frac{1}{2}, -\frac{1}{2}, 1
\end{array} \right)
\times \left( \begin{array}{c}
\frac{1}{2}, -\frac{1}{6}, 1, 1, 1, -\frac{1}{2}, -\frac{1}{2}, 1 \\
0, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, 0, 0, 0
\end{array} \right)
\times \left( \begin{array}{c}
4, -3, -\frac{7}{2}, -4, -3, -\frac{7}{2}, -\frac{9}{2}, \frac{2}{2} \\
-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}
\end{array} \right)
\times \left( \begin{array}{c}
\frac{1}{3}, 0, 0, \frac{2}{3}, 0, \frac{5}{3}, -2, 0
\end{array} \right).
\]  

(103)

A possible second order 2 Wilson line is set to zero.

The shift $V$ is defined to satisfy two criteria.

- The first criterion is the existence of a local $SO(10)$ GUT at the $T_1$ fixed points at $x_6 = 0$ in the $SO(4)$ torus (figure 12).

\[
P \cdot V = Z; \ P \in SO(10) \text{ momentum lattice.}
\]  

(104)

Since the $T_1$ twisted sector has no invariant torus and only one Wilson line along the $x_6$ direction, all states located at these two fixed points must come in complete $SO(10)$ multiplets. (For more discussion on local GUTs, see [102, 122]).

- The second criterion is that two massless spinor representations of $SO(10)$ are located at the $x_6 = 0$ fixed points.

Hence, the two complete families on the local $SO(10)$ GUT fixed points give us an excellent starting point to find the MSSM. The Higgs doublets and third family of quarks and leptons must then come from elsewhere.

Let us now discuss the effective 5D orbifold GUT [131]. Consider the orbifold $(T^2)^3 / Z_3$ plus the Wilson line $W_3$ in the $SU_3$ torus. The $Z_3$ twist does not act on the $SO_4$ torus, see figure 11. As a consequence of embedding the $Z_3$ twist as a shift in the $E_8 \times E_8$ group lattice and taking into account the $W_3$ Wilson line, the first $E_8$ is broken to $SU(6)$. This gives the effective 5D orbifold gauge multiplet contained in the $N = 1$ vector field $V$. In addition we find the massless states $\Sigma = 35$, and $20 + 20'$ and $18 (6 + 6')$ in the 6D untwisted sector and $T_2$. $T_2$ twisted sectors. Together these form a complete $N = 2$ gauge multiplet $(V + \Sigma)$ (see equations (109) and (110)) and a $20 + 18 (6 + 6')$ dimensional hypermultiplets\(^6\). In fact the massless states in this sector can all be viewed as ‘bulk’ states moving around in a large 5D space-time.

\(^6\) Note that in four dimensions massless chiral adjoints cannot be obtained at level 1 Kac-Moody algebras. Clearly this is not true in 5D.

Now consider the $Z_2$ twist and the Wilson line $W_2$ along the $x_6$ axis in the $SO_4$ torus. The action of the $Z_2$ twist breaks the gauge group to $SU(5)$, while $W_2$ breaks $SU(5)$ further to the SM gauge group $SU(3) C \times SU(2)_L \times U(1)_Y$ (the SM is represented by the red states in equation (109)).

Let us focus on those MSSM states located in the bulk. From two of the pairs of $N = 1$ chiral multiplets $6 + 6'$, which decompose as

\[2 \times (6 + 6') \supset [(1, 2)_{1/2}^{-1} + (3, 1)_{-2/3, 1/3}]
\]

(107)

we obtain the third family $b^c$ and lepton doublet, $l$. The rest of the third family comes from the $10 + 10'$ of $SU(5)$ contained in the $20 + 20'$ of $SU(6)$, in the untwisted sector.

Now consider the Higgs bosons. The bulk gauge symmetry is $SU(6)$. Under $SU(5) \times U(1)$, the adjoint decomposes as

\[35 \rightarrow 24_0 + 5_{1/2} + 5'_{-1/2} + 1_0.
\]

(108)

The Higgs doublets are represented by the blue states in equation (110). Thus the MSSM Higgs sector emerges from the breaking of the $SU(6)$ adjoint by the orbifold and the model satisfies the property of ‘gauge–Higgs unification’.

To summarize, in models with gauge–Higgs unification, the Higgs multiplets come from the 5D vector multiplet $(V, \Sigma)$, both in the adjoint representation of $SU(6)$. $V$ is the 4D gauge multiplet and the 4D chiral multiplet $\Sigma$ contains the Higgs doublets. These states transform as follows under the orbifold parities ($P, P'$):

\[
\begin{align*}
V : & \quad (++) (++) (++) (--) (--) (--) \quad (+) \\
& \quad (++) (++) (++) (--) (--) (--) \quad (+)
\end{align*}
\]

(109)

\[
\begin{align*}
\Phi : & \quad (--) (--) (--) (--) (--) (--) \quad (+) \\
& \quad (--) (--) (--) (--) (--) (--) \quad (+)
\end{align*}
\]

(110)

Hence, we have obtained doublet–triplet splitting via orbifolding.

4.4.6. **$D_4$ family symmetry.** Consider the $Z_2$ fixed points. We have four fixed points, separated into an $SU(5)$ and SM invariant pair by the $W_2$ Wilson line (see figure 13). We find two complete families, one on each of the $SO(10)$ fixed points and a small set of vector-like exotics (with fractional electric charge) on the other fixed points. Since $W_2$ is in the direction orthogonal to the two families, we find a non-trivial $D_4$ family
symmetry. This will affect a possible hierarchy of fermion masses. We will discuss the family symmetry and the exotics in more detail next.

The discrete group $D_4$ is a non-abelian discrete subgroup of $SU_2$ of order 8. It is generated by the set of $2 \times 2$ Pauli matrices
\begin{equation}
D_4 = \{\pm 1, \pm \sigma_1, \pm \sigma_3, \mp i \sigma_2\}.
\end{equation}
In our case, the action of the transformation $\sigma_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ takes $F_1 \leftrightarrow F_2$, while the action of $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ takes $F_2 \rightarrow -F_2$. These are symmetries of the string. The first is a symmetry which interchanges the two light families located at opposite sides of one cycle of the orbifolded $SO(4)$ torus. This symmetry corresponds to the geometric translation half way around the cycle which is not broken by the Wilson line lying along the orthogonal cycle. The second is a result of so-called space-group selection rules which require an even number of states at each of these two fixed points. As a result, the theory is invariant under the action of multiplying each state located at, say, the lower fixed point by minus one.

Under $D_4$ the three families of quarks and leptons transform as a doublet, $(F_1, F_2)$, and a singlet, $F_3$. As a consequence of $D_4$ (and additional U(1) symmetries), only the third family can have a tree-level Yukawa coupling to the Higgs (which is also a $D_4$ singlet). All others Yukawa couplings can only be obtained once the family symmetries are broken. Thus the string theory includes a natural Froggatt–Nielsen mechanism [58] for generating a hierarchy of fermion masses. In summary we conclude the following.

- Since the top quarks and the Higgs are derived from the $SU(6)$ chiral adjoint and $20$ hypermultiplet in the 5D bulk, they have a tree-level Yukawa interaction given by
\begin{equation}
W \supset \frac{g_5}{\sqrt{\pi R}} \int_0^{\pi R} dy^5 20^c \times 20 = g_G \, Q_3 \, H_u \, U_3^c.
\end{equation}
where $g_5$ ($g_G$) is the 5D (4D) $SU(6)$ gauge coupling constant evaluated at the string scale. (For a detailed analysis of gauge-Yukawa unification in this context, see [132].)

- The first two families reside at the $\mathbb{Z}_2$ fixed points, resulting in a $D_4$ family symmetry. Hence family symmetry breaking may be used to generate a hierarchy of fermion masses. (For a discussion of $D_4$ family symmetry and phenomenology, see [133]. For a general discussion of discrete non-abelian family symmetries from orbifold compactifications of the heterotic string, see [134].)

4.4.7. More details of ‘benchmark’ model 1 [101]. In order to analyze the theory further one must construct the superpotential and Kähler potential. We have assumed the zeroth order perturbative Kähler potential in our analysis. The superpotential on the other hand can be obtained as a power series in chiral superfields. Each monomial in this series is a holomorphic product of chiral superfields to some given order, $n$, in the fields. We have considered terms up to $n = 8$. The monomials in the superpotential are strongly constrained by string selection rules and gauge invariance. In our analysis we allow all monomials consistent with these selection rules. Although the coefficients of each term can in principle be determined by calculating the vacuum expectation value of the product of the appropriate vertex operators; such a calculation is prohibitive due to the fact that there are hundreds of terms in the polynomial at each order $n$. Thus we assume that any term that is permitted by the selection rules appears in the superpotential with order one coefficient.

Let us now consider the spectrum, exotics, $R$ parity, Yukawa couplings, and neutrino masses. In table 11 we list the states of the model. In addition to the three families of quarks and leptons and one pair of Higgs doublets, we have vector-like exotics (states which can obtain mass without breaking any SM symmetry) and SM singlets. The SM singlets enter the superpotential in several important ways. They can give mass to the vector-like exotics via effective mass terms of the form
\begin{equation}
E E^c \tilde{S}^n
\end{equation}
where $E, E^c$ $(\tilde{S})$ represent the vector-like exotics and SM singlets, respectively. We have checked that all vector-like exotics and unwanted U(1) gauge bosons obtain mass at supersymmetric points in moduli space with $F = D = 0$; leaving only the MSSM states at low energy. The SM singlets also generate effective Yukawa matrices for quarks and leptons, including neutrinos. The charged fermion Yukawa matrices are
\begin{equation}
Y_u = \begin{pmatrix} \tilde{3}_5 & \tilde{5}_5 & \tilde{7}_5 \\ \tilde{5}_5 & \tilde{3}_5 & \tilde{7}_5 \\ \tilde{7}_5 & \tilde{7}_5 & 1 \end{pmatrix}, \quad Y_d = \begin{pmatrix} \tilde{5}_3 & \tilde{5}_3 & 0 \\ \tilde{5}_3 & \tilde{5}_3 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y_e = \begin{pmatrix} \tilde{3}_6 & \tilde{3}_6 & \tilde{3}_6 \\ \tilde{3}_6 & \tilde{3}_6 & \tilde{3}_6 \\ \tilde{3}_6 & \tilde{3}_6 & 0 \end{pmatrix}.
\end{equation}
where $\tilde{S}^n$ represents a polynomial in SM singlets beginning at order $n$ in the product of fields. And we have shown that the three left-handed neutrinos get small mass due to a non-trivial see-saw mechanism involving the 16 right-handed neutrinos and their 13 conjugates. All in all, this ‘benchmark’ model looks very much like the MSSM!

In addition, one of the most important constraints in this construction is the existence of an exact low energy $R$ parity. In this model we identified a generalized $B–L$ (see table 11) which is standard for the SM states and vector-like on the vector-like exotics. This $B–L$ naturally distinguishes the Higgs and lepton doublets. Moreover, we found SM singlet states
\begin{equation}
\tilde{S} = \{h_i, \chi_i, s_i^0\}
\end{equation}
which can get vacuum expectation values preserving a matter parity $Z_2^M$ subgroup of $U(1)_{B-L}$. It is this set of SM singlets which give vector-like exotics mass and effective Yukawa matrices for quarks and leptons. The $\chi$ fields spontaneously break $B-L$ leaving over a discrete $Z_2$ matter parity under which all quarks and leptons are odd and Higgs doublets are even. This symmetry enforces an exact $R$ parity forbidding the baryon or lepton number violating operators $U D D, Q L D, L L E, L H d$.

Finally the mu term vanishes in the supersymmetric limit. This is a consequence of the fact that the coefficient of the $H_u H_d$ term in the superpotential has vacuum quantum numbers. Thus any product of SM singlets which can appear in the pure singlet superpotential can appear as an effective mu term. In fact both the mu term and the singlet superpotential vanish to order 6 in the product of fields. Hence in the supersymmetric vacuum the VEV of the superpotential and the mu term both vanish. As a consequence, when supergravity is considered, the supersymmetric vacuum is consistent with flat Minkowski space. (For a recent discussion on the vanishing of the $\mu$ term and its connection to approximate $R$ symmetries, see [135].)

Note that Yukawa couplings, gauge couplings and vector-like exotic masses are functions of moduli (along SUSY flat directions). Some of these moduli are blow-up modes for some, but not all, of the orbifold fixed points. In fact, two fixed points are not blown up!

### 4.4.8. Gauge coupling unification and proton decay

We have checked whether the SM gauge couplings unify at the string scale in the class of models similar to model 1 above [131]. All of the 15 MSSM-like models of [101] have three families of quarks and leptons and one or more pairs of Higgs doublets. They all admit an $SU(6)$ orbifold GUT with gauge–Higgs unification and the third family in the bulk. They differ, however, in other bulk and brane exotic states. We show that the KK modes of the model, including only those of the third family and the gauge sector, are not consistent with gauge coupling unification at the string scale. Nevertheless, we show that it is possible to obtain unification if one adjusts the spectrum of vector-like exotics below the compactification scale. As an example, see figure 14. Note that the compactification scale is less than the 4D GUT scale and some exotics have mass two orders of magnitude less than $M_3$, while all others are taken to have mass at $M_{\text{string}}$. In addition, the value of the GUT coupling at the string scale, $\alpha_G(M_{\text{string}}) \equiv \alpha_{\text{string}}$, satisfies the weakly coupled heterotic string relation

$$G_N = \frac{1}{4} \alpha_{\text{string}} \alpha'$$

or

$$\alpha_{\text{string}} = \frac{1}{8} \left( \frac{M_{\text{Pl}}}{M_{\text{string}}} \right)^2 .$$

In figure 15 we plot the distribution of solutions with different choices of light exotics. On the same plot we give the proton lifetime due to dimension 6 operators. Recall in these models the two light families are located on the $SU(5)$ branes, thus the proton decay rate is only suppressed by $M_3^{-2}$. Note that 90% of the models are already excluded by the Super-Kamiokande bounds on the proton lifetime. The remaining models may be tested at a next generation megaton water Čerenkov detector.

Note that dimension 5 operators are generated when integrating out color-triplet exotics. These operators may be suppressed by fine tuning. But this is not a satisfactory solution. Recently, however, it was shown that a discrete $Z_4$ $R$ symmetry can be used to naturally suppress both the $\mu$ term and dimension 5 proton decay operators. These terms are forbidden at the perturbative level, but can be generated through non-perturbative interactions. Moreover, string models with such symmetries have been constructed [136].

#### 4.5. F theory/type IIB

We now change directions and discuss some recent progress in F theory model building [90–94, 137, 138]. An $SU(5)$ GUT is obtained on a D7 ‘gauge’ brane $S \times \mathbb{R}^3$. D7 ‘matter’ branes on $S \times \mathbb{R}^3$ also exist with chiral matter in 6D on $\Sigma \times \mathbb{R}^3$ at the intersection of the gauge and matter branes (figure 16). Yukawa couplings enter at the triple intersections $\Sigma_1 \cap \Sigma_2 \cap \Sigma_3$ of matter sub-manifolds (figure 17).

$SU(5)$ is broken to the SM gauge group with non-vanishing hypercharge flux $\langle F_7 \rangle$. Note that this is not possible in the heterotic string! This is because of the term in the Lagrangian

$$\int d^{10}x \langle d B + A_T \wedge (F_7) \rangle$$

which leads to a massive hypercharge gauge boson and consequently a massive photon. In addition, $\langle F_7 \rangle$ on the Higgs brane leads to doublet–triplet splitting. Finally, spinor representations of $SO(10)$ are possible in F theory; although they are not possible in the perturbative type IIB string.

It has been argued that it is difficult to find multiply connected manifolds, with non-zero $\Pi(1)$ and thus breaking GUTs with hypercharge flux is novel advantage for F theory model building. On the other hand, it has been shown [93, 139] that breaking GUTs with hypercharge field strengths induces

| # | irrep | label | # | irrep | label |
|---|---|---|---|---|---|
| 3 | $(3, 2, 1, 1)_{(3/3, 1/3)}$ | $\tilde{q}_i$ | 3 | $(3, 1, 1, 1)_{(4/3, -1/3)}$ | $\tilde{u}_i$ |
| 3 | $(3, 1, 1, 1)_{(2, 1)}$ | $\tilde{e}_i$ | 8 | $(3, 2, 1, 1)_{(1, 0)}$ | $m_i$ |
| 4 | $(3, 1, 1, 1)_{(2, 3, -1/3)}$ | $\tilde{d}_i$ | 1 | $(3, 1, 1, 1)_{(2-3, 1/3)}$ | $d_i$ |
| 4 | $(3, 2, 1, 1)_{(1, 1)}$ | $\tilde{l}_i$ | 1 | $(3, 2, 1, 1)_{(1, 1)}$ | $\tilde{l}_i$ |
| 1 | $(3, 2, 1, 1)_{(1, 0)}$ | $\phi_i$ | 1 | $(3, 2, 1, 1)_{(1, 0)}$ | $\phi_i$ |
| 6 | $(3, 1, 1, 1)_{(2, 1/2, 1/2)}$ | $\tilde{h}_i$ | 6 | $(3, 1, 1, 1)_{(2-3, 1/2, 1/2)}$ | $\tilde{h}_i$ |
| 14 | $(1, 1, 1, 1)_{(1, 1)}$ | $\tilde{e}_i$ | 14 | $(1, 1, 1, 1)_{(1, 1)}$ | $s^+_i$ |
| 16 | $(1, 1, 1, 1)_{(0, 1)}$ | $\tilde{h}_i$ | 13 | $(1, 1, 1, 1)_{(0, 1)}$ | $n_i$ |
| 5 | $(1, 1, 1, 2)_{(0, 1)}$ | $\tilde{h}_i$ | 5 | $(1, 1, 1, 2)_{(0, 1)}$ | $n_i$ |
| 10 | $(1, 1, 1, 2)_{(0, 1)}$ | $\tilde{h}_i$ | 2 | $(1, 1, 1, 2)_{(0, 1)}$ | $y_i$ |
| 6 | $(1, 1, 1, 1)_{(0, 1)}$ | $\tilde{t}_i$ | 6 | $(1, 1, 1, 1)_{(0, 1)}$ | $\tilde{t}_i$ |
| 2 | $(1, 1, 1, 1)_{(1, 1)}$ | $\tilde{t}_i$ | 2 | $(1, 1, 1, 1)_{(1, 1)}$ | $\tilde{t}_i$ |
| 4 | $(1, 1, 1, 1)_{(0, 2)}$ | $\chi_i$ | 32 | $(1, 1, 1, 1)_{(0, 2)}$ | $s^+_i$ |
| 2 | $(3, 1, 1, 1)_{(1, 3/2, -1/2)}$ | $\tilde{u}_i$ | 2 | $(3, 1, 1, 1)_{(1, 3/2, -1/2)}$ | $\tilde{u}_i$ |
Figure 14. An example of the type of gauge coupling evolution we see in these models, versus the typical behavior in the MSSM. The ‘tail’ is due to the power law running of the couplings when towers of KK modes are involved. Unification in this model occurs at $M_{\text{STRING}} \simeq 5.5 \times 10^{17}$ GeV, with a compactification scale of $M_c \simeq 8.2 \times 10^{15}$ GeV, and an exotic mass scale of $M_{\text{EX}} \simeq 8.2 \times 10^{13}$ GeV.

Figure 15. Histogram of solutions with $M_{\text{STRING}} > M_c > M_{\text{EX}}$, showing the models which are excluded by Super-Kamiokande bounds (darker green) and those which are potentially accessible in a next generation proton decay experiment (lighter green). Of 252 total solutions, 48 are not experimentally ruled out by the current experimental bound, and most of the remaining parameter space can be eliminated in the next generation of proposed proton decay searches.

Figure 16. The figure represents three complex planes labeled by $z_i$, $i = 1, 2, 3$. The 4D blue (light) surface is the gauge brane and the matter brane is red (dark). Open strings at the intersection give chiral matter in bi-fundamental representations.

Finally, gravity decouples (i.e. $M_{\text{Pl}} \to \infty$) with a non-compact $z_1$ direction. These are so-called ‘local’ constructions. A bit of progress has also been made in ‘global’ compact constructions [94, 138].

5. Heterotic–F theory duals

F theory defined on a CY$_4$ is dual to the heterotic string defined on a CY$_3$ (figure 18). We are now attempting to construct the F theory dual to our MSSM-like models [143]. The motivation is three-fold.

(i) Three-family MSSM-like models have been found using an $SO(5) \times SO(5) \times SO(4)$ torus mod $Z_4 \times Z_2$ [144] and an $SO(4)^3$ torus mod $Z_2 \times Z_2$ [145]. This suggests a larger class of MSSM-like models. We hope to find a more general description of MSSM-like models, i.e. all models in the same universality class.
(iii) It may help us understand moduli stabilization and SUSY breaking.

(ii) It may also provide a general understanding of moduli space, since from the orbifold viewpoint we must first construct the superpotential before it is possible to identify the moduli.

(iii) It may help us understand moduli stabilization and SUSY breaking.

We should expect F theory duals of heterotic orbifold models to exist? First uplift our $E(8) \times E(8)$ orbifold models onto a smooth Calabi–Yau manifold. Recall that after the first $Z_3$ orbifold plus Wilson line $W_3$ we find a 6D $SU(6)$ orbifold GUT compactified on $(T^2)^2/Z_3 \times T_2$. The complete massless spectrum in this case (including the hidden sector) is given in table 12 [131]. This spectrum satisfies the gravity anomaly constraint $N_H - N_0^c + 29N_T = 273$, where $(N_H = 320, N_0^c = 76, N_T = 1)$ are the number of (hyper-, vector, tensor) multiplets.

Moreover, using the results of Bershadsky et al. [146] we show that the $E(8) \times E(8)$ heterotic string compactified on a smooth $K_3 \times T^2$, with instantons imbedded into $K_3$, is equivalent to the orbifolded theories. For example, with 12 instantons imbedded into an $SU(3) \times SU(2)$ subgroup of the first $E(8)$ leaves an $SU(6) 6D$ GUT with the massless hypermultiplets $(20 + \text{c.c.}) + 18(6 + \text{c.c.})$. Then imbedding 12 instantons into an $E(6)$ subgroup of the second $E(8)$ plus additional higgsing leaves an unbroken $SO(8)$ gauge symmetry with the massless hypermultiplets $4(8v + 8s + 8c + \text{c.c.)}$. This is identical to the massless spectrum of the orbifold GUT, if we neglect the additional $SU(3) \times U(1)^2$ symmetry which is broken when going to the smooth limit. In fact, it is expected that the blow-up modes necessary to smooth out the orbifold singularities carry charges under some of the orbifold gauge symmetries; spontaneously breaking these symmetries. Therefore $K_3 \times T^2$ with instantons is the smooth limit of $T^4/Z_3 \times T^2$ orbifold plus Wilson line. In addition, it was shown that F theory compactified on a Calabi–Yau 3-fold (defined in terms of a torus $T^2$ fibered over the space $F_n \times T^2$) is dual to an $E(8) \times E(8)$ heterotic string compactified on $K_3 \times T_2$ with instantons [146] (see figure 19).

Pictorially we see that the $SU(6)[SO(8)]$ gauge branes are localized at the upper [lower] points on the $z_1$ 2-sphere (figure 19). These 7 branes wrap the 4D surface $S = (z_2, z_3)$. The matter 7 branes intersect the gauge branes at points in $z_2$ and wrap the 4D surface $S' = (z_1, z_3)$. The intersection of the matter and gauge branes is along the 2D surface $\Sigma = (z_3)$. We now need to break the 6D $SU(6)$ GUT to $SU(5)$ and then to the standard model. At the same time we must break the $N = 1$ SUSY in 6D to $N = 1$ in 4D. This is accomplished by acting with the $Z_2$ orbifold on the torus and the 2-sphere. A $U(1)$ flux in $SU(6)$ on the gauge and matter branes breaks $SU(6)$ to $SU(5)$. The breaking to the standard model requires a Wilson line on the torus. However, we now encounter a possible obstruction to finding the F theory dual of our Heterotic orbifold model. We need to keep two orbifold fixed points (figure 20)

(i) otherwise hypercharge gets mass [147],

(ii) the Wilson line shrinks to a point, since $T^2/Z_3$ is topologically equivalent to a 2-sphere, and

(iii) blow-up modes on the heterotic side leave two orbifold fixed points.

In addition, on the heterotic side the two light families are located at 4D orbifold fixed points. We expect that on the F theory side they will be located on $D_3$ branes fixed at the two remaining 4D fixed points. (These problems may be absent in the model of [145] in which the GUT breaking Wilson line wraps a non-contractible cycle, even when all orbifold fixed points are blown up.)

| Multiplet Type | Representation | Number |
|----------------|---------------|--------|
| tensor         | singlet       | 1      |
| vector         | $(35, 1, 1) \oplus (1, 28, 1)$ | $35 + 28$ |
| vector         | $(1, 1, 8) \oplus 5 \times (1, 1, 1)$ | $8 + 5$ |
| hyper          | $(20, 1, 1) \oplus (1, 8, \text{vac}, 1) \oplus 4 \times (1, 1, 1)$ | $20 + 24 + 4$ |
| hyper          | $9 \times ((6, 1, 1) \oplus (6, 1, 1))$ | $108$ |
| hyper          | $9 \times ((1, 1, 3) \oplus (1, 1, 3))$ | $54$ |
| hyper          | $3 \times (1, 8, \text{vac}, 1)$ | $72$ |
| hyper          | $36 \times (1, 1, 1)$ | $36$ |
| SUGRA singlets |               | $2$    |
matter branes are located schematically at the solid point in $Z^3$. More pairs of Higgs doublets and the standard model gauge theories contain three families of quarks and leptons, one or in the supersymmetric limit. The massless spectrum of these Up until now we have focused on MSSM-like models obtained stabilisation.

At the 6D level we have a Calabi–Yau three-fold times a torus. The gauge branes are located at the points $z_1 = 0$ and $\infty$. The matter branes are located schematically at the solid point in $Z_2$.

Figure 19.

At the 4D level we fibre the last torus over the 2-sphere and retain two orbifold fixed points. These two fixed points are where the two light families are conjectured to be located. In addition, as long as the fixed points remain, the Wilson line wrapping the last torus is stable.

Figure 20.

6. Supersymmetry breaking and moduli stabilization

Up until now we have focused on MSSM-like models obtained in the supersymmetric limit. The massless spectrum of these theories contains three families of quarks and leptons, one or more pairs of Higgs doublets and the standard model gauge sector $SU(3) \otimes SU(2) \otimes U(1)_Y$. However, these theories typically include many standard model singlet fields. In the effective field theory limit, these singlet fields have flat potentials. All such fields are called moduli. Some of these moduli are geometric, describing the volume or shape of the internal six dimensions. While in orbifold models there are also blow-up modes which parametrize the smoothing out of the orbifold fixed points. The gauge and Yukawa couplings of the standard model are functions of these moduli. In addition there are, in many cases, extra gauge interactions. Typically a non-abelian hidden sector with new hidden sector quarks and anti-quarks carrying only the hidden sector gauge charge. There may be extra $U(1)$ gauge interactions and also new states which are vector-like under the standard model gauge symmetry. Blow-up modes typically have charge under some of the additional gauge symmetries. This is because these symmetries are only realized in the orbifold limit.

Supersymmetry breaking is studied in the effective supergravity field theory describing the massless states of the particular string model. If we can manage to break supersymmetry in this field theory limit, then moduli stabilization will generically manage itself. Radiative corrections to scalar potentials in non-supersymmetric theories will, in most cases, introduce curvature into tree-level flat directions. Supersymmetry prevents this, however, once supersymmetry is broken then there will be few if any flat directions left. But how to break supersymmetry? It is easy to show that if the tree-level theory is supersymmetric, then supersymmetry cannot be broken in any finite order of perturbation theory. Hence in order to break supersymmetry, one needs non-perturbative effects and non-perturbative effects come in at least two ways, either via non-local tunneling in field space or via strong interaction dynamics, such as gaugino condensates.

There is one additional constraint in any successful SUSY breaking solution. The vacuum energy must be nearly zero in Planck units, i.e. $(T_{\mu \nu}) = M_{Pl}^2 \Delta g_{\mu \nu}$ where the observed cosmological constant has a value $\Lambda \sim O(10^{-120} M_{Pl}^4)$, corresponding to a vacuum energy density of order $(10^{-3} \text{eV})^4$. The scalar potential in supergravity has the form

$$V = e^K (K^{-1}_{ij}(D_i V)(D_j V)^{\ast} - 3 |V|^2) + \sum_a D_a^2$$

where $\mathcal{K}$ is the Kähler potential, $W$ is the superpotential, $D_a$ is the $D$ term associated with the gauge group $a$ and $D_i = \frac{\partial}{\partial \Phi_i}$. $K_{ij}$ is the Kähler covariant derivative with respect to the chiral field $\Phi_i$. Note that $K_{ij} = \frac{\partial \mathcal{K}}{\partial \Phi_i^\ast \partial \Phi_j}$, $K_{ij} = \frac{\partial \mathcal{K}}{\partial \Phi_i \partial \Phi_j^\ast}$. Supersymmetry breaking requires that $D_i W \neq 0$ and/or $D_i \neq 0$. In the supersymmetric limit, the vacuum energy is less than or equal to zero, where $V < 0$ if $\langle V \rangle \neq 0$. Typically, even when supersymmetry is broken, one requires some up-lifting of the vacuum energy in order to obtain a small positive cosmological constant. The study of supersymmetry breaking via non-perturbative effects goes back to the work of Witten and others [148]. Supersymmetry is broken in a hidden sector and then transmitted to the visible SM sector via SM gauge interactions. The messengers of supersymmetry breaking couple directly to the hidden sector and they are also charged under SM gauge interactions. This mechanism for supersymmetry breaking is known as gauge mediated supersymmetry breaking (GMSB). No self-consistent supersymmetry breaking model of this kind existed until the work of Affleck et al. [149]. These authors showed that instanton effects can generate non-perturbative contributions to the superpotential. The non-perturbative terms, along with perturbative contributions, can sometimes lead to supersymmetry breaking. Simple gauge mediated SUSY breaking models were found in [150]. However, the most general understanding of non-perturbative effects in supersymmetric gauge theories came with the work of Seiberg [151]. All of the aforementioned progress was in the context of
global SUSY. It was, however, shown in local supersymmetry, i.e. supergravity, byNilles and others [152] that gaugino condensates lead to SUSY breaking. This SUSY breaking is then transmitted to the SM sector via supergravity, with the low energy SUSY breaking scale set by the gravitino mass, m_{3/2}. In recent years, more mechanisms for supersymmetry breaking have been discovered, including anomaly mediated SUSY breaking and gaugino mediated SUSY breaking. In summary, non-perturbative mechanisms for supersymmetry breaking have a long history. Some, or all, of these mechanisms may be applied in a string theory context. I will only discuss some of the more recent literature on SUSY breaking in string models.

Moduli stabilization and supersymmetry breaking in type II string models have been considered in [153, 154]. Most of the geometric moduli in these theories can be stabilized via gauge fluxes on internal manifolds. In the context of type II string models, Kachru–Kallosh–Linde–Trivedi (KKLT) [154] use fluxes to stabilize most geometric moduli and SUSY breaking comes from non-perturbative contributions to the superpotential. The superpotential has the form

$$ W = W_0 + C e^{-aT} $$

where the constant $W_0$ is a function of integer valued flux contributions which stabilize internal volumes in CY3. Note that $W_0$ sets the scale of supersymmetry breaking and for $m_{3/2} \sim 10$ TeV one needs $W_0 \sim 10^{-16}$ (in Planck units) [155]. $T$ is the volume modulus of a 4 cycle wrapped by a $D_3$-brane. The non-perturbative contribution to the superpotential can arise either from an instanton configuration or a gaugino condensate. There is a problem, however, with this simple set-up, i.e. the vacuum energy in this simplest case is negative. Thus one also needs to introduce an up-lifting sector. KKLT add an anti-$D_3$-brane which adds a positive contribution to the vacuum energy, $V_D = \frac{1}{2} T^2$, and the constant $D$ can be fine-tuned to obtain a small positive cosmological constant.

Soft SUSY breaking contributions to squark, slepton and gaugino masses have been calculated in [155]. They find that matter VEVs dominate over moduli VEVs leading to a scenario for SUSY breaking termed, mirage mediation. In this case, scalar masses are of order $m_{3/2}/(aT) \sim 10^2$–$10^3$ GeV (with $aT \sim \ln(M_{Pl}/m_{3/2}) \sim O(80^2)$), while gaugino masses, similarly suppressed, obtain significant contributions from both moduli and anomaly mediation. As a result, gaugino masses measured at the LHC may appear to unify at a mirage scale which is intermediate between a TeV and GUT scales.

In F theory constructions [156–159] the authors have used the vector-like exotics under SM gauge symmetries as mediators of supersymmetry breaking to the MSSM states. This is known as gauge mediated SUSY breaking and typically results in a gravitino LSP and a long-lived stau or bino NLSP. The idea is to generate a Polonyi potential to induce SUSY breaking. The models contain two $D_7$-branes wrapping 4 cycles and intersecting along a 2D curve $\Sigma$. With the appropriate cohomology, there is one massless chiral multiplet, $X$, at the intersection. Then including a $D_3$-instanton wrapping one of the 4 cycles, a Polonyi term of the form $W \supset F_X X$ is generated with $F_X$ suppressed as compared with the compactification scale of the 4 cycle. The phenomenology is somewhat model dependent. In [157], the authors obtain a gravitino mass $m_{3/2} \sim \frac{1}{\sqrt{M}} \sim 1$ GeV, with a messenger mass $M \sim M_{GUT}$. In [159], on the other hand, the authors obtain a gravitino mass $m_{3/2} \sim \frac{1}{\sqrt{M}} \sim 10$–$100$ MeV, with a messenger mass $M \sim 10^{12}$ GeV. It should be noted that earlier and similar SUSY breaking constructions, in the context of type II strings can be found in [160, 161] or type I in [162].

Stabilization of moduli and SUSY breaking within the context of the heterotic string has a long history. The racetrack mechanism has been discussed the most [163]7. In this scenario, two gaugino condensates generate a non-perturbative superpotential of the form

$$ W_{NP} = \sum_{a=1,2} [C_a \Lambda_a(S, T)]^3 $$

where

$$ \Lambda_a(S, T) = e^{-\frac{3}{2} \frac{S}{T}} f_a(S, T) $$

is the strong dynamics scale, $f_a \sim S$ (at zeroth order) is the gauge kinetic function and $\beta_a = 3N_a$ for $SU(N_a)$ is the coefficient of the one-loop beta function.

In section 4.4.5 we discussed ‘benchmark model 1’ of the ‘mini-landscape’ of heterotic orbifold constructions [101] which has properties very close to that of the MSSM. This model has been analyzed in the supersymmetric limit. It contains an MSSM spectrum with three families of quarks and leptons, one pair of Higgs doublets and an exact R parity. In the orbifold limit, it also contains a small number of vector-like exotics and extra $U(1)$ gauge interactions felt by standard model particles. This theory also contains a large number of standard model singlet fields, some of which are moduli, i.e. thinspace blow-up modes of the orbifold fixed points. It was shown that all vector-like exotics and additional $U(1)$ gauge bosons acquire mass at scales of order the string scale at supersymmetric minima satisfying $F_I = D_a = 0$ for all chiral fields labeled by the index $I$ and all gauge groups labeled by the index $a$. In addition, the value of the gauge couplings at the string scale and the effective Yukawa couplings are determined by the presumed values of the vacuum expectation values (VEVs) for moduli including the dilaton, $S$, the bulk volume and complex structure moduli, $T_I, I = 1, 2, 3$ and $U$ and the SM singlet fields containing the blow-up moduli [147, 166]. Finally the theory also contains a hidden sector $SU(4)$ gauge group with QCD-like chiral matter.

Supersymmetry breaking and moduli stabilization in ‘benchmark model 1’ has been discussed in [167]. In general, the model has a perturbative superpotential satisfying modular invariance constraints, an anomalous $U(1)_A$ gauge symmetry with a dynamically generated Fayet–Iliopoulos $D$-term and the hidden QCD-like non-abelian gauge sector generating a non-perturbative superpotential. The model also has of order 50 chiral singlet moduli. The perturbative superpotential is a polynomial in products of chiral superfields with hundreds of terms. It is not possible at the present time to analyze this model in complete detail. Thus in [167] a simple model

7 The racetrack mechanism has also been discussed in the context of F theory [164] and more recently in terms of M theory [165].
with a dilaton, \( S \), one volume modulus, \( T \), and three standard model singlets was studied. The model has only one gaugino condensate, as is the case for the ‘benchmark model 1.’ We obtain a ‘hybrid KKLT’ kind of superpotential that behaves like a single-condensate for the dilaton \( S \), but as a racetrack for the \( T \) and, by extension, also for the \( U \) moduli. An additional matter \( F \) term, driven by the cancelation of an anomalous \( U(1)_{A} \) \( D \)-term, is the seed for successful up-lifting. Note that similar analyses in the literature have also used an anomalous \( U(1)_{A} \) \( D \)-term in coordination with other perturbative or non-perturbative terms in the superpotential to accomplish SUSY breaking and up-lifting [168, 169]. The discussion in [169] is closest in spirit to that of [167].

The structure of the superpotential of the form \( \mathcal{W} \sim w_{0}e^{-bT} + \phi_{2} e^{-aS-b_{2}T} \) gives the crucial progress:

(i) a ‘hybrid KKLT’ kind of superpotential that behaves like a single-condensate for the dilaton \( S \), but as a racetrack for the \( T \) and, by extension, also for the \( U \) moduli, and

(ii) an additional matter \( F_{\phi} \) term driven by the cancelation of the anomalous \( U(1)_{A} \) \( D \)-term seeds SUSY breaking with successful uplifting.

Note that the constants \( b, b_{2} \) can have either sign. For the case with \( b, b_{2} > 0 \) the superpotential for \( T \) is racetrack-like. However for \( b, b_{2} < 0 \) the scalar potential for \( T \) diverges as \( T \) goes to zero or infinity and compactification is guaranteed [170]. Soft SUSY breaking terms for squarks, sleptons and gauginos were evaluated in this simple model. Note that the gravitino mass scale is set by the constant \( w_{0} \) which was shown to be naturally small in Planck units due to approximate \( R \) symmetries [171].

6.1. Some thoughts on the string landscape and the anthropic principle

The energy balance of the Universe includes the order of 4% visible matter, 23% dark matter and 73% dark energy. The first corresponds to the matter we are made of, while the second is postulated to be made of stable particles with mass of order one TeV; perhaps the lightest supersymmetric particle. The quantity known as dark energy is quite frankly a complete conundrum. It is most simply described as a cosmological constant. It is quite frankly a complete conundrum. It is most simply described as a cosmological constant or vacuum energy density. Why is it so very small has no good explanation (in natural units \( (M_{P}) \) the cosmological constant is of order \( 10^{-120} \)). So perhaps the best explanation to date is known as the anthropic principle. Only if the cosmological constant is small enough can galaxies form and thus provide the \( \text{a priori} \) conditions for humans to be around to observe this Universe [172]. If there is an ensemble of possible Universes with all possible values of the cosmological constant, then we might expect a cosmological constant of the observed value. Indeed this may be the case, and it is good to know that string theory provides such a large ensemble of possible vacua, i.e. the string landscape [79, 80, 173]. The string landscape thus provides a possible explanation for why the cosmological constant is so small. But what other questions does it answer? Can it explain why the observed standard model gauge interactions with three families of quarks and leptons and a Higgs? As discussed earlier in section 4.1, random searches in the string landscape suggest that the standard model is very rare. This may also suggest that string theory cannot make predictions for low energy physics. This would be a shame, but perhaps we are being too hasty. One possibility is that some or all of the above questions are also answered by anthropic arguments. Attempts in this direction have been taken to try and understand the weak scale and the Higgs mass [174] and also the top quark mass [175] or axion cosmology [176].

On the other hand, perhaps understanding the value of the cosmological constant is premature. After all string theory, as such, is a work in progress. In particular, the most complete description of string theory is in terms of the quantization of conformal field theories (CFT) defined on the 2D string world sheets. The fields themselves span, in the case of superstrings, a 10D space-time background. In fact, space-time itself is an emergent quantity. Perturbative calculations rely on assuming a particular classical background configuration. Non-perturbative physics comes in through strong non-abelian gauge interactions (fluxes and/or condensates), instantons (gauge and world sheet) and \( D \)-branes. Whether the string CFT is defined on a lattice or in terms of fermions and/or bosons is an arbitrary choice. But each choice determines a different massless sector. The massless sector can be described by an effective field theory, but how does one distinguish one arbitrary choice of background from another? Each effective field theory has its own calculable ground state, but how does one compare one ground state for one background to another with a different background choice? Why, in practice, are four space-time dimensions large, while the other six are curled up into an, as yet, unobservable ‘ball’? Why should we have \( N = 1 \) SUSY in four dimensions? All of these questions are crucial to an understanding of our Universe and deeply profound, but they have proven to be extremely difficult nuts to crack.

However, there is much more data in our low energy phenomenology than just the cosmological constant. Much in these data suggest more symmetry—the family structure and charge quantization, the hierarchy of fermion masses and mixing angles, and the absence of large flavor violating processes all suggest grand unification with family symmetries. Perhaps string theory can be predictive, if we understood the rules for choosing the correct position in the string landscape. At the moment, these rules are not understood, so the best guess relies on statistical analyses. However, we have argued in these pages that finding the standard model (or minimal supersymmetric standard model) is best achieved by requiring SUSY GUTs at the first step. I would like to argue that if we could better understand the string landscape in these ‘fertile patches’ we might then be able to understand the rules needed to choose this region of the landscape. Of course, a major caveat in this whole discussion is the assumption that our low energy Universe is describable by the minimal supersymmetric standard model. This assumption will soon be tested at the Large Hadron Collider located at CERN near Geneva, Switzerland.

One small hint in this direction is provided by the heterotic orbifold models discussed in section 4.4. These heterotic
orbifold models have some amazing properties.

(i) They incorporate local GUTs with two complete families localized at orbifold fixed points;
(ii) They incorporate a 5D SU(6) orbifold GUT with gauge-Higgs unification and the third family in the bulk;
(iii) As a consequence, they have gauge-Yukawa unification for the top quark (thus explaining why the top quark is heavy);
(iv) They incorporate doublet–triplet splitting with a \( \mu \) term which is naturally small;
(v) They have an exact \( R \) parity. (Moreover, recently it was discovered that similar models can incorporate a \( Z^R_4 \) symmetry which allows all Yukawa interactions and neutrino masses while forbidding the \( \mu \) term and dimension 5 baryon and lepton number violating operators at the perturbative level \[136\]. The \( Z^R_4 \) symmetry is possible due to the final \( Z_2 \) orbifold);
(vi) As a consequence of the \( Z_2 \) orbifold, the model has a \( D_4 \) family symmetry which can ameliorate problems with flavor changing neutral currents while at the same time accommodating a hierarchy of quark and lepton masses;
(vii) Approximate \( R \) symmetries naturally generate a small constant contribution to the superpotential, setting the scale for the gravitino mass once supersymmetry breaking is generated.

Such models are generically related to smooth heterotic models compactified on \((K_3 \times T^2)/Z_2\). Perhaps a clue to why the standard model is special can be found in this class of heterotic models.

7. Conclusion

In this paper, we have discussed an evolution of SUSY GUT model building. We saw that 4D SUSY GUTs have many virtues. However, there are some problems which suggest that these models may be difficult to derive from a more fundamental theory, i.e. string theory. We then discussed orbifold GUT field theories which solve two of the most difficult problems of 4D GUTs, i.e. GUT symmetry breaking and Higgs doublet–triplet splitting. We then showed how some orbifold GUTs can find an ultra-violet completion within the context of heterotic string theory.

The flood gates are now wide open. In recent work \[101\] we have obtained many models with features like the MSSM: the SM gauge group with three families, and vector-like exotics which can, in principle, obtain large mass. The models have an exact \( R \) parity and non-trivial Yukawa matrices for quarks and leptons. In addition, neutrinos obtain mass via the see-saw mechanism. We also showed that gauge coupling unification can be accommodated.

Of course, this is not the end of the story. It is just the beginning. In order to obtain predictions for the LHC, one must stabilize the moduli and break supersymmetry. In fact, these two conditions are not independent, since once SUSY is broken, the moduli will be stabilized. The scary fact is that the moduli have to be stabilized at just the right values to be consistent with low energy phenomenology. Interesting first steps in the construction of complete and phenomenologically testable string models have been taken.

Acknowledgments

This work was accomplished with partial support from DOE grant DOE/ER/01545-888. I also wish to acknowledge support from CERN, where most of this article was written.

References

[1] Georgi H 1975 AIP Conf. Proc. 23 575
Fritzsch H and Minkowski P 1975 Ann. Phys. 93 193
[2] Pati J and Salam A 1973 Phys. Rev. D 8 1240
For more discussion on the standard charge assignments in this formalism, see Davidson A 1979 Phys. Rev. D 20 776
[3] Mohapatra R N and Marshak R E 1980 Phys. Lett. B 91 222
[4] Georgi H and Glashow S L 1974 Phys. Rev. Lett. 32 438
Henry T 1981 Nucl. Phys. B 188 513
Dimopoulos S and Raby S 1981 Nucl. Phys. B 192 353
Dine M, Fischler W and Srednicki M 1981 Nucl. Phys. B 189 575
[5] Dimopoulos S, Raby S and Wilczek F 1981 Phys. Rev. D 24 1681
Dimopoulos S and Georgi H 1981 Nucl. Phys. B 193 150
Ibanez L and Ross G G 1981 Phys. Lett. B 105 439
Sacai N 1981 Z. Phys. C 11 153
[6] Einhorn M B and Jones D R T 1982 Nucl. Phys. B 196 475
Marciano W J and Senjanovic G 1982 Phys. Rev. D 25 3092
[7] Yao W-M et al (Particle Data Group) 2006 J. Phys. G: Nucl. Part. Phys. 33 1 (special issue)
[8] Langacker R and Polonsky N 1995 Phys. Rev. D 52 3081
Alcami M L, Feruglio F, Lin Y and Varagnolo A 2005 J. High Energy Phys. JHEP03(2005)054
[9] Carena M S, Pokorski S and Wagner C E M 1993 Nucl. Phys. B 406 59
[10] Raby S, Ratz M and Schmidt-Hoberg K 2010 Phys. Lett. B 687 342
[11] Gell-Mann M, Ramond P and Slansky R 1979 Supergavity ed P van Nieuwenhuizen and Freedman D Z (Amsterdam: North-Holland) p 315
[12] Weinberg S 1982 Phys. Rev. D 26 287
[13] Sakai N and Yanagida T 1982 Nucl. Phys. B 197 533
[14] Farrar G and Fayet P 1978 Phys. Lett. B 76 575
[15] Dimopoulos S, Raby S and Wilczek F 1982 Phys. Lett. B 112 133
Ellis J, Nanopoulos D V and Rudaz S 1982 Nucl. Phys. B 202 43
[16] Ibanez L E and Ross G G 1992 Nucl. Phys. B 368 3
[17] Barbieri R, Strumia A and Berezhiani Z 1997 Phys. Lett. B 407 250
[18] Giudice G F and Rattazzi R 1997 Phys. Lett. B 406 321
[19] Barbier R et al 2005 Phys. Rep. 420 1
[20] For a recent discussion, see Aulakh C S, Bajc B, Melfo A, Ellis J, Nanopoulos D V and Rudaz S 1982 Nucl. Phys. B 236 1
Blazek T, Carena M, Raby S and Wilczek F 1997 Phys. Rev. D 54 2261
[21] Lucas V and Raby S 1996 Phys. Rev. D 54 2261
[22] Altarelli G, Feruglio F and Masina I 2000 J. High Energy Phys. JHEP11(2000)040
[23] Dermisek R, Malti A and Raby S 2001 Phys. Rev. D 63 035001
[24] Babu K S, Pati J C and Wilczek F 2000 Nucl. Phys. B 566 33
[25] Dimopoulos S and Wilczek F 1981 Incomplete multiplets in supersymmetric unified models NSF Report NSF-ITP-82-07 (unpublished)
[26] Babu K S and Barr S M 1993 Phys. Rev. D 48 5534
[27] See talk by Miura M 2010 (Super-Kamiokande Collaboration) ICHEP 2010 (Paris, France)
[28] See talks by Matthew Earl 2000 NNN Workshop, Talk (Irvine, February 2000)
Totsuka Y 2000 SUSY2K (CERN, June 2000) Suzuki Y 2000 Int. Workshop on Neutrino Oscillations and their Origins (Tokyo, Japan, December 2000) Suzuki Y 2001 arXiv:hep-ex/0110005 Kobayashi K (Super-Kamiokande Collaboration) 2001 Search for nuclear decay from Super-Kamiokande, Prepared for 27th Int. Cosmic Ray Conf. (ICRC 2001) (Hamburg, Germany, 7–15 August 2001) For published results see Hayato Y et al (Super-Kamiokande Collaboration) 1999 Phys. Rev. Lett. 83 1529 Kobayashi K et al (Super-Kamiokande Collaboration) 2005 arXiv:hep-ex/0502026 [28] Goto T and Nihei T 1999 Phys. Rev. D 59 115009 Murayama H and Pierce A 2002 Phys. Rev. D 65 055009 [29] Choi K S 2008 Phys. Lett. B 668 392 [30] Chanowitz M, Ellis J and Gaillard M K 1978 Nucl. Phys. B 135 66 For the corresponding SUSY analysis, see Einhorn M and Jones D R T 1982 Nucl. Phys. B 196 475 Inoue K, Kakuto A, Komatsu H and Takeshita S 1982 Prog. Theor. Phys. 67 1889 Ibanez L E and Lopez C 1983 Phys. Lett. B 126 54 Ibanez L E and Lopez C 1984 Nucl. Phys. B 233 511 [31] Georgi H and Hanopulos D V 1979 Nucl. Phys. B 159 16 [32] Harvey J, Ramond P and Reiss D B 1980 Phys. Lett. B 92 309 Harvey J, Ramond P and Reiss D B 1982 Nucl. Phys. B 199 223 [33] Banks T 1988 Nucl. Phys. B 303 172 Olechowski M and Pokorski S 1988 Phys. Lett. B 214 393 Pokorski S 1990 Nucl. Phys. B: Proc. Suppl. 13 606 Ananthanarayan B, Lazarides G and Shafi Q 1991 Phys. Rev. D 43 1613 Shafi Q and Ananthanarayan B 1992 1991 Summer School in High Energy Physics and Cosmology (17 June–9 August, 1991, Trieste, Italy) (Singapore: World Scientific) Dimopoulos S, Hall J. L and Raby S 1992 Phys. Rev. Lett. 68 1984 Dimopoulos S, Hall J. L and Raby S 1992 Phys. Rev. D 45 4192 Anderson G et al 1993 Phys. Rev. D 47 3702 Ananthanarayan B, Lazarides G and Shafi Q 1993 Phys. Lett. B 300 245 Anderson G et al 1994 Phys. Rev. D 49 3660 Ananthanarayan B, Shafi Q and Wang X M 1994 Phys. Rev. D 50 5980 Lazarides G, Shafi Q and Wetterich C 1981 Nucl. Phys. B 181 287 Clark T E, Kuo T K and Nakagawa N 1982 Phys. Lett. B 115 28 Babu K S and Mohapatra R N 1993 Phys. Rev. Lett. 70 2845 Bajc B, Senjanovic G and Vissani F 2003 Phys. Rev. Lett. 90 051802 Goh H S, Mohapatra R N and Ng S P 2003 Phys. Lett. B 570 215 Goh H S, Mohapatra R N and Ng S P 2003 Phys. Rev. D 68 115008 Dutta B, Mimura Y and Mohapatra R N 2004 Phys. Rev. D 70 015013 [37] LEP Higgs Working Group, ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration and OPAL Collaboration 2001 arXiv:hep-ex/0107030 [38] Carena M and Haber H E 2003 Prog. Part. Nucl. Phys. 50 63 [39] Hall I J, Rattazzi R and Sarid U 1994 Phys. Rev. D 50 7048 Carena M, Olechowski M, Pokorski S and Wagner C E M 1994 Nucl. Phys. B 419 213 Rattazzi R and Sarid U 1997 Nucl. Phys. B 501 297 [40] Blazek T, Dermisek R and Raby S 2002 Phys. Rev. Lett. 88 111804 Blazek T, Dermisek R and Raby S 2002 Phys. Rev. D 65 115004 [41] Tobe K and Wells J D 2003 arXiv:hep-ph/0301015 Auto D, Baer H, Balazs C, Belyaev A, Ferrandis J and Tata X 2003 arXiv:hep-ph/0302155 [42] Baer H, Kraml S, Sekmen S and Summy H 2008 J. High Energy Phys. JHEP0803(2008)056 [43] Balazs C and Dermisek R 2003 arXiv:hep-ph/0303161 [44] Bagger J A, Feng J L, Polonsky N and Zhang R J 2000 Phys. Lett. B 473 264 [45] Gabbiani F, Gabrielli E, Masiero A and Silvestrini L 1996 Nucl. Phys. B 477 321 Besmer T, Greub C and Huth T 2001 Nucl. Phys. B 609 359 [46] Georgi H and Jarlskog C 1979 Phys. Lett. B 86 297 Babu K S and Mohapatra R N 1995 Phys. Rev. Lett. 74 2418 Blazek T, Carena M, Raby S and Wagner C 1997 Phys. Rev. D 56 6919 Barbieri R, Hall L J, Raby S and Romano R A 1997 Nucl. Phys. B 493 3 Blazek T, Raby S and Tobe K 1999 Phys. Rev. D 60 113001 Blazek T, Raby S and Tobe K 2000 Phys. Rev. D 62 055001 Shafi Q and Tavartkiladze Z 2000 Phys. Rev. D 48 1475 Albright C H and Barr S M 2000 Phys. Rev. Lett. 85 244 Altarelli G, Feruglio F and Masina I 2000 J. High Energy Phys. JHEP11(2000)040 Bereziani Z and Rossi A 2001 Nucl. Phys. B 594 113 Albright C H and Barr S M 2001 Phys. Rev. D 64 073010 Dermisek R and Raby S 2005 Phys. Lett. B 622 327 [49] Weinberg S 1977 A Festschrift for I I Rabi ed L Motz and I I Rabi (New York: New Academic of Sciences) Wilczek F and Zee A 1977 Phys. Lett. B 70 418 Fritzsch H 1977 Phys. Lett. B 70 436 [50] Kim H D, Raby S and Schradin L 2004 Phys. Rev. D 69 092002 [51] Hall L J and Rasin A 1993 Phys. Lett. B 315 164 [52] Minkowski P 1977 Phys. Lett. B 67 421 [53] Yanagida T 1979 Proc. Workshop on the Unified Theory and the Baryon Number of the Universe (Tsukuba, Japan) ed O Sawada and A Sugamoto KEK Report no 79-19 Glasshow S 1980 Quarks and leptons Proc. Cargese Lectures ed M Levy (New York: Plenum) Mohapatra R N and Senjanovic G 1980 Phys. Rev. Lett. 44 912 [54] Dermisek R and Raby S 2005 Phys. Lett. B 622 327 [55] Dermisek R, Harada M and Raby S 2006 Phys. Rev. D 74 035011 [56] Albrecht M, Altmannshofer W, Buras A J, Guadagnoli D and Straub D M 2007 J. High Energy Phys. JHEP10(2007)055 [57] Altmannshofer W, Guadagnoli D, Raby S and Straub D M 2008 arXiv:0801.4363 [hep-ph] [58] Froggatt C D and Nielsen H B 1979 Nucl. Phys. B 147 277 [59] Dermisek R, Raby S, Roszkowski L and Ruiz De Austri R 2003 J. High Energy Phys. JHEP04(2003)037 Dermisek R, Raby S, Roszkowski L and Ruiz De Austri R 2005 J. High Energy Phys. JHEP09(2005)029 [60] Barr S M and Raby S 1997 Phys. Rev. Lett. 79 4748 [61] Masiero A, Nanopoulos D V, Tamvakis K and Yanagida T 1982 Phys. Lett. B 115 380 Grinstein B 1982 Nucl. Phys. B 206 387 [62] Dermisek R, Raby S and Nandi S 2002 Nucl. Phys. B 641 327 [63] Hall L J and Nomura Y 2001 Phys. Rev. D 64 055003 [64] Dienes K R, Dudas E and Gherghetta T 1999 Nucl. Phys. B 537 47 [65] Contino R, Pilo L, Rattazzi R and Trincherini E 2002 Nucl. Phys. B 622 227 [66] Ghilencea D M and Groot Nibbelink S 2002 Nucl. Phys. B 641 35 [67] Hall L J and Nomura Y 2002 Phys. Rev. D 66 075004
Hall L J, Salem M P and Watari T 2007 Phys. Rev. D 76 093001

Hall L J and Nomura Y 2008 Phys. Rev. D 78 035001

Bousso R, Hall L J and Nomura Y 2009 Phys. Rev. D 80 063510

[174] Arkani-Hamed N and Dimopoulos S 2005 J. High Energy Phys. JHEP06(2005)073

Arkani-Hamed N, Dimopoulos S, Giudice G F and Romanino A 2005 Nucl. Phys. B 709 3

[175] Feldstein B, Hall L J and Watari T 2006 Phys. Rev. D 74 095011

[176] Wilczek F 2004 arXiv:hep-ph/0408167

Hertzberg M P, Tegmark M and Wilczek F 2008 Phys. Rev. D 78 083507