PAPER

A stronger case for superunification post Higgs boson discovery

To cite this article: Pran Nath and Raza M Syed 2017 Phys. Scr. 92 124005

View the article online for updates and enhancements.
A stronger case for superunification post Higgs boson discovery

Pran Nath1,3 and Raza M Syed1,2

1 Department of Physics, Northeastern University, Boston, MA 02115-5000, United States of America
2 Department of Physics, American University of Sharjah, PO Box 26666, Sharjah, United Arab Emirates
3 Author to whom any correspondence should be addressed.

E-mail: nath@neu.edu and rsyed@aus.edu

Received 27 September 2017
Accepted for publication 20 October 2017
Published 17 November 2017

Abstract
Supersymmetry and more specifically supergravity grand unification allow one to extrapolate physics from the electroweak scale up to the grand unification scale consistent with electroweak data. Here we give a brief overview of their current status and show that the case for supersymmetry is stronger as a result of the Higgs boson discovery with a mass measurement at \( \sim 125 \text{ GeV} \) consistent with the supergravity grand unification prediction that the Higgs boson mass lies below 130 GeV. Thus the discovery of the Higgs boson and the measurement of its mass provide a further impetus for the search for sparticles to continue at the current and future colliders.

Keywords: supersymmetry, supergravity, unification

1. Introduction

The standard model of particle physics [1] based on the gauge group \( SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \) and three generations of quarks and leptons is a highly successful model at low energies up to the electroweak scale. One of the basic elements of the model is that it is anomaly free. Specifically, the quarks and the leptons have the \( SU(3)_C, SU(2)_L, U(1)_Y \) quantum numbers so that \( q(3, 2, \frac{2}{3}), \ u^c(3, 1, -\frac{1}{3}), \ d^c(3, \frac{1}{3}, 1), \ L(1, 2, -\frac{1}{2}), \ e^c(1, 1, 1), \) where the first two entries refer to the \( SU(3)_C, SU(2)_L \) representations, and the last entry refers to the hypercharge defined so that \( Q = T_3 + Y. \) The anomaly free condition in this case implies that one has \( \sum_i f_i Y_i = 0, \) where \( f_i \) is a product of multiplicity and color factor. Here one generation of quarks and leptons exactly satisfies the anomaly cancellation condition. The interesting phenomenon is that while the leptons have integral charge, \( Q = -1 \) for charged lepton, \( Q = 0 \) for the neutrino, the quarks have fractional charge, \( 2/3 \) for the up quarks and \( -1/3 \) for the down quark. The charge assignment appears intriguing and leads one to ask if there exist a larger framework within which one may understand such charge assignments. Such a framework must be more unified and exist at a larger scale. There are other aspects which point to the possibility that a more unified framework may exist such as the product nature of standard model group. Here it requires three gauge coupling constants \( g_1, g_2, g_Y \) to describe the interactions. One may speculate if they are remnants of a single coupling. Such issues were the subject of investigations in the early seventies. Thus in 1973–74 Pati and Salam [2, 3] proposed that the standard model was remnant of the group \( G(4, 2, 2) \equiv SU(4) \otimes SU(2)_L \otimes SU(2)_R. \) Here leptons and quarks are unified with the leptons arising as a fourth color. Soon after the work of [2] Georgi and Glashow [4] proposed the group \( SU(5) \) which in addition to unifying the quarks and the leptons, also unifies the gauge coupling constants. A group which gives an even greater unification was proposed subsequently. This is the group \( SO(10) \) [5]. It has the benefit of unifying a full generation of quarks and leptons in one irreducible representation of \( SO(10). \) The general criteria for grand unification is that one needs those unification groups which have chiral representations. Here the relevant groups are \( SU(N); \ SO(4N + 2), N \geq 1; \) \( E_6. \) As noted one of the constraints on model building is that of anomaly cancellation. Here the groups \( E_6 \) and \( SO(4N + 2), N \geq 1 \) are automatically anomaly free while for \( SU(N) \) one needs combinations of representations which are anomaly free.

However, as is well known non-supersymmetric models have a serious fine tuning problem [6]. Quantum loop corrections to the Higgs boson mass-squared give contributions which are quadratically divergent in the cutoff. In grand unified theories that cutoff would be the GUT scale \( \sim 10^{16} \) GeV, which is much larger than the electroweak scale. A
cancellation of the loop term would require a fine tuning of one part in $10^{28}$. The cancellation of the quadratic divergence occurs naturally in supersymmetry [7]. It is desirable then to formulate unified models using supersymmetry. One persistent problem here concerns breaking of supersymmetry which is essential for building a viable phenomenology. This is problematic in global supersymmetry. In order to break supersymmetry in a phenomenologically viable way one needs local supersymmetry/supergravity [8–10] (for a more extensive discussion see [11]). Indeed grand unified models based on local supersymmetry provide the appropriate framework for unifying the strong and the electroweak interactions [12]. Gauge coupling unification is an important touchstone of unified model [13]. An important success of supersymmetry models is the unification of gauge coupling constants consistent with LEP data [14–17]. We note in passing that Planck scale physics could affect the predictions at the grand unification scale [18] (see [19, 20]). Further, a significant feature of supergravity grand unification is that it is also the appropriate vehicle for the analysis of string based models since supergravity is the low energy limit of strings, i.e., at scales $E < M_P$ (see, e.g., [21]).

The outline of the rest of the paper is as follows: in section 2 we discuss the first works on unification beyond the standard model. These include the quark-lepton unification and the unification of gauge coupling constants. As noted the group $SO(10)$ is now the preferred unification group for the unification of the electroweak and the strong interactions. This is discussed in section 3. In section 4 we discuss an alternative possibility for a unifying group, i.e., $E_6$. The flavor puzzle which relates to the hierarchy of quark and lepton masses and mixings is discussed in section 5. In section 6 we discuss supergravity grand unification which provides the modern framework for realistic analyses of grand unified models and allows one to extrapolate physics from the electroweak scale to the grand unification scale. Unification in strings is discussed in section 7. One of the hallmarks of grand unified models is the prediction for the existence of monopoles and we discuss it in section 8. We note that no signature of monopole has thus far been detected and confirmation of its existence remains an outstanding experimental question. Since grand unification implies quark-lepton unification, another important prediction of grand unification is the decay of the proton. We discuss the current status of proton stability in GUTs and strings in section 9. In section 10 we argue that the discovery of the Higgs boson mass at $\sim 125$ GeV lends further support for SUSY/SUGRA/Strings, and consequently for the discovery of sparticles at colliders. The role of future colliders for testing superunification is discussed in section 11. Conclusions are given in section 12.

2. $G(4, 2, 2)$ and $SU(5)$ unification

As mentioned in section 1 a significant step toward unification beyond the standard model was taken by Pati and Salam [2] in 1974 when they proposed an extension of the standard model gauge group to the group $G(4, 2, 2) \equiv SU(4) \otimes SU(2)_L \otimes SU(2)_R$. Here $(4, 2, 1)$ and $(\bar{4}, 1, 2)$ representations of $G(4, 2, 2)$ contain a full generation of quarks and leptons. Since quarks and leptons reside in the same multiplet, $G(4, 2, 2)$ represents unification of quarks and leptons. This phenomenon has a direct consequence in that it allows conversion of quarks into leptons and thus one might expect the proton to become unstable and decay. The feature above is shared by essentially all unified models and thus proton lifetime limits act as a strong constraint on unified models of particle unifications. We also note that in $(4, 2, 1) + (\bar{4}, 1, 2)$ one has one more particle, i.e., $\nu'$, which does not appear in the standard model. $\nu'$ enters in the so called seesaw mechanism [22] that gives mass to the neutrinos. In $G(4, 2, 2)$ the charge operator takes the form $Q_{em} = T_{3R} + T_{3L} + \frac{\nu - \nu'}{2}$, where $T_{3R}$ and $T_{3L}$ are the generators of $SU(2)_R$ and $SU(2)_L$ and $B$ and $L$ are baryon and lepton numbers. To break the $G(4, 2, 2)$ symmetry one introduces heavy Higgs representations $(4, 2, 1)_{h^2} + (\bar{4}, 1, 2)_{h^2}$. In this case when $\nu_R$ and $\nu_R'$ develop VEVs, i.e., $\langle \nu_R \rangle = \langle \nu_R' \rangle = M_{h^2}$, $G(4, 2, 2)$ breaks to the standard model gauge group and the charge operator takes the familiar form $Q_{em} = T_{3L} + Y$ where $Y = T_{3R} + \frac{\nu - \nu'}{2}$. The group $G(4, 2, 2)$ can be broken further down to $SU(3)_C \otimes U(1)_{em}$ by use of $(1, 2, 2)$ Higgs representation. A comprehensive review of $G(4, 2, 2)$ was recently given in [23].

$SU(5)$ [4] is the lowest rank unified group which can accommodate the standard model gauge group. Here a full generation of quarks and leptons can be accommodated in its representations $5 \oplus 10$. Under the standard model gauge group they decompose so that $5 = (3, 1, -\frac{1}{3}) \oplus (1, 2, \frac{2}{3})$ and $10 = (3, 2, \frac{1}{2}) \oplus (3, 1, -\frac{2}{3}) \oplus (1, 1, 1)$, where one identifies $(\nu_R, e_L)$ with $(1, 2, -\frac{1}{3})$, $\nu'_{R}$ with $(1, 1, 1)$, $(u_R, d_L)$ with $(3, 2, \frac{1}{2})$, $u'_{R}$ with $(3, 1, -\frac{1}{2})$ and $d'_{L}$ with $(3, 1, \frac{1}{2})$. The GUT symmetry is broken by a 24-plet of heavy field $\Sigma(i, j = 1, 2, \cdots, 5)$ which breaks $SU(5)$ down to the standard model gauge group. To break the symmetry further to the residual gauge group $SU(3)_C \otimes U(1)_{em}$ one introduces a 5-plet of Higgs $H$ in the non-supersymmetric case. Here one immediate issue concerns the so-called doublet–triplet problem, i.e., how to keep the doublet of the 5-plet of Higgs light while making the triplet of the 5-plet superheavy. A concrete way to see this problem is to consider the scalar potential $V = M^2 \text{Tr} \Sigma^2 + \frac{\lambda_3}{4} \text{Tr} (\Sigma^2)^2 + \frac{\lambda_3}{2} \text{Tr} (\text{det} \Sigma)^2 + \mu \text{Tr} (\Sigma^2) + \frac{1}{\lambda_3} \text{Tr} H^2 \Sigma^2 + \frac{1}{\lambda_3} \text{Tr} \Sigma^2 H^2 + \frac{1}{4} (H^2 H^2)^2$. In order to break $SU(5)$ down to the standard model gauge group we need to have the VEV formation of $\Sigma$ so that $(\Sigma) = \text{diag}(2, 2, 2, -3, -3)\nu$. Here the spontaneous breaking of the symmetry gives the constraint $M^2 + (7\lambda_3 + 30\lambda_4)\nu^2 - \frac{1}{2} \mu \nu^2 = 0$. The breaking generates a mass for the Higgs doublet which is superheavy whereas electroweak symmetry breaking requires that the Higgs doublet be light. In order to achieve a light Higgs doublet we need the constraint $10\lambda_3 + 3\lambda_4 = 0$. This constraint must be satisfied to one part in $10^{10}$, which is a high degree of fine tuning.

In $SU(5)$ GUT the hypercharge coupling $g_Y$ is related to the $SU(5)$ invariant coupling $g_5$ so that $g_Y = \sqrt{3/5} g_5$. Thus
SU(5) predicts the weak angle at the GUT scale so that $\sin^2 \theta_W = g_\tau^2 / (g_\tau^2 + g_\nu^2) = 3/8$, where we set $g_\tau = g_\nu$. One of the problems of the minimal SU(5) model is that it generates undesirable relationships among the quark and the lepton masses. Thus consider the SU(5) Yukawa couplings $L_Y = h_d \psi_u \psi_i^c H^i + h_s \epsilon_{ijk \ell mn} \bar{\psi}_u \psi_i^c \psi_j^c H^m + h_c$, where $\psi_u$ and $\psi_i$ are the 5-plet and 10-plet of fermions. In the above we have suppressed the generation index. More explicitly each generation will have their own Yukawa couplings $h_d$ and $h_s$. After spontaneous breaking when the Higgs doublet gets a VEV one finds the following mass relation $m_u = m_d$, $m_s = m_s$, $m_t = m_t$. For the first two generations one has at the GUT scale the equality $m_u / m_d = m_d / m_s$. These ratios are independent of the scale to one loop order and they hold at the electroweak scale to this order. However, the relation is badly in conflict with data. For the third generation $m_b = m_t$ also does not quite work for the non-supersymmetric case although it does for the supersymmetric case. This means that the minimal SU(5) model is not sufficient to explain the flavor structure of the standard model. In SU(5) there are 24 gauge bosons of which 12 are the gauge bosons of the standard model consisting of the gluon, $W^\pm$, $Z$, $\gamma$. The remaining gauge bosons are superheavy leptons-quarks ($X^{1/3}$, $Y^{1/3}$), ($X^{-4/3}$, $Y^{-1/3}$) where the super-scripts give the charge assignments. However, non-supersymmetric SU(5) does not produce unification of gauge coupling constants consistent with LEP data.

We turn now to the supersymmetric version of SU(5) [24]. Here the superpotential for the minimal supersymmetric SU(5) is given by $W = \lambda_1 [1/5 \text{Tr} \Sigma^3 + 1/2 \text{Tr} \Sigma^2] + \lambda_2 H_u |\Sigma|^2 + 2 M^\delta |\Sigma|^3 H_d^\delta + f_1 H_u \cdot 5 \cdot 10 + f_2 H_2 \cdot 10 \cdot 10$. The GUT symmetry breaks as in the non-supersymmetric case with $\Sigma$ developing a VEV. Here as in the non-supersymmetric case a fine tuning is needed to get a light doublet. Specifically under the constraint $M^\delta = M$ the Higgs doublets are light while the Higgs triplets are superheavy. While a fine tuning is needed in the minimal SU(5) model to get light Higgs doublets, it is possible to achieve the doublet–triplet splitting without fine tuning in non-minimal or extended SU(5) models. One example is the so called missing partner mechanism [25, 26] and the other flipped SU(5) $\otimes U(1)$ model [27]. In the missing partner mechanism one uses a 75-plet representation to break the GUT symmetry and additionally 50 + 50 representations to give masses to the color triplets. The superpotential involving the 50, 50 and 75 has the form $W_5 = \lambda_1 \Phi_{ik}^m \Phi_{lm}^m H_d^i + \lambda_2 \Phi_{lm}^m \Phi_{ik}^m H_u^i + M \Phi_{ik}^m \Phi_{lm}^m + W_5(\Sigma)$, where $50 = \Phi_{ik}^m$, 50 = $\overline{\Phi}^{lm}_m$, and 75 = $\Sigma_{ik}^m$. The superpotential $W_5(\Sigma)$ is chosen to allow a GUT symmetry breaking where $\Sigma_{ik}^m$ develops a non-vanishing VEV of $O(M)$. The SU(3) $\otimes$ SU(2) $\otimes$ $U(1)$ decomposition of the 75-plet contains the singlet field $(1, 1, 0)$ and its VEV formation breaks the SU(5) GUT symmetry. The Higgs fields 5 and 50-plet have the SU(3) $\otimes$ SU(2) $\otimes$ $U(1)$ decomposition so that 5 = $(1, 2)(3 \oplus 3, 1)$ and 50 = $(1, 1)(-12) \oplus (3, 1)(-2) \oplus (3, 2)(-7) \oplus (6, 3)(-2) \oplus (6, 1)(8) \oplus (8, 2)(3)$. Here it is interesting to note that the 5-plet contains a Higgs triplet and similarly 5 contains an anti-triplet, while 50-plet contains a Higgs triplet and similarly 50 contains a Higgs anti-triplet. Thus the interaction 5 $\otimes$ 50 $\otimes$ 75 gives superheavy masses to the triplets and anti-triplet by matching the triplet in 5 to the anti-triplet in 50 and similar tying the anti-triplet in 5 with the triplet in 50. On the other hand, 50 and 50 do not have Higgs doublets and thus the Higgs doublets of 5 and 50 remain light.

3. SO(10) unification

The group SO(10) as the framework for grand unification appears preferred over SU(5). The group SO(10) contains both $G(4, 2, 2)$ and SU(5) $\otimes U(1)$ as subgroups, i.e., SO(10) has the branchings SO(10) $\rightarrow$ SU(4)c $\otimes$ SU(2)hc $\otimes$ SU(2)hc and SO(10) $\rightarrow$ SU(5) $\otimes$ U(1). It possesses a spinor representation which is $2^\delta = 32$ dimensional and which splits into $16 \otimes \overline{16}$. A full generation of quarks and leptons can be accommodated in a single 16-plet representation. Thus the 16-plet has the decomposition in SU(5) $\otimes U(1)$ so that $16 = (10)(-1) \oplus (5)(3) \oplus (1)(-5)$. As noted the combination 5 $\oplus$ 10 in SU(5) is anomaly free and further 1(-5) in the 16-plet decomposition is a right handed neutrino which is a singlet of the standard model gauge group and thus the 16-plet of matter in SO(10) is anomaly free. The absence of anomaly in this case is the consequence of a more general result for SO(N) gauge theories. Thus in general anomalies
arise due to the non-vanishing of the trace over the product of three group generators in some group representation \( \text{Tr}(T_a T_b T_c) \). For \( SO(10) \) one will have \( \text{Tr}((\Sigma_{\alpha}^g, \Sigma_{\alpha}^d, \Sigma_{\alpha}^u)) \). However, there is no invariant tensor to which the above quantity can be proportional which then automatically guarantees vanishing of the anomaly for \( SO(10) \). This analysis extends to other \( SO(N) \) groups. One exception is \( SO(6) \) where there does exist a six index invariant tensor \( \epsilon_{\alpha \beta \gamma \delta \epsilon \zeta} \) and so in this case vanishing of the anomaly is not automatic.

The group \( SO(10) \) is rank 5 where as the standard model gauge group is rank 4. The rank of the group can be reduced by either using \( 16 \oplus 10 \) of Higgs fields or \( 126 \oplus 45 \) of Higgs. Since under \( SU(5) \otimes U(1) \) one has \( 16 \supset 1(\pm 5) \) we see that a VEV formation for the singlet will reduce the rank of the group. Similarly \( 126 \supset 1(\pm 10) \) under the above decomposition. Thus when the singlets in \( 16 \oplus 10 \) of Higgs or \( 126 \oplus 45 \) get VEVs, the \( SO(10) \) gauge symmetry will break reducing its rank. However, we still need to reduce the remaining group symmetry to the standard model gauge group. For this we need to have additional Higgs fields such as \( 45, 54, 210 \). Further to get the residual gauge group \( SU(3)_C \otimes U(1)_{em} \) we need to have 10-plet of Higgs fields. Thus the breaking of \( SO(10) \) down to \( SU(3)_C \otimes U(1)_{em} \) requires at least three sets of Higgs representations: one to reduce the rank, the second to break the rest of the gauge group to the standard model gauge group and then at least one 10-plet to break the electroweak symmetry. As discussed above one can do this by a combination of fields from the set: \( 10, 16 \oplus 10, 45, 54, 120, 126 \oplus 45, 210 \). To generate quark and lepton masses we need to couple two 16-plets of matter with Higgs fields. To see which Higgs fields couple we expand the product \( 16 \otimes 16 \) as a sum over the irreducible representations of \( SO(10) \). Here we have \( 16 \otimes 16 = 10 \oplus \bar{10} \oplus 126 \oplus \bar{126} \), where the \( s(t) \) refer to symmetric (anti-symmetric) under the interchange of the two 16-plets. The array of Higgs bosons available lead to a large number of possible \( SO(10) \) models. For some of the works utilizing large representations see [28].

As discussed above, the conventional models have the drawback that one needs several sets of Higgs fields to accomplish complete breaking. One recent suggestion to overcome this drawback is to use \( 144 \oplus 144 \) of Higgs. This combination can allow one to break \( SO(10) \) symmetry all the way to \( SU(3)_C \otimes U(1) \). This can be seen as follows: in \( SU(5) \otimes U(1) \) decomposition one finds that \( 144 \supset 24(\pm 5) \) which implies that spontaneous symmetry breaking which gives VEV to the standard model singlet of 24 would also reduce the rank of the group. Thus in one step one can break \( SO(10) \) gauge symmetry down to the standard model gauge group. Further, there exist several Higgs doublet in 144 which have the quantum numbers of the standard model Higgs and one may use fine tuning to make one of the Higgs doublets light which is needed to give masses to the quarks and the leptons. A VEV formation that breaks the \( SO(10) \) symmetry can be achieved by taking a combination of dimension two and dimension four terms in the potential so that \( V = -M^2 \text{Tr}(\Sigma^2) \frac{1}{2} \text{Tr}(\Sigma^2 \Sigma^2) + \frac{1}{2} \text{Tr}(\Sigma^2 \Sigma^2) + \frac{1}{2} \text{Tr}(\Sigma^2) + \frac{1}{2} \text{Tr}(\Sigma^2 \Sigma^2) + \frac{1}{2} \text{Tr}(\Sigma^2 \Sigma^2) + \frac{1}{2} \text{Tr}(\Sigma^2 \Sigma^2) + \frac{1}{2} \text{Tr}(\Sigma^2 \Sigma^2) \). Symmetry breaking allows for local minima which lead to the standard model gauge group for the vacuum structure so that \( (\Sigma) = (\Sigma^i) = v \text{diag}(2, 2, 2, -3, -3) \), where \( v^2 = M^2 / (\kappa_1 + \kappa_2) + 30(\kappa_1 + \kappa_2) \). This local minimum will be a global minimum for some some range of the parameters of the potential which implies that spontaneous symmetry breaking does indeed break \( SO(10) \) down to \( SU(3)_{\text{c}} \otimes SU(2)_{\text{b}} \otimes U(1)_{\text{y}} \). Identical conclusions can be arrived at for the case of supersymmetric \( SO(10) \) model. Here the potential will become the supersymmetric standard model gauge group.

As mentioned in section 2, grand unified models typically have the doublet–triplet problem. In \( SU(5) \) aside from fine tuning two other avenues to overcome this problem were mentioned, one being the missing partner mechanism and the other flipped \( SU(5) \otimes U(1) \). In \( SO(10) \) one early suggestion to resolve the doublet–triplet problem is the missing VEV method where using a 45-plet of Higgs fields, one breaks the \( SO(10) \) gauge symmetry in the \( (B - L) \)-preserving direction, which results in generation of Higgs triplet masses while the Higgs doublets from a 10-plet remain massless. Another possibility is a missing partner mechanism in \( SO(10) \) which is in spirit is similar to the one for \( SU(5) \). In [37] a missing partner mechanism for \( SO(10) \) was constructed and the heavy fields used were \( 126 \oplus 45 \) along with a set of light fields. The reason for this choice is to simulate the missing partner mechanism of \( SU(5) \) in the following way: \( 126 \oplus 45 \) are chosen because they contain \( 50 \oplus 50 \) of \( SU(5) \) and 210 is chosen because it contains the 75 of \( SU(5) \). This parallels the analysis in \( SU(5) \) and leads to a pair of light Higgs doublets and heavy Higgs triplets.

In [38] a more comprehensive analysis of the missing partner mechanism for \( SO(10) \) was given. Here in addition to the missing partner model consisting of (i) Heavy \( \{126 \oplus 210 \} + \text{Light } [2 \times 10 + 120] \) one has others: (ii) Heavy \( \{126 \oplus 45 \} + \text{Light } [10 + 120] \), (iii) Heavy \( \{126 \oplus 210 \} + \text{Light } [10 \times 120] \), (iv) Heavy \( \{560 \oplus 560 \} + \text{Light } [2 \times 10 + 320] \). Models (i), (ii) and (iii) are anchored in the heavy fields \( 126 + 210 \) since they contain the \( 50 \oplus 50 \) of \( SU(5) \). However, model (iv) is anchored by a pair of 560 \( \oplus 560 \) Higgs fields which also contain \( 50 \oplus 50 \) of \( SU(5) \). Interestingly these are the next higher dimensional representations in \( SO(10) \) after \( 126 \oplus 210 \) which contain an excess of color triplets over \( SU(2)_{\text{b}} \) doublets. Further, it turns out that \( 560 \) also contains the \( SU(5) \) representations \( 1(\pm 5) \oplus 24(\pm 5) \oplus 75(\pm 5) \). After spontaneous breaking these fields acquire VEVs. They reduce the
rank of the group from 5 to 4 and further break the gauge symmetry down to the symmetry of the standard model gauge group. That means that the $SO(10)$ gauge symmetry breaks to the standard model gauge group in one step. This is reminiscent of the unified Higgs case discussed earlier in this section where $144 \oplus \overline{144}$ break the $SO(10)$ gauge symmetry in one step.

Next we explain how the missing partner mechanism works in this case. To this end we exhibit the complete $SU(5) \otimes U(1)$ decomposition of 560. Here one has $560 = 1 \oplus 5 \oplus 10 \oplus 45 \oplus 545 \oplus 550 \oplus 70 \oplus 175 \oplus 405 \oplus 540 \oplus 570 \oplus 70$ for $SU(5) \otimes U(1)$. The light fields must be in $(2 \times 10 \oplus 320)$ representations. A very stringent condition for the missing partner mechanism to work is that all the exotic fields must become heavy and thus in the entire array of Higgs fields only a pair of Higgs doublet fields must survive and remain light. To exhibit that this indeed is the case let us consider the $SU(5) \otimes U(1)$ decomposition of 320 and of 10’s. For the 320 we have $320 = 5(2) + 5(-2) + 40(6) + 45(2) + 45(-2) + 70(2) + 70(-2)$, and for the 10-plet we have $10 = 5(2) + 3(-2)$. We also note that 45 of $SU(5)$ under $SU(3)_C \otimes SU(2)_L \times U(1)$ has the decomposition so that $45 = (1, 2) (3) \oplus (3, 1) (-2) \oplus (3, 3) (-2) \oplus (\overline{3}, 1) (8) \oplus (\overline{3}, 2) (-7) \oplus (6, 1) (-2) \oplus (8, 2) (3)$, which exhibits the doublet/triplet content of the 45-plet. One now finds that the exotic light modes become superheavy by mixing with the heavy exotics in $560 \oplus \overline{560}$ and only a pair of light Higgs doublets remain. Also remarkably the gauge group $SO(10)$ breaks directly to the standard model gauge group right at the GUT scale so that we have one step breaking of the gauge symmetry.

Higher dimensional operators appear in effective theories and allow one to explore the nature of physics beyond the standard model. They have been explored in significant depth in the literature. These operators also include the ones that violate $B - L$. Such operators appear in $n - \bar{n}$ oscillations and more recently they have gained further attention as they may be helpful in generating baryogenesis [39]. It is interesting to investigate the type of $B - L$ violating operators that arise in grand unified theories. The minimal $SU(5)$ grand unification under the assumption of $R$ parity conservation and renormalizable interactions does not possess any $B - L$ violating operators. However, $SO(10)$ models do generate $B - L$ violation. Recently an analysis of $B - L = -2$ operators has been given in [40] for a class of $SO(10)$ models where there is doublet–triplet splitting using the missing partner mechanism [38] (for related work see [41]). The $\Delta(B - L) = \pm 2$ operators lead to some unconventional proton decay modes such as $p \rightarrow \pi^+ \nu, n \rightarrow e^+ K^+$ and $p \rightarrow e^+ \pi^-$. The $\Delta(B - L) = \pm 2$ operators also allow for GUT scale baryogenesis. The baryogenesis produced by such operators is not wiped out by sphaleron processes and survives at low temperatures [39].

4. $E_6$ unification

Among the exceptional groups only $E_6$, $E_7$, $E_8$ are possible candidates for unification. However, $E_7$, $E_8$ are eliminated as they do not have chiral representations which leaves $E_6$ as the only possible candidate for unification among the exceptional groups. The lowest representation of $E_6$ is 27-plet and $27 \otimes 27 = 1 \oplus 78 \oplus 650$ where 78-plet is the adjoint representation. One notes that $27 \otimes 27 = \overline{3} \oplus 3 \overline{3} \oplus 3 \overline{3}$. This result leads to the remarkable feature of $E_6$ models that the only cubic interaction one can have for 27 is $W_{27} = \lambda \, 27 \otimes 27$. $E_6$ has many possible breaking schemes such as $E_6 \rightarrow SO(10) \otimes U(1)$, $SU(3)_C \otimes SU(2)_L \otimes SU(2)_R$, $SU(6) \otimes SU(2)$. One of the most investigated symmetry breaking schemes is $SU(3)_C$. One can exhibit the spectrum of $E_6$ in terms of representations of $SU(3)_C$ so that $27 = (1, 3, 3) \oplus (\overline{3}, 1, 3)$ and the particle content of the model consists of nonets of leptons, quarks and conjugate quarks where $L = (1, 3, 3)$, $Q = (\overline{3}, 3, 1)$ and $Q' = (\overline{3}, 1, 3)$. The symmetry of the group can be broken by the combination of Higgs fields $27 \otimes 27 \rightarrow \overline{3} \oplus 3 \overline{3} \oplus 3 \overline{3}$. Extensive analyses exist in the literature and it is shown that with appropriate symmetry breaking schemes $E_6$ can produce a low energy theory consistent with data (see, e.g., [42, 43]). Investigation of $E_6$ as the unification group within string theory has a long history. In models of this type $E_6$ is broken down to the standard model gauge group by a combination of flux breaking and breaking by VEVs of Higgs fields. In one breaking sequence one has $E_6 \rightarrow SO(10) \otimes U(1)_Y$, $SO(10) \rightarrow SU(5) \otimes U(1)_Y$, $SO(10) \rightarrow SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. Phenomenology of such breaking and of other scenarios leads to some distinctive signatures. More recently $E_6$ unification has also been investigated within F-theory (see, e.g., [44] and the references therein).

5. The flavor puzzle

The observed structure of quarks and leptons exhibits some very interesting properties. There are at least two broad features which may be summarized as follows: (i) the quarks and leptons show a hierarchy in masses for different flavors; (ii) the mixing among quarks in going from flavor to mass diagonal basis is small, while the mixing among neutrinos is large. The constraints on the neutrino mixing come from the solar and the atmospheric neutrino oscillation data which is sensitive only to the differences of neutrino mass squares, i.e., $\Delta m_{32}^2 = |m_3^2 - m_2^2|$. A fit to the neutrino data gives [45] $\Delta m_{32}^2 = (5.4 - 9.5) \times 10^{-5}$ eV$^2$, $\Delta m_{21}^2 = (1.4 - 3.7) \times 10^{-3}$ eV$^2$, $\sin^2 \theta_{12} = (0.23 - 0.39)$, $\sin^2 \theta_{23} = (0.31 - 0.72)$, $\sin^2 \theta_{13} < 0.054$, where $\Delta m_{21}^2 = |m_2^2 - m_1^2|$, $\Delta m_{32}^2 = |m_3^2 - m_2^2|$. Here the mixing angles $\theta_{12}$ and $\theta_{23}$ are large while $\theta_{13}$ is small. The constraints on the absolute value of the masses themselves arise from neutrino less double beta decay and from cosmology. Thus from cosmology one has $\sum_i |m_i| < (0.7 - 1) eV$.

One avenue to decode the flavor structure is to assume that while it looks complex at low scales, there could be...
simplicity at high scales. An example of this is the ratio of the mass of the $b$ quark versus the mass of the $\tau$ lepton which is approximately three at low energy but one can explain this ratio starting from the equality of the $b$ and $\tau$ Yukawa couplings at the GUT scale. This result holds in supergravity GUTs but not in non-supersymmetric unification. In fact, one can also explain the ratio of the top mass to the $b$-quark mass starting with equality of the bottom and the top Yukawa coupling at the GUT scale if one assumes large $\tan \beta$ [49]. One of the early works in correlating low energy data on quark and lepton flavors with textures at the GUT scale is by Georgi and Jarlskog (GJ) [46]. Aside from the GJ textures there are several other textures that can generate the desired flavor structure at low energy (for early works see [47, 48]).

The question then is what are the underlying models which can produce the desired textures. One possibility is that they arise from cubic and higher dimensional operators. In the analysis of [50] it was shown that in $SU(5)$, higher dimensions operators in an expansion in $\Sigma/M$, where $\Sigma$ is the 24-dimensional field that breaks the GUT symmetry, can produce the GJ textures. This analysis also revealed that similar textures exist in dimension five operators which enter in proton decay, which are different from the ones that appear in the Yukawa couplings.

Within renormalizable interactions one approach is to expand the superpotential in all allowed couplings and try to fit the data with the assumed set of couplings. For example, for $SO(10)$ one could use the Higgs fields $10$, $120$, $\overline{126}$ and assume general form of Yukawas consistent with these couplings, use renormalization group evolution from the GUT scale down to low energy and fit the quark and lepton masses and mixings [51]. Other approaches involve flavor symmetries. One of the common flavor symmetry used is $S_3$ of which there are a large number of variants, see, e.g., [52, 53] and the references therein.

The flavor structure can be understood in another way in a class of $SO(10)$ models within the unified Higgs sector [29, 35]. As discussed in section 3 in unified Higgs models, one uses a single pair of vector–spinor representation $144 + \overline{144}$ which breaks the $SO(10)$ gauge symmetry to the standard model gauge group. Here the matter fields can couple to the Higgs sector only at the quartic level, i.e., the interaction involves terms such as $16 \cdot 16 \cdot 144 \cdot 144$. This must be suppressed by a heavy mass. For this reason the Yukawa couplings arising from this interaction are naturally small and can provide an appropriate description of the masses and mixings of the first two generations. However, for the third generation, one needs cubic couplings and one can obtain much larger couplings in a natural way by including additional matter in $10$, $45$ and $120$. The additional matter representations allow one to have couplings of the type $16 \cdot 144 \cdot 10$, $16 \cdot 144 \cdot 45$, and $16 \cdot 144 \cdot 120$. Specifically using $16 \cdot 144 \cdot 45$, and $16 \cdot 144 \cdot 120$ one finds that cubic couplings of size appropriate for the third generation arise [35]. One can obtain $b - \tau$ as well as $b - t$ unification even at low values of $\tan \beta$. The formalism also correctly generates the charged lepton and neutrino masses which arise from a type I see-saw mechanism [35]. There are also a variety of other approaches. For some recent ones see [54].

6. Supergravity grand unification

As mentioned in the introduction, supersymmetry, which is a global symmetry, cannot be broken in a phenomenologically viable fashion. Supergravity grand unification overcomes this problem and allows one to build realistic models with spontaneous breaking of supersymmetry which lead to sparticles with calculable masses which can be searched for at colliders. The formulation of supergravity grand unification utilizes the framework of applied supergravity where one couples an arbitrary number of $N = 1$ chiral superfields, which we denote by $Z$, to $N = 1$ vector superfield belonging to the adjoint representation of the gauge group and then couple them to $N = 1$ supergravity [12, 57–59].

The applied $N = 1$ supergravity lagrangian depends on three arbitrary functions consisting of the superpotential $W (Z)$, the Kahler potential potential $K (Z, Z^\dagger)$ and the gauge kinetic energy function $f_{\alpha \beta}$ which transforms like the symmetric product of adjoint representations. However, before discussing the breaking of supersymmetry we begin by considering the breaking of a grand unified group with no breaking of supersymmetry in the framework of supergravity grand unification. As an illustration let us consider the simplest case where one has a 24-plet of $SU(5)$ field $S_{\alpha \beta}$ $(a, b = 1 - 5)$ and a superpotential that is given by $W = \left[ \frac{2}{3} M Tr \Sigma^2 + \frac{1}{7} Tr \Sigma^3 \right]$. After a spontaneous breaking of $SU(5)$ occurs and $\Sigma_{\alpha \beta}$ develops a VEV, one has three possibilities: (i) $\langle \Sigma_{\alpha \beta} \rangle = 0$, (ii) $\langle \Sigma_{\alpha \beta} \rangle = \frac{2}{3} M \delta_{\alpha \beta} + 5 \delta_{\alpha 5} \delta_{\beta 5}$, and (iii) $\langle \Sigma_{\alpha \beta} \rangle = \frac{2}{3} M \delta_{\alpha \beta} + 5 \delta_{\alpha 5} \delta_{\beta 5}$]. Here (i) gives no breaking of gauge symmetry, (ii) gives the breaking of $SU(5)$ to $SU(4) \otimes U(1)$, and (iii) gives the breaking to $SU(3) \otimes SU(2) \otimes U(1)$.

In global supersymmetry these are flat directions and the potential vanishes for all the three cases. Thus all three vacua are degenerate. However, for the case of supergravity the potential does not vanish and one finds that the potential at the minimum is given by [12]$ V_0 (\Sigma_0, \Sigma_{ab}^0) = -\frac{1}{3} b^2 \varepsilon_1 (\Sigma_0, \Sigma_{ab}^0) W (\Sigma_0)^2$. We note now that unlike the case of global supersymmetry the potential is no longer degenerate for the three vacua [12, 55]. Suppose we add a term to the superpotential and make the vacuum energy for one case vanish. In that case it is easy to check that the vacuum energy for other cases will be negative, i.e., that the vacua will be anti-deSitter. This would imply that the Minkowskian vacuum would be unstable in each case. However, it turns out that vacuum stability is helped by gravity [56]. In fact it has been shown that the Minkowski vacuum will be stable against any finite size perturbations [55]. It should be noted that vacuum degeneracy is not lifted in all cases when gauge symmetry breaks even in the presence of gravity.

We turn now to the breaking of supersymmetry. In supergravity grand unification the breaking of supersymmetry can be generated by a superhiggs potential similar to the breaking of a gauge symmetry by a Higgs potential. A general form of the superpotential for the superhiggs is given by $W_{SH}(Z) = m^2 \kappa^{-1} f_1(\kappa Z)$ where the function $f_1(\kappa Z)$ depends on the dimensionless product $\kappa Z$. Here the breaking of supersymmetry gives $\langle Z \rangle \sim O(\kappa^{-1})$ and $f_1(\kappa Z) \sim O(1)$. The
gravitino mass in this case is $m_{3/2} \sim O(\alpha m^2)$. If we take $m$ to be intermediate scale of order $10^{10-12}$ GeV, then $m_{3/2}$ lies in the range 1–100 TeV. In early works a simple choice for $W_{3\text{H}}$ was made, i.e., $W_{3\text{H}} = m^2 (c + B)$. The VEV of the field $Z$ which breaks supersymmetry has no direct interaction with the visible sector since $\langle Z \rangle \sim \kappa^{-1}$ and thus any direct interaction between the superhiggs field $Z$ and the visible sector fields would generate a mass for the visible sector fields which would be $O(\kappa^{-1})$.

To shield the visible sector from such large mass growths the breaking of supersymmetry is communicated to the visible sector by gravity mediation. A simple way to see this is to write the superpotential including superhiggs and matter fields so that [12, 60] $W_{\text{tot}} = W(Z_u) + W_{3\text{H}}(Z)$, where $W(Z_u)$ depends on the matter fields $Z_u$ and $W_{3\text{H}}(Z)$ depends only on the superhiggs field $Z$ which breaks supersymmetry. The two are, however, connected via the supergravity scalar potential [12, 58], i.e.,

$$V = e^{2k} \left[ (K^{-1})^{\mu \nu} (\bar{\Omega}_{\bar{\nu}} + \kappa^2 K_{\bar{\nu}} W) \left( \bar{\Omega}_{\mu} + \kappa^2 K_{\mu} W \right) - 3\kappa^2 |W|^2 \right] + V_0,$$

where $V_0$ is the $D$-term potential. Integrating out the superhiggs field, one finds that the low energy theory in the visible sector does contain soft breaking generated by gravity mediation which are size $m \sim \sim m^2$ and thus the soft breaking is free of the Planck mass. However, supergravity models with grand unification contain another heavy mass, i.e., the GUT mass $M_G$ in addition to the Planck mass. Including $M_G$ in the analysis in a grand unified supergravity model brings in another type of hierarchy problem, i.e., one would have mass terms of the form $m_m M_G \cdots \alpha^m M_G$, where $\alpha \equiv (\kappa M_G)$. A remarkable aspect of supergravity unified models is that all items of the type above cancel or vanish [12]. In spontaneously broken supergravity the sum rule for masses is given by [57, 59]

$$\sum_{i=0}^N (-1)^i (2J + 1) m_i^2 = 2(N - 1) m_{\text{eff}}^2,$$

where $N$ is the number of chiral superfields.

In the universal supergravity models the effective low energy theory after integration over the heavy fields is thus given by [12]

$$V_{\text{eff}} = \left[ \frac{\Omega^2}{\bar{\Omega}} \right] + m_0^2 Z_u Z_u + (B_0 W^{(2)} + A_0 W^{(3)} + \text{h.c.}) + D \text{ term},$$

where $W^{(2)}$ and $W^{(3)}$ are parts of the superpotential which are quadratic and trilinear in the scalar fields. Here the soft parameters consist of the universal scalar mass $m_0$, and the dimensioned parameters $A_0$ and $B_0$. In the presence of non-universality, one may have flavor dependence on the scalar masses as well as on the trilinear couplings. In the analysis above the vacuum energy has been adjusted to vanish. We note in passing that there exist models in which this is not the case and the minimum has a large negative energy [61]. The class of supergravity unification is rather large, as it encompasses a vast variety of particle physics models depending on the choice of the Kahler potential, the superpotential and the gauge kinetic energy function. Further, strings in low energy limit are describable by supergravity models (for a review see [11]). Similarly M-theory [62] in low energy limit can be described by 11D supergravity [63] which can be reduced further to 4D. Some examples of low energy string based supergravity models are KKLT [64] and the Large Volume Scenario [65].

Soft terms also arise for gaugino masses. One such source is via loops through their gauge interactions with the fields in the heavy sector [66, 67]. Thus the gaugino masses $\tilde{m}_i$, $i = 1, 2, 3$ for the $U(1), SU(2), SU(3)$ gauge groups are given by $\tilde{m}_i = \frac{1}{2} m_{3/2} CD(R)/D(A)$, where $D(R)$ is the dimensionality of the exchanged representation, $D(A)$ is the dimensionality of the adjoint representation and $C$ is the quadratic Casimir of the fields that contribute. One can also generate gaugino masses at the tree level via a field dependent gauge kinetic energy function $f_{\alpha \beta}$ after spontaneous breaking of supersymmetry and one has $L_{\alpha \beta} = -(1/2) m_{3/2} \tilde{X}^\lambda \tilde{\lambda}$. Here $f_{\alpha \beta}$ transforms like the symmetric product of adjoint representations. For the case when $f_{\alpha \beta}$ transforms like a singlet of the gauge group $L_{\alpha \beta} = -(1/2) m_{3/2} \tilde{X}^\lambda \tilde{\lambda}$. In general, one would have non-universality of gaugino masses when $f_{\alpha \beta}$ does not transform as a pure singlet [68–71]. SUSY breaking can also arise from gaugino condensation such that $\langle \lambda^\alpha \rangle = 0$ [72]. This kind of breaking is often used in string model building. The Higgs boson mass measurement at 125 GeV gives strong support to supergravity models with gravity breaking (see, e.g., [73]).

Astrophysical evidence suggests the existence of dark matter in great abundance in our universe. Thus up to 95% of the Universe may be constituted of dark matter or dark energy. Here we will focus on dark matter. Some leading candidates for dark matter include the weakly interacting massive particles (WIMPs), the extra weakly interacting particles [74], and axions among many others. In supergravity models several neutral particles exist which could be possible candidates for dark matter, such as the neutralino, the sneutrino, and the gravitino [75, 76]. We will focus here on neutralino type WIMP which is an odd $R$ parity particle. It has been shown that such a particle turns out to be the lightest supersymmetric particle in a large part of the parameter space of supergravity models [77] and if $R$-parity is conserved, it becomes a candidate for dark matter. Indeed the neutralino was proposed as a possible candidate for dark matter soon after the formulation of supergravity grand unified models [78].

The relic density of dark matter turns out to be an important constraint on model building. In supergravity grand unified models under the constraint that the weak SUSY scale is high perhaps lying in the several TeV region, the neutralino turns out to be most often a bino. The annihilation cross section for bino-like neutralino is rather small which implies that the neutralinos cannot efficiently annihilate to produce the desired relic density consistent with the data from WMAP [79] and from PLANCK [80] experiments. In this case we need coannihilation [81–84]. Coannihilation is a process in which one or more sparticles other than the neutralino enter in the annihilation process modifying the Boltzman equation that controls the relic density. Analysis shows that if the next to the lightest supersymmetric particle (NLSP) lies close to the LSP in mass then there can be a significant enhancement in the annihilation of the LSP allowing one to satisfy the relic density constraints consistent with the WMAP and the
As discussed in section 6, $N = 1$ supergravity grand unified models with a hidden sector lead to the breaking of supersymmetry in a phenomenologically viable manner and with three generations of quarks and leptons allow an extrapolation of physics from the electroweak scale to the grand unification scale. Above this scale one expects quantum gravity effects to become important. The next step then is to look for a theory of quantum gravity whose low energy limit is supergravity grand unification. Possible candidates are string theories which come in several varieties: Type I, Type IIA, Type IIB, heterotic $SO(32)$ and heterotic $E_8 \otimes E_8$ [90]. The Type I and Type II strings contain D branes. The D branes can support gauge groups and chiral matter can exist at the intersection of D branes. The various string Types can arise from the so called ‘missing satellite’ problem where CDM predicts too many dwarf galaxies. More detailed analyses indicate that both problems could be solved by inclusion of complex dynamics and baryonic physics [86]. Aside from complex dynamics and baryonic matter, another possibility involves particle physics models such as (repsilive) self-interactions or a dark particle of de Broglie wavelength of $\sim 1$ kpc and a mass which lies in the range of $10^{-21}$–$10^{-22}$ eV [87].

7. Unification in strings

As discussed in section 6, $N = 1$ supergravity grand unified models with a hidden sector lead to the breaking of supersymmetry in a phenomenologically viable manner and with three generations of quarks and leptons allow an extrapolation of physics from the electroweak scale to the grand unification scale. Above this scale one expects quantum gravity effects to become important. The next step then is to look for a theory of quantum gravity whose low energy limit is supergravity grand unification. Possible candidates are string theories which come in several varieties: Type I, Type IIA, Type IIB, heterotic $SO(32)$ and heterotic $E_8 \otimes E_8$ [90]. The Type I and Type II strings contain D branes. The D branes can support gauge groups and chiral matter can exist at the intersection of D branes. The various string Types can arise from the so called M-theory whose low energy limit is 11D supergravity. Most of the model building in string theory has occurred in heterotic $E_8 \otimes E_8$/Horava–Witten theory [91], where the Horava–Witten theory arising from the low energy limit of M theory on $\mathbb{R}^{10} \times S^1/\mathbb{Z}_2$ may be viewed as the strong coupling limit of $E_8 \otimes E_8$ and Type IIB/F theory, where F theory [92] may be viewed as the strong coupling limit of Type IIB strings. One can also compactify M-theory on other manifolds such as on a manifold $X$ of $G_2$ holonomy [93–95]. One problem in working with string models is that they possess a huge number of vacuum states [96], as many as $10^{500}$, and the search for the right vacuum state that describes our universe is a daunting task.

One possible way to proceed then is to thin out the landscape of vacuum states by imposing phenomenological constraints. The most obvious ones are the emergence of groups which support chiral matter which limits the groups to $SU(N)$, $SO(4N + 2)$, $N \geq 1$ and $E_6$, after reduction to four dimensions, an $N = 1$ supersymmetry in 4D, a hidden sector that allows for breaking of supersymmetry, and three generation of chiral fermions which correspond to the three observed generations of quarks and leptons. For the emergence of supergravity grand unified model it is also of relevance that the grand unification scale $M_G$ emerge in some natural way from the Planck scale where $M_G$ is expected to be close to the scale where the 10-dimensional theory reduces to four dimensions.

The $E_8 \otimes E_8$ heterotic string was one of the first to be investigated at significant length [97]. After compactification to four dimensions it has a rank which can be up to 22. This allows for many possibilities for model building which have been pursued in the context of using free field constructions, orbifolds, and Calabi–Yau compactifications (for a sample of the early phenomenological analyses see [98]). Most of the analyses are within Kac–Moody level 1. Here one can achieve unified groups such as [99] $SU(5)$, $SO(10)$ or $E(6)$. However, one does not have scalars in the adjoint representation to break the gauge symmetry. At level 2, adjoint scalars are achievable but three massless generations are not easy to get. At level 3, it is possible to have scalars in the adjoint representation and also 3 massless generations. However, here one finds that the quark-lepton textures are rank 3 and thus difficult to get realistic quark-lepton masses [100]. In addition to the heterotic string constructions, a large number of string model constructions have since appeared, and a significant number within Type II (for reviews see [101–104]). Of course in such models which have an effective theory with $N = 1$ supersymmetry, one would need to break supersymmetry to make contact with observable physics (see, e.g., [105]).

We note in passing that while it is desirable that grand unification arise from strings (for a recent review see [106]) it is not essential that it do so since string theory is already a unified theory and it is not necessary for us to insist on grand unification in the effective low energy theory. Rather we may have the standard model gauge group emerging directly from string theory without going through grand unification. In this case we will have $g^2 k_i = g_{\text{string}} = 8\pi G_N / \alpha'$, where $G_N$ is the Newtonian constant, $\alpha'$ is the Regge slope, and $k_i$ are the Kac–Moody levels of the subgroups. There are positive integers for non-abelian gauge groups [107] while for $U(1)$ the normalization of $k$ is arbitrary.

Of course string theory is supposed to unify gravity along with other forces, and one may look at the evolution of the dimensionless parameter $\alpha_{GR} = G_N E^2$ along with the other three couplings (for a review see [108]) In the normal evolution of the gauge and gravity couplings using the spectrum of MSSM, one finds that while the gauge couplings do unify at the scale $M_G \sim 10^{16}$ GeV, the gravity coupling $\alpha_{GR}$ does not. There are several possible ways to address this lack of unification. One possibility is that as we evolve the couplings above the scale $M_G$, $\alpha_G$ and $\alpha_{GR}$ will unify. Another possible way is that at some scale below $M_G$, a new dimension of space opens up. Here if matter resides on the 4 dimensional wall while gravity resides in the bulk, the evolution of $\alpha_{GR}$ will be much rapid and there may be unification
of gauge and gravity couplings at a scale much below the Planck scale. A further modification of this idea is strings where the string scale is lowered to lie at the TeV scale. In this case gravity can get strong at a much lower scale and unification of gauge and gravity coupling can occur in the few TeV region [109–111]. More recent progress in model building has come from F-theory which as noted earlier can be viewed as a strong coupling limit of Type IIB string. For some recent work in F-theory model building see, e.g. [44, 112–115].

8. Monopoles

It was shown by Dirac [116] that in Maxwell electrodynamics the existence of a monopole of magnetic charge $g$ would imply a quantization of the electric charge so that e.g. $\frac{q}{n\hbar}$. In unified theories while $SU(2) \times U(1)$ does not possess a magnetic monopole, $SO(3)$ does and it arises as a consequence of solution to the field equations. The quantization condition in this case is e.g. $\frac{q}{n\hbar}$ which is the Schwinger quantization [117]. Grand unified theories also possess monopoles and they appear again as solutions to field equations. Unlike the Maxwell electrodynamics where the monopole may or may not exist, in grand unified theories where a $U(1)$ arises as a reduced symmetry, the monopole is a necessity and a prediction. However, the monopoles in grand unified theories would be superheavy with a mass of size the GUT scale. One problem associated with these monopoles is that monopoles produced in the early universe would contribute a matter density in excess of the critical relic density which would over close the Universe. Inflationary cosmology solves this problem. In some models the magnetic monopoles can be much lighter. Monopoles also appear in intersecting D-brane models where they appear along with color singlet fractionally charged states [118]. For the current experimental status of magnetic monopoles see [119].

9. Proton decay in GUTs and strings

One of the hallmarks of most unified models is the prediction that proton will decay. It is also a possible discriminant of models based on GUTs versus strings. One of the predictions in both supersymmetric and non-supersymmetric grand unification is the proton decay mode $p \rightarrow e^+\pi^0$. This mode arises from dimension six operators and involves the exchange of leptoquarks (for reviews see [120–122]). A rough estimate of the decay width is $\Gamma(p \rightarrow e^+\pi^0) \simeq \frac{\alpha_G^2}{16\pi} \frac{m^2}{M^2}$, where $\alpha_G = g_4^2/4\pi$ with $g_4$ being the unified coupling constant, and $M$ the lepto-quark mass. It leads to a partial lifetime of $\tau(p \rightarrow e^+\pi^0) \simeq 10^{36.1} \text{yr}$. This mode allows the possibility of discrimination among models based on GUTs versus those based on strings. Thus a generic analysis of D brane models allows only $10^3 \cdot \prod SU(5)$ type dimension six operators which leads to the decay $p \rightarrow \pi^0 e^0$ [123]. In $SU(5)$ grand unification, one has in addition the operator $10 \cdot \prod SU(5)$ which allows $p \rightarrow \pi^0 e^0$. Further, generic $D$ brane models do not allow the decay $p \rightarrow \pi^0 \nu$ while $SU(5)$ grand unification does. It has been pointed out, however, that special regions of intersecting $D$ brane models exist which allow the operator $10 \cdot \prod SU(5)$ and the purely stringy proton decay rate can be of the same order as the one from $SU(5)$ GUTs including the mode $p \rightarrow \pi^0 \nu$ [124]. The current experimental status of proton decay for this partial decay mode is the following: Superkamiokande gives the limit $\tau(p \rightarrow e^+\pi^0) > 2 \times 10^{34} \text{yr}$ [125] while in the future hyper-K is expected to achieve a sensitivity of $\tau(p \rightarrow e^+\pi^0) > 1 \times 10^{35} \text{yr}$ [126].

In supersymmetric unified models there are additional operators that can generate proton decay. Thus if $R$ parity is not conserved, proton decay can proceed with baryon and lepton number violating dimension four operators. Such a decay is too rapid and must be eliminated by imposition of $R$ parity conservation. In this case we still have baryon and lepton number violating dimension five operators arising from the exchange of Higgs triplets, which give rise to proton decay [120, 127]. Here the dominant decay mode is $p \rightarrow K^0 \bar{\nu}$ and could also be dangerous in terms of proton stability [128–135]. The SUSY decay modes depend sensitively on the sparticle spectrum as well as on CP phases [136, 137]. The current experimental limit from super-Kamiokande is $\tau(p \rightarrow \bar{\nu}K^0) > 4 \times 10^{32} \text{yr}$, while in the future it is expected that hyper-K may reach a sensitivity of $\tau(p \rightarrow \bar{\nu}K^0) > 2 \times 10^{34} \text{yr}$. It is known that proton decay lifetime from baryon and lepton number violating dimension five operators in SUSY GUTs with a low sparticle spectrum would lie below the current experimental limits. One possibility for stabilizing the proton is via a cancellation mechanism [120, 138]. The other possibility is via using a heavy sparticle spectrum which enters in the loop diagrams. Very roughly the proton decay from dimension five operators has the sparticle mass dependence of $m^2 \chi^+ \left/ m^2 \tilde{q} \right.$ where $\chi^+$ is the chargino and the $\tilde{q}$ is the squark. This means that a suppression of proton decay can be achieved with a large sfermion mass. As will be discussed in section 10, the discovery of the Higgs boson with a mass of $\sim 125 \text{GeV}$ points to a high SUSY scale and a high SUSY scale implies a larger proton decay lifetime for the SUSY modes. In fact one finds a direct correlation between the proton lifetime and the Higgs boson mass [139] which shows that the experimental lower limit on the proton lifetime for the SUSY mode can be easily met under the Higgs boson mass constraint.

10. A stronger case for SUSY/SUGRA after Higgs Boson discovery

The measurement of the Higgs boson at 125 GeV gives further support for supersymmetry. One reason for that is vacuum stability. For large field configurations where $h \gg \nu$ the Higgs potential is governed by the quartic term $V_5 \sim \frac{m^2}{4}h^4$. For vacuum stability $\lambda_{\text{eff}}$ must be positive. In the
standard model analyses done using NLO and NNLO corrections show that with a 125 GeV Higgs boson, the vacuum can be stable only up to about \(10^{10} - 10^{11}\) GeV [140]. For vacuum stability up to the Planck scale one needs \(m_{0} > 129.4 \pm 1.8\) GeV. Vacuum stability depends critically on the top mass. A larger top mass makes the vacuum more unstable. An advanced precision analysis [141] including two-loop matching, three-loop renormalization group evolution, and pure QCD corrections through four loops gives an upper bound on the top pole mass for the stability of the standard model vacuum up to the Planck mass scale of \(m_{t}^{\text{top}} = (171.54 \pm 0.30^{+0.26}_{-0.41})\) GeV, while the experimental value of the top is \(m_{t}^{\text{exp}} = (173.21 \pm 0.51 \pm 0.71)\) GeV. Though an improvement on previous analyses the stability of the standard model vacuum is still not fully guaranteed. In models based on supersymmetry with a Higgs mass of 125 GeV, the vacuum is stable up to the Planck scale. Of course one has to choose the parameter space of supergravity models appropriately.

As discussed in section 6 the Higgs boson mass measurement at 125 GeV [142, 143] indicates that the loop correction to the Higgs boson mass in supersymmetry is large. This in turn indicates that the weak SUSY scale, and specifically the scalar masses, must be large lying in the several TeV region [73, 144–146]. It turns out that these large scalar masses arise naturally in the radiative breaking of supersymmetry. Thus the radiative breaking in supersymmetry unified models obeys the relation (for a review see [147]) \(m_{0}^{2} + 2\mu^{2} = (1 - 3D)m_{\tilde{t}}^{2} + C(m_{0}, A_{0}, \tan \beta) \beta > 0\), where \(C\) is a polynomial in \(m_{0}\) and \(A_{0}\) and \(D\) depends on loop corrections, Yukawas and the weak SUSY scale. For the case when \(D < 1/3\) there is an upper bound on sparticles masses for fixed \(\mu\). This the case of elliptoidal geometry (EB) where for a given \(\mu\), one has an upper limit on how large the soft parameters can get. When \(D > 1/3\) the geometry becomes hyperbolic (HB) and for a fixed \(\mu\) the scalar masses get large [149].

The HB contains focal curves and focal surfaces (see Akula et al in [149]). The transition point between the two branches, i.e., between EB and HB, is \(D = 1/3\) referred to as the focal point. Here \(\mu\) essentially becomes independent of \(m_{0}\). Thus in general on the focal point, focal curves and focal surfaces, \(\mu\) can remain small while scalar masses get large. On HB the gauginos can be light and discoverable at the LHC even if the SUSY scale \(M_{3}\) is in the TeV region. Further, the weak SUSY scale is approximated by \(M_{3} \approx \sqrt{m_{\tilde{t}_{1}} m_{\tilde{t}_{2}}}\), where \(\tilde{t}_{1}, \tilde{t}_{2}\) are the stop masses. Thus \(m_{\tilde{t}_{1}}\) may be a few hundred GeV while \(m_{\tilde{t}_{2}}\) is order several TeV which leads to \(M_{3}\) in the TeV region [150]. This means that even for the case when the weak SUSY scale is large, one of the stops could have a mass that lies in the few hundred GeV region and may be discoverable. Of course, for non-universal supergravity models the sleptons can be much lighter than the squarks. Here we note that in searches for supersymmetry optimization of signal analysis beyond what is employed in simplified models is often necessary as important regions of the parameter space can otherwise be missed (see, e.g., [151–154]).

It is possible, however, that the weak scale of SUSY is much larger than previously thought and could be upwards of 10 TeV. A weak SUSY scale of this size is consistent with radiative breaking of the electroweak symmetry, the Higgs boson mass constraint and the relic density constraint [157]. It can also resolve the so called gravitino decay problem in supergravity and string theories. Thus an intrinsic element of supergravity unified models is the existence of a gravitino which is a supersymmetric partner of the graviton and becomes massive after spontaneous breaking of supersymmetry. The gravitino could be either stable or unstable. If it is stable it would be the lowest mass odd \(R\) parity object and thus contribute to dark matter. In this case one finds that the mass of the gravitino must be less than 1 KeV in order that it not over close the Universe. If the gravitino is not the LSP, it would be unstable and decay and there are strong constraints on the gravitino in this case. Here one needs to make certain that the gravitino which has only gravitational interactions does not decay too late, i.e., does not occur after the BBN time \((1 - 10^{5}) s\), which would upset the successes of the BBN. As already noted one of the implications of the Higgs boson measurement, is that the weak SUSY scale which is typically set by the gravitino mass is high lying in the several TeV region. However, the scalar masses though scaled by the gravitino mass need not be equal to the gravitino mass.

We note that supergravity models have a large landscape of soft parameters [155] which include non-universalities in the gaugino sector and as well as non-universalities in the matter and Higgs sectors [156]. First there could be non-universalities which are model dependent and split the scalar masses and second that renomalization group evolution has significant effect on the scalar masses when evolved down to the electroweak scale. Thus in general the scalar masses could be much lower than the gravitino mass. Further, the gaugino masses could be significantly smaller than the scalar masses. This would allow the gravitino to have decay modes \(\tilde{g} \to \tilde{g} G, \tilde{g} \to \tilde{g} W^{\pm}, \tilde{g} \to \tilde{g} \gamma, \tilde{g} \to \tilde{g} Z\). We note in passing that supergravity theories arising from string compactification with stabilized moduli often lead to a gravitino mass which may lie in the range 10 TeV and above (see, e.g., [89, 158, 159]).

For the gravitino mass lying in the 50–100 TeV region, the decay of the gravitino occurs significantly before the BBN time [157]. There is, however, one more constraint that one needs to attend to, which is that under the assumption of \(R\) parity conservation, each of the gravitino decay will result in an LSP neutralino which will contribute to the relic density of the Universe. Thus here the total relic density of the neutralinos will be given by \(\Omega_{\chi_{0}}^{\text{a}} = \Omega_{\chi_{0}^{\text{a}}}^{G} + \Omega_{\chi_{0}^{\text{a}}}^{W}\), where \(\Omega_{\chi_{0}^{G}}^{\text{a}}\) is the regular relic density arising from the thermal production of neutralinos after the freeze out and \(\Omega_{\chi_{0}^{W}}^{\text{a}}\) is the relic density arising from the decay of the gravitinos. We may write \(\Omega_{\chi_{0}^{a}}^{G}, \chi_{0}^{a}\) so that \(\Omega_{\chi_{0}^{a}}^{G} = (m_{\tilde{g}} / m_{G_{0}}) \Omega_{G_{0}}\), which implies that we need to compute the quantity \(\Omega_{G_{0}}\). We assume that the gravitinos produced in the early universe before the start of inflation have been inflated away and the relevant relic density of
Detailed analyses of this production exists in the literature and one finds that the relic density of neutralinos generated by the decay of the gravitinos produced by reheating is given by

\[ \Omega_\chi_0 h^2 = \sum_{i=1}^3 \omega_i m_i^2 / 3m_\tilde{g} \ln(k/f)(m_\tilde{g}/100 \text{ GeV})(T_R/10^{10} \text{ GeV}). \]

Here \( g_i, m_i (i = 1, 2, 3) \) are the gauge couplings and the gaugino masses for the gauge groups \( U(1)_Y, SU(2)_L \) and \( SU(3)_C \) which are evaluated at the reheat value \( T_R \) using renormalization group to evolve their values from the GUT scale. Further, \( \omega_i (i = 1, 2, 3) = (0.018, 0.044, 0.177) \) [160]. Analysis shows that negligible amount of the relic density arises from the decay of the gravitino up to reheating temperatures of \( 10^{10} \text{ GeV} \) [157].

11. Testing supergravity unification at future colliders

As mentioned in section 10, the case for SUSY/SUGRA is much stronger as a consequence of the Higgs boson discovery and the measurement of its mass at 125 GeV. Currently there is no other paradigm that can replace supergravity grand unification as we extrapolate physics from the electroweak scale to the GUT scale. For these reasons the search for sparticles must continue. There is a good chance that we will find sparticles at the LHC by the time all its runs are over. The discovery of even one sparticle will open up a new era for particle spectroscopy including the discovery of the remaining sparticles, and precision measurement of their masses and couplings. For these higher energy colliders are essential. For the future several proposals are under consideration both for high energy \( e^+ e^- \) colliders as well as for high energy proton–proton colliders. For the \( e^+ e^- \) colliders the possibilities are: (a) ILC: International Linear Collider, (b) CEPC: Circular Collider, (c) FCC-ee: Future Circular Collider. The ILC is under consideration in Japan, CEPC in China and FCC-ee at CERN. These colliders are essentially Higgs factories which are likely to run at an energy around 240 GeV which gives the optimal cross section for Higgs-strahlung, i.e., \( e^+ e^- \rightarrow Zh \). The \( Zh \) final state is the preferred mode rather than \( hh \) since \( Z \) can be efficiently detected via \( Z \rightarrow l^+ l^- \). The Higgs factories can do precision physics related to the Higgs boson specifically the couplings of the Higgs bosons to fermions and other electroweak parameters with great accuracy. Some supersymmetric effects could show up in these high precision experiments. The sparticles most likely to be probed would be electroweak gauginos and light sleptons if they are low mass. For the study of most other sparticles one needs proton–proton colliders. Here the possibilities are (i) SppC: 70–100 TeV pp collider, (ii) VLHC: 100 TeV pp collider. The SppC if built would be in China and VLHC is a possibility for the future at CERN.

12. Conclusion

There are a variety of reasons why supersymmetry is desirable when we think of high scale physics. One reason is the well known hierarchy problem [6] arising from loop corrections to the scalar masses. Thus while a loop correction to a fermion mass is proportional to the fermion mass, i.e., \( b m_f \propto m_f \), for the scalars the corrections to the scalar mass \( m \) is of the form \( b m^2 \propto \Lambda^2 \), where \( \Lambda \) is a cutoff scale. In a quantum field theory the cutoff scale could be order the Planck mass and thus the correction is very large. One of the ways to overcome this problem is supersymmetry where the loop correction from the squark exchange cancels the loop correction from the quark exchange. This leads to a natural cancellation of 1 part in \( 10^{28} \). The situation is similar to the \( \Delta S = 1 \) neutral current in the standard model where the charm quark loop cancels the u quark loop consistent with experiment \( Br(K^+ \rightarrow \mu^+ \nu) / Br(K^- \rightarrow \mu^- \bar{\nu}) = (6.84 \pm 0.11) \times 10^{-9} \) [161]. In this case the cancellation required is order one part in \( 10^9 \) and it leads to the discovery of the charm quark. In comparison for the case of the Higgs boson mass, the cancellation is 1 part in \( 10^{28} \) between the squark loop and the quark loop and thus there is an overwhelming reason for supersymmetry to be discovered.

In addition to the above let us recount some of the other successes of the supersymmetry/supergravity models. One of the early ones includes the fact that supersymmetry/supergravity models provide the right number of extra particles needed for the unification of couplings using the LEP data as one moves from the electroweak scale to the grand unification scale. Supergravity models provide an explanation for the tachyonic Higgs mass term that is central to accomplish spontaneous breaking of the electroweak symmetry in the standard model. Supergravity grand unification predicted the Higgs boson mass to lie below 130 GeV which is consistent with the current measurement of the Higgs boson. Currently there is no good alternative to supergravity grand unification if we wish to extrapolate physics from the electroweak scale to the grand unification scale. So the search for sparticles must continue. The most likely sparticle candidates for discovery are the neutralino \( \tilde{\chi}^0 \), the chargino \( \tilde{\chi}^\pm \), the gluino \( \tilde{g} \), the stop \( \tilde{t} \), and the stau \( \tilde{\tau} \). There is also the possibility of discovering the additional Higgs bosons that appear in SUSY extensions of the standard model. There is a good chance that with the full capability of the LHC (\( \mathcal{L} = 3000 \text{ fb}^{-1}, \sqrt{s} = 14 \text{ TeV} \)) one discovers sparticles and/or additional Higgs boson. At the same time a good idea to look ahead and plan for a 100 TeV pp super collider.

Acknowledgments

This research was supported in part by NSF Grant PHY-1620526.
References

Salam A 1969 Elementary Particle Theory ed N Svartholm (Stockholm: Almqvist and Wiksell) p 367
Hoof G ’71 1971 Nucl. Phys. B 33 173
Hoof G ’ and Veltman M J G 1972 Nucl. Phys. B 44 189
Volkov D V and Akulov V P 1972 JETP Lett. 16 438
Phys. Rev. D 13 3214
Ellis J, Kelley S and Nonopoulos D V 1990 Phys. Lett. 249B 441
Shafl Q and Wetterich C 1984 Phys. Rev. Lett. 52 875
Yanagida T 1979 KEK Report 79-18 Tsukuba p 95
Glashow S L 1980 Quarks and Leptons, Cargèse ed M Lévy et al (New York: Plenum) p 707
Wilczek F and Zee A 1982 Phys. Rev. D 25 553
Nath P and Syed R M 2004 Nucl. Phys. B 676 64
Enomoto S and Maekawa N 2011 Phys. Rev. D 84 096007
Nath P and Syed R M 2016 Phys. Rev. D 93 055005
Aulakh C S, Asawat R L and Krishna S 2017 Pramana 89 51
Georgi H and Jarlskog C 1979 Phys. Lett. 86B 297
Froggatt C D and Nielsen H B 1979 Nucl. Phys. B 147 277
Ibanez L and Ross G 1994 Phys. Lett. 332 100


Kors B and Nath P 2004 Nucl. Phys. B \textbf{681} 77


Raby S 2011 Rep. Prog. Phys. \textbf{74} 036901


Dirac P A M 1948 Phys. Rev. \textbf{74} 817

Schwinger J S 1966 Phys. Rev. \textbf{144} 1087


Abe K et al (Super-Kamiokande Collaboration) 2017 Phys. Rev. D \textbf{96} 012003

Abe K et al 2011 arXiv:1109.3262 [hep-ex]

Weinberg S 1982 Phys. Rev. D \textbf{26} 287


Lucas V and Raby S 1997 Phys. Rev. D \textbf{55} 6986


Ibrahim T and Nath P 2008 Rev. Mod. Phys. \textbf{80} 577


Nath P and Syed R M 2008 Phys. Rev. D \textbf{77} 015015

