16. Grand Unified Theories

Written April 2002 by S. Raby (Ohio State University).

16.1. Grand Unification

16.1.1. *Standard Model: An Introduction*:

In spite of all the successes of the Standard Model [SM], it is unlikely to be the final theory. It leaves many unanswered questions. Why the local gauge interactions $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, and why 3 families of quarks and leptons? Moreover, why does one family consist of the states $[Q, u^c, d^c; L, e^c]$ transforming as $[(3, 2, 1/3)$, $(\bar{3}, 1, -4/3)$, $(\bar{3}, 1, 2/3)$; $(1, 2, -1)$, $(1, 1, 2)]$, where $Q = (u, d)$, and $L = (\nu, e)$ are $\mathrm{SU}(2)_L$ doublets, and u^c, d^c, e^c are charge conjugate $\mathrm{SU}(2)_L$ singlet fields with the $\mathrm{U}(1)_Y$ quantum numbers given? [We use the convention that electric charge $Q_{EM} = T_{3L} + Y/2$ and all fields are left-handed.] Note the SM gauge interactions of quarks and leptons are completely fixed by their gauge charges. Thus, if we understood the origin of this charge quantization, we would also understand why there are no fractionally charged hadrons. Finally, what is the origin of quark and lepton masses, or the apparent hierarchy of family masses and quark mixing angles? Perhaps if we understood this, we would also know the origin of CP violation, the solution to the strong CP problem, the origin of the cosmological matter-antimatter asymmetry, or the nature of dark matter.

The SM has 19 arbitrary parameters; their values are chosen to fit the data. Three arbitrary gauge couplings: g_3, g, g' (where g, g' are the $\mathrm{SU}(2)_L$, $\mathrm{U}(1)_Y$ couplings, respectively) or equivalently, $\alpha_s = (g_3^2/4\pi)$, $\alpha_{EM} = (e^2/4\pi)$ $(e = g \ \sin\theta_W)$, and $\sin^2\theta_W = (g')^2/(g^2 + (g')^2)$. In addition, there are 13 parameters associated with the 9 charged fermion masses and the four mixing angles in the CKM matrix. The remaining 3 parameters are v, λ [the Higgs VEV (vacuum expectation value) and quartic coupling] (or equivalently, M_Z, m_h^0), and the QCD θ parameter. In addition, there are hints of new physics beyond the SM, such as neutrino masses. With 3 light Majorana neutrinos, there are at least 9 additional parameters in the neutrino sector; 3 masses and 6 mixing angles. In summary, the SM has too many arbitrary parameters, and leaves open too many unresolved questions to be considered complete. These are the problems which grand unified theories hope to address.

16.1.2. *Charge Quantization*:

In the Standard Model, quarks and leptons are on an equal footing; both fundamental particles without substructure. It is now clear that they may be two faces of the same coin; unified, for example, by extending QCD (or $\mathrm{SU}(3)_C$) to include leptons as the fourth color, $\mathrm{SU}(4)_C$ [1]. The complete Pati-Salam gauge group is $\mathrm{SU}(4)_C \times \mathrm{SU}(2)_L \times \mathrm{SU}(2)_R$, with the states of one family $[(Q, L), (Q^c, L^c)]$ transforming as $[(4, 2, 1)$, $(\bar{4}, 1, \bar{2})]$, where $Q^c = (d^c, u^c)$, $L^c = (e^c, \nu^c)$ are doublets under $\mathrm{SU}(2)_R$. Electric charge is now given by the relation $Q_{EM} = T_{3L} + T_{3R} + 1/2(B-L)$, and $\mathrm{SU}(4)_C$ contains the subgroup $\mathrm{SU}(3)_C \times (B-L)$ where B (L) is baryon (lepton) number. Note ν^c has no SM quantum numbers and is thus completely "sterile." It is introduced to complete the $\mathrm{SU}(2)_R$ lepton doublet. This additional state is desirable when considering neutrino masses.

Although quarks and leptons are unified with the states of one family forming two irreducible representations of the gauge group, there are still 3 independent gauge couplings (two if one also imposes parity, *i.e.*, $L \leftrightarrow R$ symmetry). As a result, the three low-energy gauge couplings are still independent arbitrary parameters. This difficulty is resolved by embedding the SM gauge group into the simple unified gauge group, Georgi-Glashow SU(5), with one universal gauge coupling α_G defined at the grand unification scale M_G [2]. Quarks and leptons still sit in two irreducible representations, as before, with a $\mathbf{10} = [Q, u^c, e^c]$ and $\bar{\mathbf{5}} = [d^c, L]$. Nevertheless, the three low energy gauge couplings are now determined in terms of two independent parameters : α_G and M_G. Hence, there is one prediction.

In order to break the electroweak symmetry at the weak scale and give mass to quarks and leptons, Higgs doublets are needed which can sit in either a $\mathbf{5_H}$ or $\bar{\mathbf{5}}_\mathbf{H}$. The additional 3 states are color triplet Higgs scalars. The couplings of these color triplets violate baryon and lepton number, and nucleons decay via the exchange of a single color triplet Higgs scalar. Hence, in order not to violently disagree with the non-observation of nucleon decay, their mass must be greater than $\sim 10^{10-11}$ GeV. Note, in supersymmetric GUTs, in order to cancel anomalies, as well as give mass to both up and down quarks, both Higgs multiplets $\mathbf{5_H}$, $\bar{\mathbf{5}}_\mathbf{H}$ are required. As we shall discuss later, nucleon decay now constrains the color triplet Higgs states in a SUSY GUT to have mass significantly greater than M_G.

Complete unification is possible with the symmetry group SO(10), with one universal gauge coupling α_G, and one family of quarks and leptons sitting in the 16-dimensional-spinor representation $\mathbf{16} = [\mathbf{10} + \bar{\mathbf{5}} + \mathbf{1}]$ [3]. The SU(5) singlet $\mathbf{1}$ is identified with ν^c. In Table 16.1 we present the states of one family of quarks and leptons, as they appear in the $\mathbf{16}$. It is an amazing and perhaps even profound fact that all the states of a single family of quarks and leptons can be represented digitally as a set of 5 zeros and/or ones or equivalently as the tensor product of 5 "spin" 1/2 states (see Table 16.1). The first three "spins" correspond to $SU(3)_C$ color quantum numbers, while the last two are $SU(2)_L$ weak quantum numbers. In fact, an $SU(3)_C$ rotation just raises one color index and lowers another, thereby changing colors $\{r,\ b,\ y\}$. Similarly an $SU(2)_L$ rotation raises one weak index and lowers another, thereby flipping the weak isospin from up to down or vice versa. In this representation, weak hypercharge Y is given by the simple relation $Y = 2/3(\sum$ color spins$)-(\sum$ weak spins$)$ where the sum is over the spin values $\{\pm 1/2\}$. SU(5) rotations then raise (or lower) a color index, while at the same time lowering (or raising) a weak index. It is easy to see that such rotations can mix the states $\{Q,\ u^c,\ e^c\}$ and $\{d^c, L\}$ among themselves, and ν^c is a singlet. The new SO(10) rotations [not in SU(5)] are then given by either raising or lowering any two spins. For example, by lowering the two weak indices ν^c rotates into e^c, etc.

SO(10) has two inequivalent maximal breaking patterns: $SO(10) \rightarrow SU(5) \times U(1)_X$ and $SO(10) \rightarrow SU(4)_C \times SU(2)_L \times SU(2)_R$. In the first case, we obtain Georgi-Glashow SU(5) if Q_{EM} is given in terms of SU(5) generators alone, or so-called flipped SU(5) [4] if Q_{EM} is partly in $U(1)_X$. In the latter case, we have the Pati-Salam symmetry. If SO(10) breaks directly to the SM at M_G, then we retain the prediction for gauge coupling unification. However, more possibilities for breaking (hence more breaking scales

Table 16.1: The quantum numbers of the **16** dimensional representation of SO(10) are represented as a tensor product of 5 "spin" 1/2 states with the values $\pm$ denoting the spin states $|\pm\frac{1}{2}>$ and with the condition that we have an even number of $-$ spins.

State	Y	Color	Weak
$\boldsymbol{\nu}^c$	0	$+\,+\,+$	$+\,+$
$\boldsymbol{e}^c$	2	$+\,+\,+$	$-\,-$
$\boldsymbol{u}_r$	1/3	$-\,+\,+$	$+\,-$
$\boldsymbol{d}_r$	1/3	$-\,+\,+$	$-\,+$
$\boldsymbol{u}_b$	1/3	$+\,-\,+$	$+\,-$
$\boldsymbol{d}_b$	1/3	$+\,-\,+$	$-\,+$
$\boldsymbol{u}_y$	1/3	$+\,+\,-$	$+\,-$
$\boldsymbol{d}_y$	1/3	$+\,+\,-$	$-\,+$
$\boldsymbol{u}^c_r$	$-4/3$	$+\,-\,-$	$+\,+$
$\boldsymbol{u}^c_b$	$-4/3$	$-\,+\,-$	$+\,+$
$\boldsymbol{u}^c_y$	$-4/3$	$-\,-\,+$	$+\,+$
$\boldsymbol{d}^c_r$	2/3	$+\,-\,-$	$-\,-$
$\boldsymbol{d}^c_b$	2/3	$-\,+\,-$	$-\,-$
$\boldsymbol{d}^c_y$	2/3	$-\,-\,+$	$-\,-$
$\boldsymbol{\nu}$	-1	$-\,-\,-$	$+\,-$
$\boldsymbol{e}$	-1	$-\,-\,-$	$-\,+$

and more parameters) are available in SO(10). Nevertheless with one breaking pattern SO(10) $\rightarrow$ SU(5) $\rightarrow$ SM, where the last breaking scale is M_G, the predictions from gauge coupling unification are preserved. The Higgs multiplets in minimal SO(10) are contained in the fundamental $\mathbf{10_H} = [\mathbf{5_H},\ \bar{\mathbf{5}}_\mathbf{H}]$ representation. Note only in SO(10) does the gauge symmetry distinguish quark and lepton multiplets from Higgs multiplets.

Finally, larger symmetry groups have been considered. For example, $E(6)$ has a fundamental representation **27**, which under SO(10) transforms as a $[\mathbf{16} + \mathbf{10} + \mathbf{1}]$. The breaking pattern $E(6) \rightarrow \mathrm{SU}(3)_C \times \mathrm{SU}(3)_L \times \mathrm{SU}(3)_R$ is also possible. With the additional permutation symmetry $Z(3)$ interchanging the three SU(3)s, we obtain so-called "trinification [5]," with a universal gauge coupling. The latter breaking pattern has been used in phenomenological analyses of the heterotic string [6]. Note, in larger symmetry groups, such as $E(6)$, SU(6), etc., there are now many more states which have not been observed and must be removed from the effective low-energy theory. In particular, three families of **27**s in $E(6)$ contain three Higgs type multiplets transforming as **10**s of SO(10). This makes these larger symmetry groups unattractive starting points for model building.

16.1.3. *Gauge coupling unification*:

The biggest paradox of grand unification is to understand how it is possible to have a universal gauge coupling g_G in a grand unified theory [GUT], and yet have three unequal gauge couplings at the weak scale with $g_3 > g > g'$. The solution is given in terms of the concept of an effective field theory [EFT] [7]. The GUT symmetry is spontaneously broken at the scale M_G, and all particles not in the SM obtain mass of order M_G. When calculating Green's functions with external energies $E \gg M_G$, we can neglect the mass of all particles in the loop and hence all particles contribute to the renormalization group running of the universal gauge coupling. However, for $E \ll M_G$, one can consider an effective field theory including only the states with mass $< E \ll M_G$. The gauge symmetry of the EFT is $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, and the three gauge couplings renormalize independently. The states of the EFT include only those of the SM; 12 gauge bosons, 3 families of quarks and leptons, and one or more Higgs doublets. At M_G, the two effective theories [the GUT itself is most likely the EFT of a more fundamental theory defined at a higher scale] must give identical results; hence we have the boundary conditions $g_3 = g_2 = g_1 \equiv g_G$, where at any scale $\mu < M_G$, we have $g_2 \equiv g$ and $g_1 = \sqrt{5/3}\ g'$. Then using two low-energy couplings, such as $\alpha_s(M_Z)$, $\alpha_{EM}(M_Z)$, the two independent parameters α_G, M_G can be fixed. The third gauge coupling, $\sin^2\theta_W$ in this case, is then predicted. This was the procedure up until about 1991 [8,9]. Subsequently, the uncertainties in $\sin^2\theta_W$ were reduced tenfold. Since then, $\alpha_{EM}(M_Z)$, $\sin^2\theta_W$ have been used as input to predict α_G, M_G, and $\alpha_s(M_Z)$ [10].

Note, the above boundary condition is only valid when using one-loop-renormalization group [RG] running. With precision electroweak data, however, it is necessary to use two-loop-RG running. Hence, one must include one-loop-threshold corrections to gauge coupling boundary conditions at both the weak and GUT scales. In this case, it is always possible to define the GUT scale as the point where $\alpha_1(M_G) = \alpha_2(M_G) \equiv \tilde{\alpha}_G$ and $\alpha_3(M_G) = \tilde{\alpha}_G\ (1 + \epsilon_3)$. The threshold correction ϵ_3 is a logarithmic function of all states with mass of order M_G and $\tilde{\alpha}_G = \alpha_G + \Delta$, where α_G is the GUT coupling constant above M_G, and Δ is a one-loop-threshold correction. To the extent that gauge coupling unification is perturbative, the GUT threshold corrections are small and calculable. This presumes that the GUT scale is sufficiently below the Planck scale or any other strong coupling extension of the GUT, such as a strongly coupled string theory.

Supersymmetric grand unified theories [SUSY GUTs] are an extension of non-SUSY GUTs [11]. The key difference between SUSY GUTs and non-SUSY GUTs is the low-energy effective theory. The low-energy effective field theory in a SUSY GUT is assumed to satisfy $N = 1$ supersymmetry down to scales of order the weak scale, in addition to the SM gauge symmetry. Hence, the spectrum includes all the SM states, plus their supersymmetric partners. It also includes one pair (or more) of Higgs doublets; one to give mass to up-type quarks, and the other to down-type quarks and charged leptons. Two doublets with opposite hypercharge Y are also needed to cancel fermionic triangle anomalies. Note, a low-energy SUSY-breaking scale (the scale at which the SUSY partners of SM particles obtain mass) is necessary to solve the gauge hierarchy problem.

Simple non-SUSY SU(5) is ruled out, initially by the increased accuracy in the measurement of $\sin^2\theta_W$, and by early bounds on the proton lifetime (see below) [9].

However, by now LEP data [10] has conclusively shown that SUSY GUTs is the new Standard Model; by which we mean the theory used to guide the search for new physics beyond the present SM. SUSY extensions of the SM have the property that their effects decouple as the effective SUSY-breaking scale is increased. Any theory beyond the SM must have this property simply because the SM works so well. However, the SUSY-breaking scale cannot be increased with impunity, since this would reintroduce a gauge hierarchy problem. Unfortunately there is no clear-cut answer to the question, "When is the SUSY-breaking scale too high?" A conservative bound would suggest that the third generation squarks and sleptons must be lighter than about 1 TeV, in order that the one-loop corrections to the Higgs mass from Yukawa interactions remain of order the Higgs mass bound itself.

At present, gauge coupling unification within SUSY GUTs works extremely well. Exact unification at M_G, with two-loop-RG running from M_G to M_Z, and one-loop-threshold corrections at the weak scale, fits to within 3 σ of the present precise low-energy data. A small threshold correction at M_G ($\epsilon_3 \sim -4\%$) is sufficient to fit the low-energy data precisely.* This may be compared to non-SUSY GUTs, where the fit misses by $\sim 12\ \sigma$, and a precise fit requires new weak-scale states in incomplete GUT multiplets, or multiple GUT-breaking scales.**

16.1.4. *Nucleon Decay*:

Baryon number is necessarily violated in any GUT [15]. In SU(5), nucleons decay via the exchange of gauge bosons with GUT scale masses, resulting in dimension-6 baryon-number-violating operators suppressed by $(1/M_G^2)$. The nucleon lifetime is calculable and given by $\tau_N \propto M_G^4/(\alpha_G^2\ m_p^5)$. The dominant decay mode of the proton (and the baryon-violating decay mode of the neutron), via gauge exchange, is $p \to e^+\ \pi^0$ ($n \to e^+\ \pi^-$). In any simple gauge symmetry, with one universal GUT coupling and scale (α_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. Experimental searches for nucleon decay began with the Kolar Gold Mine, Homestake, Soudan, NUSEX, Frejus, HPW, and IMB detectors [8]. The present experimental bounds come from Super-Kamiokande and Soudan II. We discuss these results shortly. Non-SUSY GUTs are also ruled out by the non-observation of nucleon decay [9]. In SUSY GUTs,

---

* This result implicitly assumes universal GUT boundary conditions for soft SUSY-breaking parameters at M_G. In the simplest case, we have a universal gaugino mass $M_{1/2}$, a universal mass for squarks and sleptons m_{16}, and a universal Higgs mass m_{10}, as motivated by SO(10). In some cases, threshold corrections to gauge coupling unification can be exchanged for threshold corrections to soft SUSY parameters. See for example, Ref. 12 and references therein.

** Non-SUSY GUTs with a more complicated breaking pattern can still fit the data. For example, non-SUSY SO(10) $\to$ SU(4)$_C \times$ SU(2)$_L \times$ SU(2)$_R$ $\to$SM, with the second breaking scale of order an intermediate scale, determined by light neutrino masses using the see-saw mechanism, can fit the low-energy data for gauge couplings [13], and at the same time survive nucleon decay bounds [14], discussed in the following section.

the GUT scale is of order 3×10^{16} GeV, as compared to 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 [11] with $\tau_p \sim 10^{34-38}$ yrs.

However, in SUSY GUTs, there are additional sources for baryon-number violation—dimension-4 and -5 operators [16]. Although the 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 $(\mathbf{10}\ \bar{\mathbf{5}}\ \bar{\mathbf{5}}) \supset (u^c\ d^c\ d^c) + (Q\ L\ d^c) + (e^c\ L\ L)$, and $(\mathbf{10}\ \mathbf{10}\ \mathbf{10}\ \bar{\mathbf{5}}) \supset (Q\ Q\ Q\ L) + (u^c\ u^c\ d^c\ e^c) + B$ 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 order $(1/M_G)$.

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 a $\mathbf{5_H}$, $\bar{\mathbf{5}}_\mathbf{H}$, and R parity distinguishes the $\bar{\mathbf{5}}$ (quarks and leptons) from $\bar{\mathbf{5}}_\mathbf{H}$ (Higgs). R parity [17] (or more precisely, its cousin, family reflection symmetry (see Dimopoulos and Georgi [11] and DRW [18]) takes $F \to -F$, $H \to H$ with $F = \{\mathbf{10},\ \bar{\mathbf{5}}\}$, $H = \{\bar{\mathbf{5}}_\mathbf{H},\ \mathbf{5_H}\}$. This forbids the dimension-4 operator $(\mathbf{10}\ \bar{\mathbf{5}}\ \bar{\mathbf{5}})$, but allows the Yukawa couplings of the form $(\mathbf{10}\ \bar{\mathbf{5}}\ \bar{\mathbf{5}}_\mathbf{H})$ and $(\mathbf{10}\ \mathbf{10}\ \mathbf{5_H})$. It also forbids the dimension-3, lepton-number-violating operator $(\bar{\mathbf{5}}\ \mathbf{5_H}) \supset (L\ H_u)$, with a coefficient with dimensions of mass which, like the μ parameter, could be of order the weak scale and the dimension-5, baryon-number-violating operator $(\mathbf{10}\ \mathbf{10}\ \mathbf{10}\ \bar{\mathbf{5}}_\mathbf{H}) \supset (Q\ Q\ Q\ H_d) + \cdots$.

Note, 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 [19]. However, these symmetries cannot be embedded in a GUT. Thus, in a SUSY GUT, only R parity can prevent unwanted dimension four 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.

Note also, R parity 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 Z(4) center of SO(10) distinguishes **10**s from **16**s, and can be associated with R parity [20].

Dimension-5 baryon-number-violating operators may be forbidden at tree level by symmetries in SU(5), 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. Hence, the color triplet partners of Higgs doublets must necessarily obtain mass of order the GUT scale. The dominant decay modes from dimension-5 operators are $p \to K^+ \, \overline{\nu}$ $(n \to K^0 \, \overline{\nu})$. This is due to a simple symmetry argument; the operators $(Q_i \; Q_j \; Q_k \; L_l)$, $(u_i^c \; u_j^c \; d_k^c \; e_l^c)$ (where i, j, k, $l = 1, 2, 3$ are family indices, and color and weak indices are implicit) must be invariant under $\mathrm{SU}(3)_C$ and $\mathrm{SU}(2)_L$. 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 member must exist in the final state [18].

Recent Super-Kamiokande bounds on the proton lifetime severely constrain these dimension-6 operators with dimension-5 operators with $\tau_{(p \to e^+ \pi^0)} > 5.0 \times 10^{33}$ yrs (79.3 ktyr exposure), $\tau_{(n \to e^+ \pi^-)} > 5 \times 10^{33}$ yrs (61 ktyr), and $\tau_{(p \to K^+ \overline{\nu})} > 1.6 \times 10^{33}$ yrs (79.3 ktyr), $\tau_{(n \to K^0 \overline{\nu})} > 1.7 \times 10^{32}$ yrs (61 ktyr) at (90% CL) based on the listed exposures [21]. These constraints are now sufficient to rule out minimal SUSY SU(5) [22]. Non-minimal Higgs sectors in SU(5) or SO(10) theories still survive [24,25]. The upper bound on the proton lifetime from these theories is approximately a factor of 5 above the experimental bounds. They are also being pushed to their theoretical limits. Hence, if SUSY GUTs are correct, then nucleon decay must be seen soon.

Is there a way out of this conclusion? String theories, and recent field theoretic constructions [26,27], contain grand unified symmetries realized in higher dimensions. In most heterotic string models, when compactifying all but four of these extra dimensions, only the MSSM is recovered as a symmetry of the effective four dimensional field theory. [Of course, this is not required by string theory, and string theory models exist whose low-energy field theory is a SUSY GUT [28].] In the process of compactification and GUT symmetry breaking, color triplet Higgs states are removed (projected out of the massless sector of the theory). In addition, the same projections, in heterotic string models, typically rearrange the quark and lepton states so that the massless states which survive emanate from different GUT multiplets. In these models, proton decay due to dimension-5 operators can be severely suppressed, or eliminated completely. In addition, proton decay due to dimension-6 operators may be enhanced due to threshold corrections at the GUT scale which effectively lower the GUT scale [27], or eliminate it altogether, if the states of one family come from different irreducible representations. Hence, the observation of proton decay may distinguish extra-dimensional GUTs from four-dimensional ones.

Before concluding the topic of baryon-number violation, consider the status of $\Delta B = 2$ neutron- anti-neutron oscillations. Generically, the leading operator for this process is the dimension-9 six-quark operator $G_{(\Delta B = 2)} \; (u^c \; d^c \; d^c \; u^c \; d^c \; d^c)$, with dimensionful coefficient $G_{(\Delta B = 2)} \sim 1/M^5$. The present experimental bound $\tau_{n - \overline{n}} \geq 0.86 \times 10^8$ sec. at 90% CL [30] probes only up to the scale $M \leq 10^6$ GeV. For $M \sim M_G$, $n - \overline{n}$ oscillations appear to be unobservable for any GUT (for a recent discussion see Ref. 29).

16.1.5. *Yukawa coupling unification*:

16.1.5.1. *3rd generation, b–τ or t–b–τ unification:*

If quarks and leptons are two sides of the same coin, related by a new grand unified gauge symmetry, then that same symmetry relates the Yukawa couplings (and hence the masses) of quarks and leptons. In SU(5), there are two independent renormalizable Yukawa interactions given by λ_t $(\mathbf{10\,10\,5_H})$ + λ $(\mathbf{10\,\bar{5}\,\bar{5}_H})$. These contain the SM interactions λ_t $(\mathbf{Q\,u^c\,H_u})$ + λ $(\mathbf{Q\,d^c\,H_d}\ +\ \mathbf{e^c\,L\,H_d})$. Hence, at the GUT scale, we have the tree-level relation, $\lambda_b = \lambda_\tau \equiv \lambda$ [31]. In SO(10), there is only one independent renormalizable Yukawa interaction given by λ $(\mathbf{16\,16\,10_H})$, which gives the tree-level relation, $\lambda_t = \lambda_b = \lambda_\tau \equiv \lambda$ [32,33]. Note, in the discussion above, we assume the minimal Higgs content, with Higgs in $\mathbf{5}$, $\mathbf{\bar{5}}$ for SU(5) and $\mathbf{10}$ for SO(10). With Higgs in higher-dimensional representations, there are more possible Yukawa couplings.

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

Including one-loop-threshold corrections at M_Z, and additional RG running, one finds the top, bottom, and τ-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, $\tan\beta > 2.4$ [34]. 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 [35]. Yukawa unification (consistent with low energy data) is only possible in a restricted region of SUSY parameter space with important consequences for SUSY searches [36].

16.1.5.2. *Three families:*

Simple Yukawa unification is not possible for the first two generations, of quarks and leptons. Consider the SU(5) GUT scale relation $\lambda_b = \lambda_\tau$. If extended to the first two generations, one would have $\lambda_s = \lambda_\mu$, $\lambda_d = \lambda_e$, which gives $\lambda_s/\lambda_d = \lambda_\mu/\lambda_e$. The last relation is a renormalization group invariant, and is thus satisfied at any scale. In particular, at the weak scale, one obtains $m_s/m_d = m_\mu/m_e$, which is in serious disagreement with the data, namely $m_s/m_d \sim 20$ and $m_\mu/m_e \sim 200$. An elegant solution to this problem was given by Georgi and Jarlskog [37]. Of course, a three-family model must also give the observed CKM mixing in the quark sector. Note, 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. Three-family models exist which fit all the data, including neutrino masses and mixing [38].

16.1.6. *Neutrino Masses*:

Atmospheric and solar neutrino oscillations require neutrino masses. Adding three "sterile" neutrinos ν^c with the Yukawa coupling λ_ν $(\nu^{\mathbf{c}}\ \mathbf{L}\ \mathbf{H_u})$, one easily obtains three massive Dirac neutrinos with mass $m_\nu = \lambda_\nu\ v_u$. However, in order to obtain a tau neutrino with mass of order 0.1 eV, one needs $\lambda_{\nu_\tau}/\lambda_\tau \leq 10^{-10}$. The see-saw mechanism, on the other hand, can naturally explain such small neutrino masses [15,39]. Since ν^c has no SM quantum numbers, there is no symmetry (other than global lepton number) which prevents the mass term $\frac{1}{2}\ \nu^c\ M\ \nu^c$. Moreover, one might expect $M \sim M_G$. 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 $\frac{1}{2}\ (\mathbf{L}\ \mathbf{H_u})\ \lambda_\nu^T\ M^{-1}\ \lambda_\nu\ (\mathbf{L}\ \mathbf{H_u})$. This then leads to a 3×3 Majorana neutrino mass matrix $\mathbf{m} = m_\nu^T\ M^{-1}\ m_\nu$.

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

16.1.7. *Selected Topics*:

16.1.7.1. *Magnetic Monopoles:*

In the broken phase of a GUT, there are typically localized classical solutions carrying magnetic charge under an unbroken U(1) symmetry [40]. These magnetic monopoles with mass of order M_G/α_G are produced during the GUT phase transition in the early universe. The flux of magnetic monopoles is experimentally found to be less than $\sim 10^{-16}$ $\mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{sr}^{-1}$ [41]. Many more are predicted however, hence the GUT monopole problem. In fact, one of the original motivations for an inflationary universe is to solve the monopole problem by invoking an epoch of rapid inflation after the GUT phase transition [42]. This would have the effect of diluting the monopole density as long as the reheat temperature is sufficiently below M_G. Parenthetically, it was also shown that GUT monopoles can catalyze nucleon decay [43].

16.1.7.2. *Baryogenesis via Leptogenesis:*

Baryon-number-violating operators in SU(5) or SO(10) preserve the global symmetry $B-L$. Hence, the value of the cosmological $B-L$ density is an initial condition of the theory, and is typically assumed to be zero. On the other hand, anomalies of the electroweak symmetry violate $B+L$ while also preserving $B-L$. Hence, thermal fluctuations in the early universe, via so-called sphaleron processes, can drive $B+L$ to zero, washing out any net baryon number generated in the early universe at GUT temperatures.

One way out of this dilemma is to generate a net $B-L$ dynamically in the early universe. We have just seen that neutrino oscillations suggest a new scale of physics of order 10^{14} GeV. This scale is associated with heavy Majorana neutrinos with mass M. If in the early universe, the decay of the heavy neutrinos is out of equilibrium and violates both lepton number and CP, then a net lepton number may be generated. This lepton number will then be partially converted into baryon number via electroweak processes [44].

16.1.7.3. *GUT symmetry breaking:*

The grand unification symmetry is necessarily broken spontaneously. Scalar potentials (or superpotentials) exist whose vacua spontaneously break SU(5) and SO(10). These potentials are ad hoc (just like the Higgs potential in the SM), and, therefore it is hoped that they may be replaced with better motivated sectors. Gauge coupling unification now tests GUT-breaking sectors, since it is one of the two dominant corrections to the GUT threshold correction ϵ_3. The other dominant correction comes from the Higgs sector and doublet-triplet splitting. This latter contribution is always positive $\epsilon_3 \propto \ln(M_T/M_G)$ (where M_T is an effective color triplet Higgs mass), while the low-energy data requires $\epsilon_3 < 0$. Hence, the GUT-breaking sector must provide a significant (of order -8%) contribution to ϵ_3 to be consistent with the Super-K bound on the proton lifetime [23,24,25,38].

In string theory (and GUTs in extra-dimensions), GUT breaking may occur due to boundary conditions in the compactified dimensions [26,27]. This is still ad hoc. The major benefits are that it does not require complicated GUT-breaking sectors, and it can suppress dimension-5 baryon-violating operators.

16.1.7.4. *Doublet-triplet splitting:*

The Minimal Supersymmetric Standard Model has a μ problem: why is the coefficient of the bilinear Higgs term in the superpotential $\mu\,(\mathbf{H_u}\,\mathbf{H_d})$ of order the weak scale when, since it violates no low-energy symmetry, it could be as large as M_G? In a SUSY GUT, the μ problem is replaced by the problem of *doublet-triplet* splitting—giving mass of order M_G to the color triplet Higgs, and mass μ to the Higgs doublets. Several mechanisms for natural doublet-triplet splitting have been suggested, such as the sliding singlet, missing partner or missing VEV [45], and pseudo-Nambu-Goldstone boson mechanisms. Particular examples of the missing partner mechanism for SU(5) [25], the missing VEV mechanism for SO(10) [24,38], and the pseudo-Nambu-Goldstone boson mechanism for SU(6) [46], have been shown to be consistent with gauge coupling unification and proton decay. There are also several mechanisms for explaining why μ is of order the SUSY-breaking scale [47]. Finally, for a recent review of the μ problem and some suggested solutions in SUSY GUTs and string theory, see Ref. 48 and references therein.

16.2. Conclusion

Grand unification of the strong and electroweak interactions at a unique high energy scale $M_G \sim 3 \times 10^{16}$ GeV requires

- gauge coupling unification,
- low-energy supersymmetry [with a large SUSY desert], and
- nucleon decay.

The first prediction has already been verified. Perhaps the next two will soon be seen. Whether or not Yukawa couplings unify is more model dependent. Nevertheless, the "digital" 16-dimensional representation of quarks and leptons in SO(10) is very compelling, and may yet lead to an understanding of fermion masses and mixing angles.

In any event, the experimental verification of the first three pillars of SUSY GUTs would forever change our view of Nature. Moreover, the concomitant evidence for a vast SUSY desert would expose a huge lever arm for discovery. For then it would become clear that experiments probing the TeV scale could reveal physics at the GUT scale and perhaps beyond.

References:

1. J. Pati and A. Salam, Phys. Rev. **D8**, 1240 (1973);

   For more discussion on the standard charge assignments in this formalism, see A. Davidson, Phys. Rev. **D20**, 776 (1979); and R.N. Mohapatra and R.E. Marshak, Phys. Lett. **B91**, 222 (1980).
2. H. Georgi and S.L. Glashow, Phys. Rev. Lett. **32**, 438 (1974).
3. H. Georgi, Particles and Fields, *Proceedings of the APS Div. of Particles and Fields*, ed. C. Carlson, p. 575 (1975);

   H. Fritzsch and P. Minkowski, Ann. Phys. **93**, 193 (1975).
4. S.M. Barr, Phys. Lett. **B112**, 219 (1982).
5. A. de Rujula, H. Georgi, and S.L. Glashow, p. 88, *5th Workshop on Grand Unification*, ed. K. Kang, H. Fried, and P. Frampton, World Scientific, Singapore (1984);

   See also earlier paper by Y. Achiman and B. Stech, p. 303, "New Phenomena in Lepton-Hadron Physics," ed. D.E.C. Fries and J. Wess, Plenum, NY (1979).
6. B.R. Greene *et al.*, Nucl. Phys. **B278**, 667 (1986), and Nucl. Phys. **B292**, 606 (1987);

   B.R. Greene, C.A. Lutken, and G.G. Ross, Nucl. Phys. **B325**, 101 (1989).
7. H. Georgi, H. Quinn, and S. Weinberg, Phys. Rev. Lett. **33**, 451 (1974);

   see also the definition of effective field theories by S. Weinberg, Phys. Lett. **91B**, 51 (1980).
8. See talks on proposed and running nucleon decay experiments, and theoretical talks by P. Langacker, p. 131, and W.J. Marciano and A. Sirlin, p. 151, in *The Second Workshop on Grand Unification*, eds. J.P. Leveille, L.R. Sulak, and D.G. Unger, Birkhäuser, Boston (1981).

9. W.J. Marciano, p. 190, *Eighth Workshop on Grand Unification*, ed. K. Wali, World Scientific Publishing Co., Singapore (1987).
10. U. Amaldi, W. de Boer, and H. Fürstenau, Phys. Lett. **B260**, 447 (1991);
J. Ellis, S. Kelly and D.V. Nanopoulos, Phys. Lett. **B260**, 131 (1991);
P. Langacker and M. Luo, Phys. Rev. **D44**, 817 (1991);
P. Langacker and N. Polonsky, Phys. Rev. **D47**, 4028 (1993);
M. Carena, S. Pokorski, and C.E.M. Wagner, Nucl. Phys. **B406**, 59 (1993);
see also the review by S. Dimopoulos, S. Raby, and F. Wilczek, Physics Today, 25–33, October (1991).
11. S. Dimopoulos, S. Raby, and F. Wilczek, Phys. Rev. **D24**, 1681 (1981);
S. Dimopoulos and H. Georgi, Nucl. Phys. **B193**, 150 (1981);
L. Ibanez and G.G. Ross, Phys. Lett. **105B**, 439 (1981);
N. Sakai, Z. Phys. **C11**, 153 (1981);
M.B. Einhorn and D.R.T. Jones, Nucl. Phys. **B196**, 475 (1982);
W.J. Marciano and G. Senjanovic, Phys. Rev. **D25**, 3092 (1982).
12. G. Anderson *et al.*, in *New directions for high-energy physics*, Snowmass 1996, eds. D.G. Cassel, L. Trindle Gennari, and R.H. Siemann, `hep-ph/9609457`.
13. R.N. Mohapatra and M.K. Parida, Phys. Rev. **D47**, 264 (1993).
14. D.G. Lee *et al.*, Phys. Rev. **D51**, 229 (1995).
15. M. Gell-Mann, P. Ramond, and R. Slansky, in *Supergravity*, eds. P. van Nieuwenhuizen and D.Z. Freedman, North-Holland, Amsterdam, 1979, p. 315.
16. S. Weinberg, Phys. Rev. **D26**, 287 (1982);
N. Sakai and T Yanagida, Nucl. Phys. **B197**, 533 (1982).
17. G. Farrar and P. Fayet, Phys. Lett. **B76**, 575 (1978).
18. S. Dimopoulos, S. Raby, and F. Wilczek, Phys. Lett. **112B**, 133 (1982);
J. Ellis, D.V. Nanopoulos, and S. Rudaz, Nucl. Phys. **B202**, 43 (1982).
19. L.E. Ibanez and G.G. Ross, Nucl. Phys. **B368**, 3 (1992).
20. For a recent discussion, see C.S. Aulakh *et al.*, Nucl. Phys. **B597**, 89 (2001).
21. See talks by Matthew Earl, *NNN workshop*, Irvine, February (2000);
Y. Totsuka, *SUSY2K*, CERN, June (2000);
Y. Suzuki, *International Workshop on Neutrino Oscillations and their Origins*, Tokyo, Japan, December (2000), and *Baksan School, Baksan Valley*, Russia, April (2001), `hep-ex/0110005`. For published results see : Y. Hayato *et al.* (Super-Kamiokande Collab.), Phys. Rev. Lett. **83**, 1529 (1999).
22. H. Murayama and A. Pierce, Phys. Rev. **D65**, 055009 (2002).
23. K.S. Babu and S.M. Barr, Phys. Rev. **D48**, 5354 (1993);
V. Lucas and S. Raby, Phys. Rev. **D54**, 2261 (1996);

S.M. Barr and S. Raby, Phys. Rev. Lett. **79**, 4748 (1997) and references therein.

24. R. Dermíšek, A. Mafi, and S. Raby, Phys. Rev. **D63**, 035001 (2001);
K.S. Babu, J.C. Pati, and F. Wilczek, Nucl. Phys. **B566**, 33 (2000).
25. G. Altarelli, F. Feruglio, I. Masina, JHEP **0011**, 040 (2000);
see also earlier papers by A. Masiero *et al.*, Phys. Lett. **B115**, 380 (1982);
B. Grinstein, Nucl. Phys. **B206**, 387 (1982).
26. P. Candelas *et al.*, Nucl. Phys. **B258**, 46 (1985);
L.J. Dixon *et al.*, Nucl. Phys. **B261**, 678 (1985), and Nucl. Phys. **B274**, 285 (1986).
27. Y. Kawamura, Prog. Theor. Phys. **105**, 999 (2001);
L.J. Hall and Y. Nomura, Phys. Rev. **D64**, 055003 (2001);
R. Barbieri, L.J. Hall, and Y. Nomura, `hep-ph/0106190 (2001)`.
28. Z. Kakushadze and S.H.H. Tye, Phys. Rev. **D54**, 7520 (1996);
Z. Kakushadze *et al.*, Int. J. Mod. Phys. **A13**, 2551 (1998).
29. K.S. Babu and R.N. Mohapatra, Phys. Lett. **B518**, 269 (2001).
30. M. Baldoceolin *et al.*, Z. Phys. **C63**, 409 (1994).
31. M. Chanowitz, J. Ellis, and M.K. Gaillard, Nucl. Phys. **B135**, 66 (1978);
For the corresponding SUSY analysis, see M. Einhorn and D.R.T. Jones, Nucl. Phys. **B196**, 475 (1982);
K. Inoue *et al.*, Prog. Theor. Phys. **67**, 1889 (1982);
L.E. Ibañez and C. Lopez, Nucl. Phys. **B233**, 511 (1984).
32. H. Georgi and D.V. Nanopoulos, Nucl. Phys. **B159**, 16 (1979);
J. Harvey, P. Ramond, and D.B. Reiss, Phys. Lett. **92B**, 309 (1980);
Nucl. Phys. **B199**, 223 (1982).
33. T. Banks, Nucl. Phys. **B303**, 172 (1988);
M. Olechowski and S. Pokorski, Phys. Lett. **B214**, 393 (1988);
S. Pokorski, Nucl. Phys. (Proc. Supp.) **B13**, 606 (1990);
B. Ananthanarayan, G. Lazarides, and Q. Shafi, Phys. Rev. **D44**, 1613 (1991);
Q. Shafi and B. Ananthanarayan, ICTP Summer School lectures (1991);
S. Dimopoulos, L.J. Hall, and S. Raby, Phys. Rev. Lett. **68**, 1984 (1992), and Phys. Rev. **D45**, 4192 (1992);
G. Anderson *et al.*, Phys. Rev. **D47**, 3702 (1993);
B. Ananthanarayan, G. Lazarides, and Q. Shafi, Phys. Lett. **B300**, 245 (1993);
G. Anderson *et al.*, Phys. Rev. **D49**, 3660 (1994);
B. Ananthanarayan, Q. Shafi, and X.M. Wang, Phys. Rev. **D50**, 5980 (1994).
34. LEP Higgs Working Group and ALEPH Collab., DELPHI Collab., L3 Collab., and OPAL Collab., Preliminary results, `hep-ex/0107030 (2001)`.
35. L.J. Hall, R. Rattazzi, and U. Sarid, Phys. Rev. **D50**, 7048 (1994);

M. Carena *et al.*, Nucl. Phys. **B419**, 213 (1994);

R. Rattazzi and U. Sarid, Nucl. Phys. **B501**, 297 (1997).

36. T. Blažek, R. Dermíšek, and S. Raby, Phys. Rev. Lett. **88**, 111804 (2002) and `hep-ph/0201081`.

37. H. Georgi and C. Jarlskog, Phys. Lett. **86B**, 297 (1979).

38. K.S. Babu and R.N. Mohapatra, Phys. Rev. Lett. **74**, 2418 (1995);

V. Lucas and S. Raby, Phys. Rev. **D54**, 2261 (1996);

T. Blažek *et al.*, Phys. Rev. **D56**, 6919 (1997);

R. Barbieri *et al.*, Nucl. Phys. **B493**, 3 (1997);

T. Blazek, S. Raby, and K. Tobe, Phys. Rev. **D60**, 113001 (1999), and Phys. Rev. **D62**, 055001 (2000);

Q. Shafi and Z. Tavartkiladze, Phys. Lett. **B487**, 145 (2000);

C.H. Albright and S.M. Barr, Phys. Rev. Lett. **85**, 244 (2000);

K.S. Babu, J.C. Pati, and F. Wilczek, Nucl. Phys. **B566**, 33 (2000);

G. Altarelli, F. Feruglio, I. Masina, Ref. 25;

Z. Berezhiani and A. Rossi, Nucl. Phys. **B594**, 113 (2001).

39. T. Yanagida, in *Proceedings of the Workshop on the Unified Theory and the Baryon Number of the Universe*, eds. O. Sawada and A. Sugamoto, KEK report No. 79-18, Tsukuba, Japan, 1979;

R.N. Mohapatra and G. Senjanovic, Phys. Rev. Lett. **44**, 912 (1980).

40. G. 't Hooft, Nucl. Phys. **B79**, 276 (1974);

A.M. Polyakov, Pis'ma Zh. Eksp. Teor. Fiz. **20**, 430 (1974) [JETP Lett. **20**, 194 (1974)];

For a pedagogical introduction, see S. Coleman, in *Aspects of Symmetry*, Selected Erice Lectures, Cambridge University Press, Cambridge, (1985), and P. Goddard and D. Olive, Rep. Prog. Phys. **41**, 1357 (1978).

41. I. De Mitri, (MACRO Collab.), Nucl. Phys. (Proc. Suppl.) **B95**, 82 (2001).
42. For a review, see A.D. Linde, *Particle Physics and Inflationary Cosmology*, Harwood Academic, Switzerland (1990).
43. V. Rubakov, Nucl. Phys. **B203**, 311 (1982), Institute of Nuclear Research Report No. P-0211, Moscow (1981), unpublished;

    C. Callan, Phys. Rev. **D26**, 2058 (1982);

    F. Wilczek, Phys. Rev. Lett. **48**, 1146 (1982);

    See also, S. Dawson and A.N. Schellekens, Phys. Rev. **D27**, 2119 (1983).
44. M. Fukugita and T. Yanagida, Phys. Lett. **B174**, 45 (1986);

    see also the recent review by W. Buchmuller, `hep-ph/0107153 (2001)` and references therein.
45. S. Dimopoulos and F. Wilczek, *Proceedings Erice Summer School*, ed. A. Zichichi (1981);

    K.S. Babu and S.M. Barr, Phys. Rev. **D50**, 3529 (1994).
46. R. Barbieri, G.R. Dvali, and A. Strumia, Nucl. Phys. **B391**, 487 (1993);

    Z. Berezhiani, C. Csaki, and L. Randall, Nucl. Phys. **B444**, 61 (1995);

    Q. Shafi and Z. Tavartkiladze, Phys. Lett. **B522**, 102 (2001).
47. G.F. Giudice and A. Masiero, Phys. Lett. **B206**, 480 (1988);

    J.E. Kim and H.P. Nilles, Mod. Phys. Lett. **A9**, 3575 (1994).
48. L. Randall and C. Csaki, *Proceedings Pascos/Hopkins 1995*, `hep-ph/9508208`;

    E. Witten, `hep-ph/0201018`.