114. Grand Unified Theories

Revised August 2017 by A. Hebecker (U. Heidelberg) and J. Hisano (Nagoya U.)

114.1. The standard model

The Standard Model (SM) may be defined as the renormalizable field theory with gauge group $G_{SM} = SU(3)_C \times SU(2)_L \times U(1)_Y$, with 3 generations of fermions in the representation

$$(\mathbf{3},\mathbf{2})_{1/3} + (\bar{\mathbf{3}},\mathbf{1})_{-4/3} + (\bar{\mathbf{3}},\mathbf{1})_{2/3} + (\mathbf{1},\mathbf{2})_{-1} + (\mathbf{1},\mathbf{1})_{2} \ , \qquad (114.1)$$

and a scalar Higgs doublet H transforming as $(\mathbf{1},\mathbf{2})_1$. Here and below we use boldface numbers to specify the dimension of representations of non-Abelian groups (in this case fundamental and antifundamental) and lower indices for $U(1)$ charges. The fields of Eq. (114.1) should also be familiar as $[Q, u^c, d^c, L, e^c]$, with $Q = (u, d)$ and $L = (\nu, e)$ being the quark and lepton $SU(2)$-doublets and u^c, d^c, e^c charge conjugate $SU(2)$-singlets.† Especially after the recent discovery of the Higgs, this model is remarkably complete and consistent with almost all experimental data.

A notable exception are neutrino masses, which are known to be non-zero but are absent in the SM even after the Higgs acquires its vacuum expectation value (VEV). The minimalist attitude is to allow for the dimension-five operator $(HL)^2$, which induces (Majorana) neutrino masses. In the seesaw mechanism [1,2,3] this operator is generated by integrating out heavy singlet fermions (right-handed (r.h.) neutrinos). Alternatively, neutrinos can have Dirac masses if light singlet neutrinos are added to the SM spectrum.

Conceptual problems of the SM include the absence of a Dark Matter candidate, of a mechanism for generating the baryon asymmetry of the universe, and of any reason for the observed smallness of the θ parameter of QCD (θ_{QCD}). In addition, the apparently rather complex group-theoretic data of Eq. (114.1) remains unexplained. Together with the abundance of seemingly arbitrary coupling constants, this disfavors the SM as a candidate fundamental theory, even before quantum gravity problems arise at energies near the Planck mass M_P.

To be precise, there are 19 SM parameters which have to be fitted to data: Three gauge couplings* g_3, g_2 and g_1, 13 parameters associated with the Yukawa couplings (9 charged fermion masses, three mixing angles and one CP phase in the CKM matrix.), the Higgs mass and quartic coupling, and θ_{QCD}. In addition, Majorana neutrinos introduce 3 more masses and 6 mixing angles and phases. As we will see, the paradigm of grand unification addresses mainly the group theoretic data of Eq. (114.1) and the values of the three gauge couplings. In many concrete realizations, it then impacts also the other mentioned issues of the SM, such as the family structure and fermion mass hierarchy.

More specifically, after precision measurements of the Weinberg angle θ_W in the LEP experiments, supersymmetric GUTs (SUSY GUTs) have become the leading candidates in the search for 'Physics beyond the SM'. Supersymmetry (SUSY) is a symmetry

† In our convention the electric charge is $Q = T_3 + Y/2$ and all our spinor fields are left-handed (l.h.).

* Equivalently, the $SU(2)_L$ and $U(1)_Y$ couplings are denoted as $g = g_2$ and $g' = \sqrt{3/5}\ g_1$. One also uses $\alpha_s = \alpha_3 = (g_3^2/4\pi)$, $\alpha_{\rm EM} = (e^2/4\pi)$ with $e = g\ \sin\theta_W$ and $\sin^2\theta_W = (g')^2/(g^2 + (g')^2)$.

between bosons and fermions which requires the addition of superpartners to the SM spectrum, thereby leading to the noted prediction of θ_W [4]. The measured Higgs mass ($\sim$ 125 GeV) is in principle consistent with this picture, assuming superpartners in the region of roughly 10 TeV. Such heavy superpartners then induce radiative corrections raising the Higgs mass above the Z boson mass m_Z [5,6]. However, if SUSY is motivated as a solution to the gauge hierarchy problem (i.e. to the naturalness problem of the Higgs mass) [7], its minimal incarnation in terms of the MSSM is becoming questionable. Indeed, compared to expectations based on the minimal SUSY SM (MSSM) with superpartner masses below about 1 TeV, the Higgs mass is somewhat too high [8]. Independently, the LHC has disfavored light colored superpartners. These facts represent new hints for future work on SUSY GUTs or on GUTs without TeV-scale supersymmetry.

114.2. Basic group theory and charge quantization

Historically, the first attempt at unification was the Pati-Salam model with gauge group $G_{PS} = SU(4)_C \times SU(2)_L \times SU(2)_R$ [9]. It unifies SM fermions in the sense that one generation (plus an extra SM singlet) now comes from the $(\mathbf{4}, \mathbf{2}, \mathbf{1}) + (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{2})$ of G_{PS}. This is easy to verify from the breaking pattern $SU(4)_C \to SU(3)_C \times U(1)_{B-L}$ together with the identification of SM hypercharge as a linear combination between $B - L$ (baryon minus lepton number) and the T_3 generator of $SU(2)_R$. This model explains charge quantization, that is, why all electric charges are integer multiples of some smallest charge in the SM. However, G_{PS} is not simple (containing three simple factors), and thus it does not predict gauge coupling unification.

Since G_{SM} has rank four (two for $SU(3)_C$ and one for $SU(2)_L$ and $U(1)_Y$, respectively), the rank-four group $SU(5)$ is the minimal choice for unification in a simple group [10]. The three SM gauge coupling constants derive from a universal coupling α_G at the GUT scale M_G. Explicitly embedding G_{SM} in $SU(5)$ is straightforward, with $SU(3)_C$ and $SU(2)_L$ corresponding e.g. to the upper-left 3×3 and lower-right 2×2 blocks, respectively, in traceless 5×5 matrices for $SU(5)$ generators of the fundamental representation. The $U(1)_Y$ corresponds to matrices generated by diag$(-2/3, -2/3, -2/3, 1, 1)$ and hence commutes with $SU(3)_C \times SU(2)_L \subset SU(5)$. It is then easy to derive how one SM generation precisely comes from the $\mathbf{10} + \overline{\mathbf{5}}$ of $SU(5)$ (where $\mathbf{10}$ is the antisymmetric rank-2 tensor):

$$\mathbf{10}: \begin{pmatrix} 0 & u_b^c & -u_g^c & u_r & d_r \\ -u_b^c & 0 & u_r^c & u_g & d_g \\ u_g^c & -u_r^c & 0 & u_b & d_b \\ -u_r & -u_g & -u_b & 0 & e^c \\ -d_r & -d_g & -d_b & -e^c & 0 \end{pmatrix} \quad \text{and} \quad \overline{\mathbf{5}}: \begin{pmatrix} d_r^c \\ d_g^c \\ d_b^c \\ e \\ -\nu_e \end{pmatrix}. \tag{114.2}$$

Since $SU(5)$ has 24 generators, $SU(5)$ GUTs have 12 new gauge bosons known as X bosons (or X/Y bosons) in addition to the SM. X bosons form an $SU(3)_C$-triplet and $SU(2)_L$-doublet. Their interaction connects quarks and leptons such that baryon and lepton numbers are not conserved and nucleon decay is predicted. Furthermore, $U(1)_Y$ hypercharge is automatically quantized since it is embedded in $SU(5)$.

In order to break the electroweak symmetry at the weak scale and give mass to quarks and leptons, Higgs doublets are needed. In the minimal $SU(5)$ model, they can sit in either a $\mathbf{5_H}$ or $\bar{\mathbf{5}}_{\mathbf{H}}$. The three additional states are referred to as color-triplet Higgs scalars. Their couplings also violate baryon and lepton numbers, inducing nucleon decay. In order not to violently disagree with the non-observation of nucleon decay, the triplet mass must be greater than $\sim 10^{11}$ GeV [11]. Moreover, in SUSY GUTs [12], in order to cancel anomalies as well as give mass to both up and down quarks, both Higgs multiplets $\mathbf{5_H}$ and $\bar{\mathbf{5}}_{\mathbf{H}}$ are required. As we shall discuss later, nucleon decay now constrains the Higgs triplets to have mass significantly greater than M_G in the minimal SUSY $SU(5)$ GUT since integrating out the Higgs triplets generates dimension-five baryon-number-violating operators [13]. The mass splitting between doublet and triplet in the $\mathbf{5_H}$ (and $\bar{\mathbf{5}}_{\mathbf{H}}$) comes from their interaction with the $SU(5)$ breaking sector.

While $SU(5)$ allows for the minimal GUT models, unification is not complete: Two independent representations, $\mathbf{10}$ and $\bar{\mathbf{5}}$, are required for one SM generation. A further representation, an SU(5) singlet, has to be added to serve as r.h. neutrino in the seesaw mechanism. In this case, the r.h. neutrino masses are not necessarily related to the GUT scale. By contrast, a single **16**-dimensional spinor representation of $SO(10)$ accommodates a full SM generation together with an extra singlet, potentially providing a r.h. neutrino [14]. This is most easily understood from the breaking pattern $SO(10) \to SU(5) \times U(1)_X$ and the associated branching rule* $\mathbf{16} = \mathbf{10}_{-5} + \bar{\mathbf{5}}_3 + \mathbf{1}_{-1}$. Here the indices refer to charges under the $U(1)_X$ subgroup, which is orthogonal to $SU(5)$ and reflects the fact that $SO(10)$ has rank five. From the above, it is easy to see that $U(1)_X$ charges can be given as $2Y - 5(B-L)$. Intriguingly, all representations of $SO(10)$ are anomaly free in four dimensions (4d). Thus, the absence of anomalies in an $SU(5)$-GUT or a SM generation can be viewed as deriving from this feature.

Table 114.1 presents the states of one family of quarks and leptons, as they appear in the **16**. To understand this, recall that the Γ-matrices of the 10d Clifford algebra give rise to five independent, anticommuting 'creation-annihilation' operators $\Gamma^{a\pm} = (\Gamma^{2a-1} \pm i\Gamma^{2a})/2$ with $a = 1, ..., 5$. These correspond to five fermionic harmonic oscillators or "spin" 1/2 systems. The 32-dimensional tensor product of those is reducible since the 10d rotation generators $M_{mn} = -i[\Gamma^m, \Gamma^n]/4$ $(m, n = 1, ..., 10)$ always flip an even number of "spins". This gives rise to the **16** as displayed in Table 114.1.

Next, one also recalls that the natural embedding of $SU(5)$ in $SO(10)$ relies on 'pairing up' real dimensions, $R^{10} \equiv C^5$, similarly to the paring up of Γ^ms used above. This makes it clear how to associate one $|\pm>$ system to each complex dimension of $SU(5)$, which explains the labeling of the "spin" columns in Table 114.1: The first three and last two "spins" correspond to $SU(3)_C$ and $SU(2)_L$, respectively. In fact, an $SU(3)_C$ rotation just raises one color index and lowers another, changing colors $\{r,\ g,\ b\}$, or changes relative phases between the three spin states. 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, or changes the relative phase between the two spin states. In this representation $U(1)_Y$

* Useful references on group theory in the present context include [15] and refs. therein.

hypercharge is simply given by $Y = -2/3(\sum$ color spins$) + (\sum$ weak spins$)$. $SU(5)$ rotations corresponding to X bosons 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. Since $SO(10)$ has 45 generators, additional 21 gauge bosons are introduced including the $U(1)_X$ above. The 20 new $SO(10)$ rotations not in $SU(5)$ are then given by either raising any two spins or lowering them. With these rotations, **1** and **5** are connected with **10**. The last $SO(10)$ rotation changes phases of states with weight $2(\sum$ color spins$) + 2(\sum$ weak spins$)$, which corresponds to $U(1)_X$.

Table 114.1: Quantum numbers of **16**-dimensional representation of $SO(10)$.

state	Y	Color	Weak	$SU(5)$	$SO(10)$
ν^c	0	$---$	$--$	**1**	
e^c	2	$---$	$++$		
u_r	1/3	$+--$	$-+$		
d_r	1/3	$+--$	$+-$		
u_g	1/3	$-+-$	$-+$		
d_g	1/3	$-+-$	$+-$	**10**	
u_b	1/3	$--+$	$-+$		
d_b	1/3	$--+$	$+-$		**16**
u^c_r	$-4/3$	$-++$	$--$		
u^c_g	$-4/3$	$+-+$	$--$		
u^c_b	$-4/3$	$++-$	$--$		
d^c_r	2/3	$-++$	$++$		
d^c_g	2/3	$+-+$	$++$		
d^c_b	2/3	$++-$	$++$	$\bar{\mathbf{5}}$	
ν	-1	$+++$	$-+$		
e	-1	$+++$	$+-$		

$SO(10)$ has two inequivalent maximal subgroups and hence breaking patterns, $SO(10) \to SU(5) \times U(1)_X$ and $SO(10) \to SU(4)_C \times SU(2)_L \times SU(2)_R$. In the first case, one can carry on breaking to $G_{SM} \subset SU(5)$ precisely as in the minimal $SU(5)$ case above. Alternatively, one can identify $U(1)_Y$ as an appropriate linear combination of $U(1)_X$ and the $U(1)$ factor from $SU(5)$, leading to the so-called flipped $SU(5)$ [16] as an intermediate step in breaking $SO(10)$ to G_{SM}. In the second case, we have an intermediate Pati-Salam model thanks to the branching rule $\mathbf{16} = (\mathbf{4}, \mathbf{2}, \mathbf{1}) + (\bar{\mathbf{4}}, \mathbf{1}, \mathbf{2})$. Finally, $SO(10)$ can break directly to the SM at M_G. Gauge coupling unification remains intact in the case of this 'direct' breaking and for the breaking pattern $SO(10) \to SU(5) \to G_{SM}$ (with $SU(5)$ broken at M_G). In the case of intermediate-scale Pati-Salam or flipped $SU(5)$ models, gauge coupling predictions are modified. The Higgs multiplets in the minimal $SO(10)$ come from the fundamental representation, $\mathbf{10_H} = \mathbf{5_H} + \bar{\mathbf{5}}_\mathbf{H}$. Note, only in $SO(10)$ does the representation type distinguish SM matter from Higgs fields.

Finally, larger symmetry groups can be considered. For example, the exceptional group E_6 has maximal subgroup $SO(10) \times U(1)$ [17]. Its fundamental representation branches as $\mathbf{27} = \mathbf{16}_1 + \mathbf{10}_{-2} + \mathbf{1}_4$. Another maximal subgroup is $SU(3)_C \times SU(3)_L \times SU(3)_R \subset \mathrm{E}_6$ with branching rule $\mathbf{27} = (\mathbf{3}, \mathbf{3}, \mathbf{1}) + (\bar{\mathbf{3}}, \mathbf{1}, \bar{\mathbf{3}}) + (\mathbf{1}, \bar{\mathbf{3}}, \mathbf{3})$. Independently of any underlying E_6, the group $[SU(3)]^3$ with additional permutation symmetry Z_3 interchanging the three factors can be considered. This is known as "trinification" [18]. The $\mathrm{E}_6 \to [SU(3)]^3$ breaking pattern has been used in phenomenological analyses of the heterotic string [19]. However, 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.

Intriguingly, the logic by which G_{SM} is a maximal subgroup of $SU(5)$, which together with $U(1)_X$ is a maximal subgroup of $SO(10)$, continues in a very elegant and systematic way up to the largest exceptional group. The resulting famous breaking chain $\mathrm{E}_8 \to \mathrm{E}_7 \to \mathrm{E}_6 \to SO(10) \to SU(5) \to G_{SM}$ together with the special role played by E_8 in group and in string theory is a tantalizing hint at deeper structures. However, since all representations of E_8 and E_7 are real and can not lead to 4d chiral fermions, this is necessarily outside the 4d GUT framework.

114.3. GUT breaking and doublet-triplet splitting

In the standard, 4d field-theoretic approach to GUTs, the unified gauge group is broken spontaneously by an appropriate GUT Higgs sector. Scalar potentials (or superpotentials in SUSY GUTs) exist whose vacua spontaneously break $SU(5)$ or $SO(10)$. While these potentials are ad hoc (just like the Higgs potential in the SM), the most naive expectation is that all their dimensionful parameters are $O(M_G)$. In the simplest case of $SU(5)$, the **24** (adjoint) GUT Higgs develops a VEV along the G_{SM}-singlet direction as $\langle\Phi\rangle \propto \mathrm{diag}(-2/3, -2/3, -2/3, 1, 1)$. In order for $SO(10)$ to break to $SU(5)$, the **16** or **126**, which have a G_{SM}-singlet with non-zero $U(1)_X$ charge, get a VEV.

The masses of doublet and triplet in the $\mathbf{5_H}$ (and $\bar{\mathbf{5}}_\mathbf{H}$) generically split due to their coupling to the GUT Higgs. In addition, both the doublet and the triplet masses also get an equal contribution from an $SU(5)$-invariant GUT-scale mass term. Without any further structure, an extreme fine-tuning between two large effects is then necessary to keep the doublet mass at the electroweak scale. Supersymmetry plays an important role in forbidding large radiative correction to the doublet mass due to the non-renormalization theorem [7]. However, even in this case we have to fine tune parameters at tree level. This is the doublet-triplet splitting problem which, in the SUSY context, is clearly related the μ-term problem of the MSSM (the smallness of the coefficient of $\mu H_u\, H_d$).

Several mechanisms for natural doublet-triplet splitting have been suggested under the assumption of supersymmetry, such as the sliding singlet [20], missing partner [21] missing VEV [22], and pseudo-Nambu-Goldstone boson mechanisms [23]. Particular examples of the missing partner mechanism for $SU(5)$ [24], the missing VEV mechanism for $SO(10)$ [25,26] and the pseudo-Nambu-Goldstone boson mechanism for $SU(6)$ [27] have been shown to be consistent with gauge coupling unification and nucleon decay. From the GUT-scale perspective, one is satisfied if the triplets are naturally heavy and the doublets are massless ($\mu \simeq 0$). There are also several mechanisms for resolving the subsequent issue of why μ is of order the SUSY breaking scale [28]. * For a review of the μ problem and some suggested solutions in SUSY GUTs and string theory, see [29,30,31,32] and references therein.

In general, GUT-breaking sectors successfully resolving the doublet-triplet splitting problem, dynamically stabilizing all GUT-scale VEVs and allowing for realistic neutrino masses and Yukawa couplings (including the GUT-symmetry violation in the latter) require a number of ingredients. However, for validity of the effective theory, introduction of higher or many representations is limited, otherwise a Landau pole may appear below the Planck scale. In addition, GUTs are only effective theories below the Planck scale in the 4d field-theoretic approach. Since M_G is close to this scale, the effects of higher-dimension operators are not obviously negligible. In particular, operators including the GUT-breaking Higgs may affect low-energy predictions, such as quark and lepton masses.

Thus, especially in the context of GUT breaking and doublet-triplet splitting, models beyond 4d field theory appear attractive. While this is mainly the subject of the next section, some advantages can already be noted: In models with extra dimensions, in particular string constructions, GUT breaking may occur due to boundary conditions in the compactified dimensions [33,34,35,36]. No complicated GUT breaking sector is then required. Moreover, boundary conditions can give mass only to the triplet, leaving the doublet massless. This is similar to the 'missing partner mechanism' since the effective mass term does not 'pair up' the triplets from $\mathbf{5_H}$ and $\overline{\mathbf{5}}_{\mathbf{H}}$ but rather each of them with further fields which are automatically present in the higher-dimensional theory. This can eliminate dimension-five nucleon decay (cf. Sec. 114.6).

114.4. String-theoretic and higher-dimensional unified models

As noted earlier, the GUT scale is dangerously close to the scale of quantum gravity. It may hence be necessary to discuss unified models of particle physics in the latter, more ambitious context. Among the models of quantum gravity, superstring or M-theory stands out as the best-studied and technically most developed proposal, possessing in particular a high level internal, mathematical consistency. For our purposes, it is sufficient

* The solution of [28] relies on the absence of the fundamental superpotential term $\mu H_u\, H_d$ (or $\mu\, \mathbf{5}_H \overline{\mathbf{5}}_H$). This is ensured by a $U(1)_R$. The latter clashes with typical superpotentials for the GUT breaking sector. However, higher-dimensional or stringy GUTs, where the triplet Higgs is simply projected out, can be consistent with the $U(1)_R$ symmetry.

to know that five 10d and one 11d low-energy effective supergravity theories arise in this setting (cf. [37] and refs. therein).

Grand unification is realized most naturally in the context of the two 'heterotic' theories with gauge groups $E_8 \times E_8$ and $SO(32)$, respectively [35] (see [38] for some of the more recent results). Justified in part by the intriguing breaking path $E_8 \to \cdots \to G_{SM}$ mentioned above, the focus has historically largely been on $E_8 \times E_8$. To describe particle physics, solutions of the 10d theory with geometry $R^{1,3} \times M_6$ are considered, where M_6 is a Calabi-Yau (CY) 3-fold (with 6 real dimensions). The background solution involves expectation values of higher-dimensional components of the $E_8 \times E_8$ gauge fields. This includes both Wilson lines [33] and non-vanishing field-strength and leads, in general, to a reduced gauge symmetry and to chirality in the resulting 4d effective theory. The 4d fermions arise from 10d gauginos.

Given an appropriate embedding of G_{SM} in $E_8 \times E_8$, gauge coupling unification is automatic at leading order. Corrections arise mainly through (string)-loop effects and are similar to the familiar field-theory thresholds of 4d GUTs [39]. Thus, one may say that coupling unification is a generic prediction in spite of the complete absence* of a 4d GUT at any energy scale. This absence is both an advantage and a weakness. On the up side, GUT breaking and doublet-triplet splitting [41] are more naturally realized and dimension-five nucleon decay is relatively easy to avoid. On the down side, there is no reason to expect full GUT representations in the matter sector and flavor model building is much less tied to the GUT structure than in 4d.

One technical problem of heterotic constructions is the dependence on the numerous size and shape parameters of M_6 (the so-called moduli), the stabilization of which is poorly understood (see [42] for recent developments). Another is the sheer mathematical complexity of the analysis, involving in particular the study of (non-Abelian) gauge-bundles on CY spaces [43] (see however [44]) .

An interesting sub-chapter of heterotic string constructions is represented by orbifold models [34]. Here the internal space is given by a six-torus, modded out by a discrete symmetry group (e.g. T^6/Z_n). More recent progress is reported in [45,46], including in particular the systematic exploration of the phenomenological advantages of so-called 'non-prime' (referring to n) orbifolds. The symmetry breaking to G_{SM} as well as the survival of Higgs doublets without triplet partners is ensured by the appropriate embedding of the discrete orbifold group in $E_8 \times E_8$. String theory on such spaces, which are locally flat but include singularities, is much more calculable than in the CY case. The orbifold geometries can be viewed as singular limits of CYs.

An even simpler approach to unified models, which includes many of the advantages of full-fledged string constructions, is provided by Orbifold GUTs [36]. These are (mostly) 5d or 6d SUSY field theories with unified gauge group (e.g. $SU(5)$ or $SO(10)$), broken in the process of compactifying to 4d. To give a particularly simple example, consider $SU(5)$ on $R^{1,3} \times S^1/(Z_2 \times Z_2')$. Here the compact space is an interval of length $\pi R/2$ and the embedding of Z_2' in the hypercharge direction of $SU(5)$ realizes the breaking to G_{SM}. Concretely, 5d X bosons are given Dirichlet BCs at one endpoint of the interval and thus

* See however [40].

have no Kaluza-Klein (KK) zero mode. Their lightest modes have mass $\sim 1/R$, making the KK-scale the effective GUT scale. As an implication, the boundary theory has no $SU(5)$ invariance. Nevertheless, since the $SU(5)$-symmetric 5d bulk dominates 4d gauge couplings, unification remains a prediction. Many other features but also problems of 4d GUTs can be circumvented, especially doublet-triplet splitting is easily realized.

With the advent of the string-theory 'flux landscape' [47], which is best understood in 10d type-IIB supergravity, the focus in string model building has shifted to this framework. While type II string theories have no gauge group in 10d, brane-stacks support gauge dynamics. A particularly appealing setting (see e.g. [48]) is provided by type IIB models with D7 branes (defining 8d submanifolds). However, in the $SO(10)$ context the **16** is not available and, for $SU(5)$, the top-Yukawa coupling vanishes at leading order [49]. As a crucial insight, this can be overcome on the non-perturbative branch of type IIB, also known as F-theory [50,51]. This setting allows for more general branes, thus avoiding constraints of the Dp-brane framework. GUT breaking can be realized using hypercharge flux (the VEV of the $\mathrm{U}(1)_Y$ field strength), an option not available in heterotic models. The whole framework combines the advantages of the heterotic or higher-dimensional unification approach with the more recent progress in understanding moduli stabilization. It thus represents at this moment the most active and promising branch of theory-driven GUT model building (see e.g. [52] and refs. therein).

As a result of the flux-breaking, a characteristic 'type IIB' or 'F-theoretic' tree-level correction to gauge unification arises [53]. The fact that this correction can be rather significant numerically is occasionally held against the framework of F-theory GUTs. However, at a parametric level, this correction nevertheless behaves like a 4d threshold, i.e., it provides $\mathcal{O}(1)$ additive contributions to the inverse 4d gauge coupling $\alpha_i^{-1}(M_G)$.

A final important issue in string GUTs is the so-called string-scale/GUT-scale problem [54]. It arises since, in heterotic compactifications, the Planck scale and the high-scale value of the gauge coupling unambiguously fix the string-scale to about 10^{18} GeV. As the compactification radius R is raised above the string length, the GUT scale (identified with $1/R$) goes down and the string coupling goes up. Within the domain of perturbative string theory, a gap of about a factor ~ 20 remains between the lowest GUT scale achievable in this way and the phenomenological goal of 2×10^{16} GeV. The situation can be improved by venturing into the non-perturbative regime [54] or by considering 'anisotropic' geometries with hierarchically different radii R [54,55].

In F-theory GUTs, the situation is dramatically improved since the gauge theory lives only in four out of the six compact dimensions. This allows for models with a 'decoupling limit', where the GUT scale is parametrically below the Planck scale [51]. However, moduli stabilization may not be without problems in such constructions, in part due to a tension between the required large volume and the desirable low SUSY breaking scale.

114.5. Gauge coupling unification

The quantitative unification of the three SM gauge couplings at the energy scale M_G is one of the cornerstones of the GUT paradigm. It is obviously of direct phenomenological relevance. Gauge coupling unification is best understood in the framework of effective field theory (EFT) [56]. In the simplest case, the relevant EFT at energies $\mu \gg M_G$ has a unified gauge symmetry (say $SU(5)$ for definiteness) and a single running gauge coupling $\alpha_G(\mu)$. At energies $\mu \ll M_G$, states with mass $\sim M_G$ (such as X bosons, GUT Higgs, color-triplet Higgs) have to be integrated out. The EFT now has three independent couplings and SM (or SUSY SM) matter content. One-loop renormalization group equations readily allow for an extrapolation to the weak scale,

$$\alpha_i^{-1}(m_Z) = \alpha_G^{-1}(M_G) + \frac{b_i}{2\pi}\log\left(\frac{M_G}{m_Z}\right) + \delta_i\,, \tag{114.3}$$

(i = 1-3). Here we defined δ_i to absorb all sub-leading effects, including threshold corrections at or near the weak scale (e.g. from superpartners and the additional Higgs bosons in the case of the MSSM). We will discuss them momentarily.

It is apparent from Eq. (114.3) that the three low-scale couplings can be very different. This is due to the large energy range $m_Z \ll \mu \ll M_G$ and the non-universal β-function coefficients ($b_i^{\rm SM} = \{41/10, -19/6, -7\}$ or $b_i^{\rm MSSM} = \{33/5, 1, -3\}$). Incomplete GUT multiplets, such as gauge and Higgs bosons in the SM and also their superpartners and the additional Higgs bosons in the MSSM, contribute to the differences between the β functions. Inverting the argument, one expects that extrapolating the measured couplings to the high scale, we find quantitative unification at $\mu \sim M_G$. While this fails in the SM, it works intriguingly well in the MSSM (cf. Fig. 1).

The three equations contained in (Eq. (114.3)) can be used to determine the three 'unknowns' $\alpha_3(m_Z)$, $\alpha_G(M_G)$ and M_G, assuming that all other parameters entering the equations are given. Focusing on the SUSY case and using the $\overline{\rm MS}$ coupling constants $\alpha_{\rm EM}^{-1}(m_Z)$ and $\sin^2\theta_W(m_Z)$ from [57],

$$\alpha_{\rm EM}^{-1}(m_Z) = 127.950 \pm 0.017\,, \tag{114.4}$$

$$\sin^2\theta_W(m_Z) = 0.23129 \pm 0.00005\,, \tag{114.5}$$

as input, one determines $\alpha_{1,2}^{-1}(m_Z)$, which then gives

$$\alpha_G^{-1}(M_G) \simeq 24.3 \qquad \text{and} \qquad M_G \simeq 2\times 10^{16}\ \text{GeV}\,. \tag{114.6}$$

Here we have set $\delta_i = 0$ for simplicity. Crucially, one in addition obtains a prediction for the low-energy observable α_3,

$$\alpha_3^{-1}(m_Z) = -\frac{5}{7}\alpha_1^{-1}(m_Z) + \frac{12}{7}\alpha_2^{-1}(m_Z) + \Delta_3\,, \tag{114.7}$$

where

$$\Delta_3 = \frac{5}{7}\delta_1 - \frac{12}{7}\delta_2 + \delta_3\,. \tag{114.8}$$

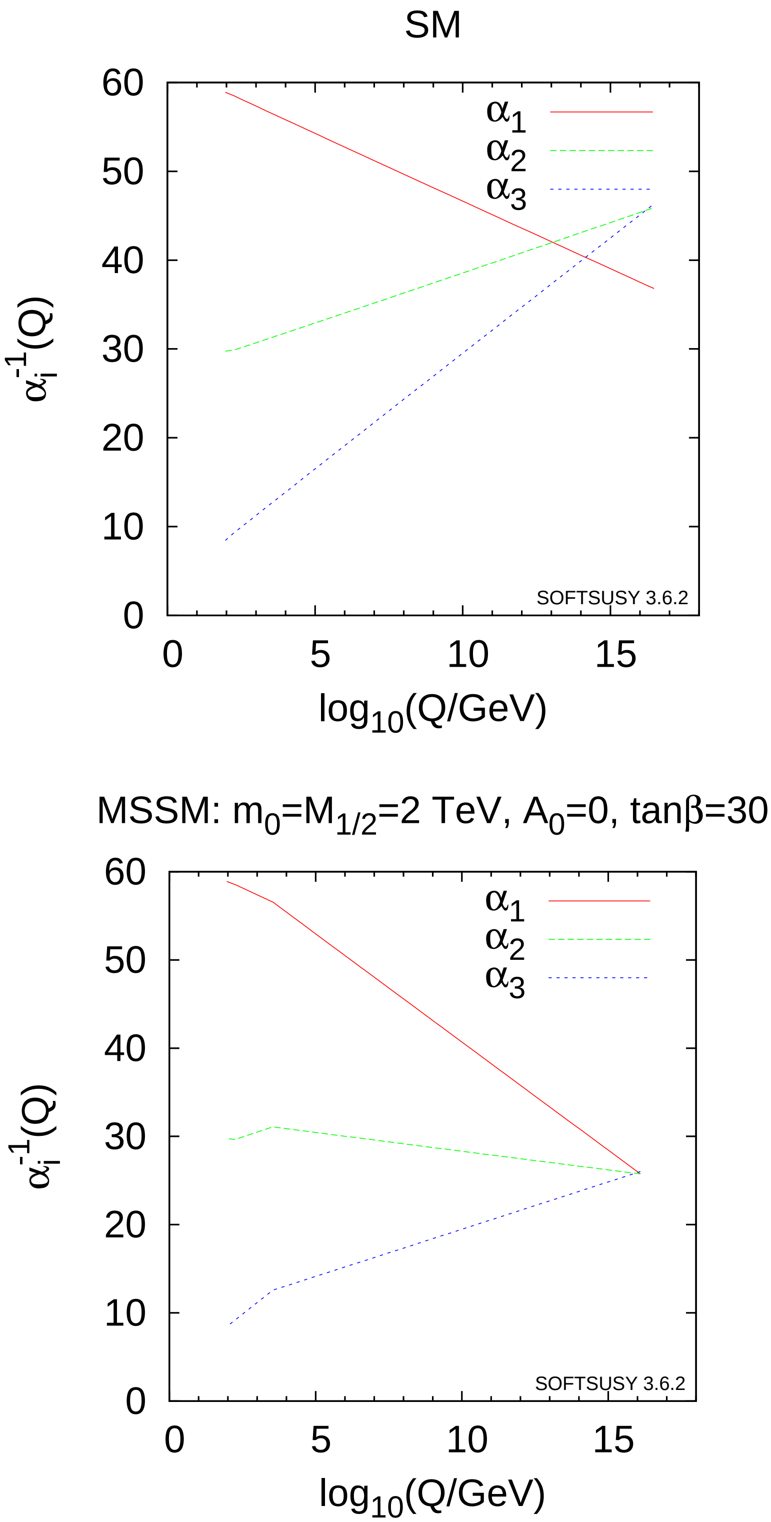

Figure 114.1: Running couplings in SM and MSSM using two-loop RG evolution. The SUSY threshold at 2 TeV is clearly visible on the MSSM side. (We thank Ben Allanach for providing the plots created using SOFTSUSY [62].)

Here we followed the elegant formulation in Ref. [58] of the classical analyses of [4]. Of course, it is a matter of convention which of the three low-energy gauge coupling parameters one 'predicts' and indeed, early works on the subject discussed the prediction of $\sin^2\theta_W$ in terms of $\alpha_{\rm EM}$ and α_3 [59,60].

Remarkably, the leading order result (i.e. Eq. (114.7) with $\delta_i = 0$) is in excellent agreement with experiments [57]:

$$\alpha_3^{\rm LO}(m_Z) = 0.117 \qquad \text{vs.} \qquad \alpha_3^{\rm EXP}(m_Z) = 0.1181 \pm 0.0011\,. \tag{114.9}$$

However, this near perfection is to some extent accidental. To see this, we now discuss the various contributions to the δ_i (and hence to Δ_3).

The two-loop running correction from the gauge sector $\Delta_3^{(2)}$ and the low-scale threshold correction $\Delta_3^{(l)}$ from superpartners can be summarized as [58]

$$\Delta_3^{(2)} \simeq -0.82 \qquad \text{and} \qquad \Delta_3^{(l)} \simeq \frac{19}{28\pi}\log\left(\frac{m_{\rm SUSY}}{m_Z}\right)\,. \tag{114.10}$$

The relevant scale $m_{\rm SUSY}$ can be estimated as [61]

$$m_{\rm SUSY} \quad \rightarrow \quad m_H^{3/19}\, m_{\widetilde{H}}^{12/19}\, m_{\widetilde{W}}^{4/19} \times \left(\frac{m_{\widetilde{W}}}{m_{\widetilde{g}}}\right)^{28/19} \left(\frac{m_{\widetilde{l}}}{m_{\widetilde{q}}}\right)^{3/19} \quad , \tag{114.11}$$

where m_H stands for the masses of non-SM Higgs states and superpartner masses are given in self-evident notation. Detailed analyses including the above effects are best done using appropriate software packages, such as SOFTSUSY [62], (or alternatively SuSpect [63] or SPheno [64]) . See also [62] for references to the underlying theoretical two-loop analyses.

To get a very rough feeling for these effects, let us assume that all superpartners are degenerate at $m_{\rm SUSY} = 1$ TeV, except for heavier gluinos: $m_{\widetilde{W}}/m_{\widetilde{g}} \simeq 1/3$. This gives $\Delta_3^{(l)} \simeq -0.35 + 0.22\ln(m_{\rm SUSY}/m_Z) \simeq 0.18$. The resulting prediction of $\alpha_3(m_Z) \simeq 0.126$ significantly upsets the perfect one-loop agreement found earlier. Before discussing this issue further, it is useful to introduce yet another important type of correction, the high or GUT scale thresholds.

To discuss high scale thresholds, let us set all other corrections to zero for the moment and write down a version of Eq. (114.3) that captures the running near and above the GUT scale more correctly. The threshold correction at one-loop level can be evaluated accurately by the simple step-function approximation for the β functions in the $\overline{\rm DR}$ scheme* [68],

$$\alpha_i^{-1}(m_Z) = \alpha_G^{-1}(\mu) + \frac{1}{2\pi}\left[b_i \ln\frac{\mu}{m_Z} + b_i^C \ln\frac{\mu}{M_C} + b_i^X \ln\frac{\mu}{M_X} + b_i^\Phi \ln\frac{\mu}{M_\Phi}\right]\,. \tag{114.12}$$

* The $\overline{\rm DR}$ scheme is frequently used in a supersymmetric regularization [65]. The renormalization transformation of the gauge coupling constants from $\overline{\rm MS}$ to $\overline{\rm DR}$ scheme is given in Ref. [66]. For an alternative treatment using holomorphic gauge couplings and NSVZ β-functions see e.g. [67].

Here we started the running at some scale $\mu \gg M_G$, including the contribution of the minimal set of states relevant for the transition from the high-scale $SU(5)$ model to the MSSM. These are the color-triplet Higgs multiplets with mass M_C, massive vector multiplets of X-bosons with mass M_X (including GUT Higgs degrees of freedom), and the remaining GUT-Higgs fields and superpartners with mass M_Φ. The coefficients $b_i^{C,X,\Phi}$ can be found in Ref. [69]. Crucially, the b_i in Eq. (114.12) conspire to make the running GUT-universal at high scales, such that the resulting prediction for α_3 does not depend on the value of μ.

To relate this to our previous discussion, we can, for example, define $M_G \equiv M_X$ and then choose $\mu = M_G$ in Eq. (114.12). This gives the high-scale threshold corrections

$$\delta_i^{(h)} = \frac{1}{2\pi}\left[b_i^C \ln\frac{M_G}{M_C} + b_i^\Phi \ln\frac{M_G}{M_\Phi}\right], \tag{114.13}$$

and a corresponding correction $\Delta_3^{(h)}$. To get some intuition for the magnitude, one can furthermore assume $M_\Phi = M_G$, finding (with $b_i^C = \{2/5, 0, 1\}$)

$$\Delta_3^{(h)} = \frac{9}{14\pi}\ln\left(\frac{M_G}{M_C}\right). \tag{114.14}$$

To obtain the desired effect of $-\Delta_3^{(2)} - \Delta_3^{(l)} \simeq +0.64$, the triplet Higgs would have to be by about a factor 20 lighter than the GUT scale. While this is ruled out by nucleon decay in the minimal model [70] as will be discussed Sec. 114.6, it is also clear that threshold corrections of this order of magnitude can, in general, be realized with a certain amount of GUT-scale model building, e.g. in specific $SU(5)$ [24] or $SO(10)$ [25,26] constructions. It is, however, a significant constraint on the 4d GUT sector of the theory.

The above analysis implicitly assumes universal soft SUSY breaking masses at the GUT scale, which directly affect the spectrum of SUSY particles at the weak scale. 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 [71] and refs. therein). For example, if gaugino masses were not unified at M_G and, in particular, gluinos were lighter than winos at the weak scale (cf. Eq. (114.11))), then it is possible that, due to weak scale threshold corrections, a much smaller or even slightly negative threshold correction at the GUT scale would be consistent with gauge coupling unification [72].

It is also noteworthy that perfect unification can be realized without significant GUT-scale corrections, simply by slightly raising the (universal) SUSY breaking scale. In this case the dark matter abundance produced by thermal processes in the early universe (if the lightest neutralino is the dark matter particle) is too high. However, even if the gaugino mass in the MSSM is about 1 TeV to explain the dark matter abundance, if the Higgsino and the non-SM Higgs boson masses are about 10-100 TeV, the effective SUSY scale can be raised [73]. This setup is realized in split SUSY [74] or the pure gravity mediation model [75] based on anomaly mediation [76]. Since the squarks and sleptons are much heavier than the gaugino masses in those setups, a gauge hierarchy problem is reintroduced. The facts that no superpartners have so far been seen at the LHC and that the observed Higgs mass favors heavier stop masses than about 1 TeV force one to accept a certain amount of fine-tuning anyway.

For non-SUSY GUTs or GUTs with a very high SUSY breaking scale to fit the data, new light states in incomplete GUT multiplets or multiple GUT breaking scales are required. For example, non-SUSY models $SO(10) \to SU(4)_C \times SU(2)_L \times SU(2)_R \to \mathrm{SM}$, 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 [77] and at the same time survive nucleon decay bounds [78]. Alternatively, one can appeal to string-theoretic corrections discussed in Sec. 114.4 to compensate for a high SUSY breaking scale. This has, for example, been concretely analyzed in the context of F-theory GUTs in [79].

In 5d or 6d orbifold GUTs, certain "GUT scale" threshold corrections come from the Kaluza-Klein modes between the compactification scale, $M_c \sim 1/R$, and the effective cutoff scale M_*. In string theory, this cutoff scale is the string scale. Gauge coupling unification at two loops then constrains the values of M_c and M_*.* Typically, one finds M_c to be lower than the 4d GUT scale. Since the X-bosons, responsible for nucleon decay, get mass at the compactification scale, this has significant consequences for nucleon decay.

Finally, it has been shown that non-supersymmetric GUTs in warped 5d orbifolds can be consistent with gauge coupling unification. This assumes (in 4d language) that the r.h. top quark and the Higgs doublets are composite-like objects with a compositeness scale in the TeV range [81].

114.6. Nucleon decay

Quarks and leptons are indistinguishable in any 4d GUT, and both the baryon (B) and lepton number (L) are not conserved. This leads to baryon-number-violating nucleon decay. In addition to baryon-number violation, lepton-number violation is also required for nucleon decay since, in the SM, leptons are the only free fermions which are lighter than nucleons. The lowest-dimension operators relevant for nucleon decay are $(B+L)$ violating dimension-six four-fermion-terms in the SM, and all baryon-violating operators with dimension less than seven preserve $(B-L)$ [82]. In $SU(5)$ GUTs, the dimension-six operators are induced by X boson exchange. These operators are suppressed by $(1/M_G^2)$, and the nucleon lifetime is given by $\tau_N \propto M_G^4/(\alpha_G^2\ m_p^5)$ (m_p is proton mass). The dominant decay mode of the proton (and the baryon-violating decay mode of the neutron), via X boson exchange, is $p \to e^+\ \pi^0$ ($n \to e^+\ \pi^-$). In any simple gauge symmetry, with one universal GUT coupling α_G and scale M_G, the nucleon lifetime from gauge

* It is interesting to note that a ratio $M_*/M_c \sim 100$, needed for gauge coupling unification to work in orbifold GUTs, is typically the maximum value for this ratio consistent with perturbativity [80].

boson 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, IMB, and Kamiokande detectors [59]. The present experimental bounds come from Super-Kamiokande. The null result on search for $p \to e^+\pi^0$ constrains M_G to be be larger than $O(10^{15})$ GeV. Non-SUSY GUTs are constrained by the non-observation of nucleon decay, while a precise and general statement is hard to make. The reason is that gauge couplings do not unify with just the SM particle content. Once extra states or large thresholds are included to ensure precision unification, a certain range of unification scales is allowed. By contrast, in SUSY GUTs one generically has $M_G \sim 2 \times 10^{16}$ GeV from the gauge coupling unification. Hence dimension-six baryon-number-violating operators are predicted to induce a lifetime of about $\tau_p \sim 10^{36}$ years.

However, in SUSY GUTs there are additional sources for baryon and/or lepton-number violation – dimension-four and five operators [13]. These arise since, in the SUSY SM, quarks and leptons have scalar partners (squarks and sleptons). Although our notation does not change, when discussing SUSY models our fields are chiral superfields and both fermionic and bosonic matter is implicitly represented by those. In this language, baryon- and/or lepton-number-violating dimension-four and five operators are given as so-called F terms of products of chiral superfields, which contain two fermionic components and the rest scalars or products of scalars. Within the context of $SU(5)$ the dimension-four and five operators have the form

$$(\mathbf{10}\ \bar{\mathbf{5}}\ \bar{\mathbf{5}}) \supset (u^c\ d^c\ d^c) + (Q\ L\ d^c) + (e^c\ L\ L),$$

$$(\mathbf{10}\ \mathbf{10}\ \mathbf{10}\ \bar{\mathbf{5}}) \supset (Q\ Q\ Q\ L) + (u^c\ u^c\ d^c\ e^c)\ + B\text{- and } L\text{-conserving terms},$$

respectively. The dimension-four operators are renormalizable, with dimensionless couplings similar to Yukawa couplings. By contrast, the dimension-five operators have a dimensionful coupling of order $(1/M_G)$. They are generated by integrating out the color-triplet Higgs with GUT-scale mass. Note that both triplet Higgsinos (due to their fermionic nature) and Higgs scalars (due to their mass-enhanced trilinear coupling with matter) contribute to the operators.

The dimension-four operators violate either baryon number or lepton number. The nucleon lifetime is extremely short if both types of dimension-four operators are present in the SUSY SM since squark or slepton exchange induces the dangerous dimension-six SM operators. Even in the case that they violate baryon number or lepton number only but not both, they are constrained by various phenomena [83]. For example, the primordial baryon number in the universe is washed out unless the dimensionless coupling constants are less than 10^{-7}. Both types of operators can be eliminated by requiring R parity, which distinguishes Higgs from ordinary matter multiplets. R parity [84] or its cousin, matter parity [12,85], act as $F \to -F$, $H \to H$ with $F = \{\mathbf{10},\ \bar{\mathbf{5}}\}$, $H = \{\bar{\mathbf{5}}_{\mathbf{H}},\ \mathbf{5}_{\mathbf{H}}\}$ in $SU(5)$. This forbids the dimension-four operator $(\mathbf{10}\ \bar{\mathbf{5}}\ \bar{\mathbf{5}})$, but allows the Yukawa couplings for quark and lepton masses of the form $(\mathbf{10}\ \bar{\mathbf{5}}\ \bar{\mathbf{5}}_{\mathbf{H}})$ and $(\mathbf{10}\ \mathbf{10}\ \mathbf{5}_{\mathbf{H}})$. It also forbids the dimension-three, lepton-number-violating operator $(\bar{\mathbf{5}}\ \mathbf{5}_{\mathbf{H}}) \supset (L\ H_u)$ as well as the dimension-five, baryon-number-violating operator $(\mathbf{10}\ \mathbf{10}\ \mathbf{10}\ \bar{\mathbf{5}}_{\mathbf{H}}) \supset (Q\ Q\ Q\ H_d) + \cdots$. In

$SU(5)$, the Higgs multiplet $\bar{\mathbf{5}}_{\mathbf{H}}$ and the matter multiplets $\bar{\mathbf{5}}$ have identical gauge quantum numbers. In E_6, Higgs and matter multiplets could be unified within the fundamental **27** representation. Only in $SO(10)$ are Higgs and matter multiplets distinguished by their gauge quantum numbers. The Z_4 center of $SO(10)$ distinguishes **10**s from **16**s and can be associated with R parity [86].

The dimension-five baryon-number-violating operators may also be forbidden at tree level by certain symmetries consistent with $SU(5)$ [13]. However, these symmetries are typically broken by the VEVs responsible for the color-triplet Higgs masses. Consequently the dimension-five operators are generically generated via the triplet Higgs exchange in SUSY $SU(5)$ GUTs, as mentioned above. Hence, the triplet partners of Higgs doublets must necessarily obtain mass of order the GUT scale. In addition, it is also important to note that Planck or string scale physics may independently generate the dimension-five operators, even without a GUT. These contributions must be suppressed by some underlying symmetry; for example, the same flavor symmetry which may be responsible for hierarchical fermion Yukawa matrices.

Dimension-five operators include squarks and/or sleptons. To allow for nucleon decay, these must be converted to light quarks or leptons by exchange of a gaugino or Higgsino in the SUSY SM. The nucleon lifetime is proportional to $M_G^2\, m_{\rm SUSY}^2/m_p^5$, where $m_{\rm SUSY}$ is the SUSY breaking scale. Thus, dimension-five operators may predict a shorter nucleon lifetime than dimension-six operators. Unless accidental cancellations are present, the dominant decay modes from dimension-five operators include a K meson, such as $p \to K^+\, \bar{\nu}$ ($n \to K^0\, \bar{\nu}$). This is due to a simple symmetry argument: The operators are given as $(Q_i\, Q_j\, Q_k\, L_l)$ and $(u_i^c\, u_j^c\, d_k^c\, e_l^c)$, where i, j, k, $l\,(=1\text{–}3)$ are family indices and color and weak indices are implicit. They must be invariant under $SU(3)_C$ and $SU(2)_L$ so that their color and weak doublet indices must be anti-symmetrized. 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 exists in the dominant modes of nucleon decay unless these modes are accidentally suppressed [85].

Recent Super-Kamiokande bounds on the proton lifetime severely constrain the dimension-six and five operators. With 306 kton-years of data they find $\tau_p/\mathrm{Br}(p \to e^+\pi^0) > 1.67 \times 10^{34}$ years and $\tau_p/\mathrm{Br}(p \to K^+\bar{\nu}) > 6.61 \times 10^{33}$ years at 90% CL [87]. The hadronic matrix elements for baryon-number-violating operators are evaluated with lattice QCD simulations [88]. The lower bound on the X boson mass from null results in nucleon decay searches is approaching 10^{16} GeV in SUSY $SU(5)$ GUTs [89]. In the minimal SUSY $SU(5)$, $\tau_p/\mathrm{Br}(p \to K^+\bar{\nu})$ is smaller than about 10^{31} years if the triplet Higgs mass is 10^{16} GeV and $m_{\rm SUSY} = 1$ TeV [90]. The triplet Higgs mass bound from nucleon decay is then in conflict with gauge coupling unification so that this model is considered to be ruled out [70].

Since nucleon decay induced by the triplet Higgs is a severe problem in SUSY GUTs, various proposals for its suppression have been made. First, some accidental symmetry or accidental structure in non-minimal Higgs sectors in $SU(5)$ or $SO(10)$ theories may suppress the dimension-five operators [25,26,21,91]. As mentioned above, the triplet Higgs mass term violates symmetries which forbid the dimension-five operators. In other words, the nucleon decay is suppressed if the Higgs triplets in $\bar{\mathbf{5}}_{\mathbf{H}}$ and $\mathbf{5}_{\mathbf{H}}$ do not have a common mass term but, instead, their mass terms involve partners from other $SU(5)$ multiplets. Second, the SUSY breaking scale may be around $\mathcal{O}(10\text{–}100)$ TeV in order to explain the observed Higgs boson mass at the LHC. In this case, nucleon decay is automatically suppressed [74,92,93]. Third, accidental cancellations among diagrams due to a fine-tuned structure of squark and slepton flavor mixing might suppress nucleon decay [94]. Last, we have also implicitly assumed a hierarchical structure for Yukawa matrices in the analysis. It is however possible to fine-tune a hierarchical structure for quarks and leptons which baffles the family structure so that the nucleon decay is suppressed [95]. The upper bound on the proton lifetime from some of these theories is approximately a factor of 10 above the experimental bounds. Future experiments with larger neutrino detectors, such as JUNO [96], Hyper-Kamiokande [97] and DUNE [98], are planned and will have higher sensitivities to nucleon decay.

Appealing to global symmetries to suppress specific interactions may not always be as straightforward as it naively seems, as a general remark, while global symmetries are introduced to control the dimension-four and five operators in SUSY GUTs. Indeed, there are two possibilities: On the one hand, the relevant symmetry might be gauged at a higher scale. Effects of the VEVs responsible for the spontaneous breaking are then in principle dangerous and need to be quantified. On the other hand, the symmetry might be truly only global. This must e.g. be the case for anomalous symmetries, which are then also violated by field-theoretic non-perturbative effects. The latter can in principle be exponentially small. It is, however, widely believed that global symmetries are always broken in quantum gravity (see e.g. [99]). One then needs to understand which power or functional form the Planck scale suppression of the relevant interaction has. For example, dimension-five baryon number violating operators suppressed by just one unit of the Planck or string scale are completely excluded.

In view of the above, it is also useful to recall that in string models 4d global symmetries generally originate in higher-dimensional gauge symmetries. Here 'global' implies that the gauge boson has acquired a Stückelberg-mass. This is a necessity in the anomalous case (Green-Schwarz mechanism) but can also happen to non-anomalous symmetries. One expects no symmetry violation beyond the well-understood non-perturbative effects. Discrete symmetries arise as subgroups of continuous gauge symmetries, such as $Z_N \subset U(1)$. In particular, non-anomalous subgroups of Stückelberg-massive $U(1)$s represent unbroken discrete gauge symmetries and as such are non-perturbatively exact (see e.g. [100]). Of course, such discrete gauge symmetries may also arise as remnants of continuous gauge symmetries after conventional 4d spontaneous breaking.

Are there ways to avoid the stringent predictions for proton decay discussed above? Orbifold GUTs and string theories, see Sec. 114.4, contain grand unified symmetries realized in higher dimensions. In the process of compactification and GUT symmetry breaking, the triplet Higgs states may be removed (projected out of the massless sector of the theory). In such models, the nucleon decay due to dimension-five operators can be severely suppressed or eliminated completely. However, nucleon decay due to dimension-six operators may be enhanced, since the gauge-bosons mediating proton decay obtain mass at the compactification scale, M_c, which is typically less than the

4d GUT scale (cf. Sec. 114.5). Alternatively, the same projections which eliminate the triplet Higgs may rearrange the quark and lepton states such that the massless states of one family come from different higher-dimensional GUT multiplets. This can suppress or completely eliminate even dimension-six proton decay. Thus, enhancement or suppression of dimension-six proton decay is model-dependent. In some complete 5d orbifold GUT models [101,58] the lifetime for the decay $\tau_p/\mathrm{Br}(p \to e^+\pi^0)$ can be near the bound of 1×10^{34} years with, however, large model-dependence and/or theoretical uncertainties. In other cases, the modes $p \to K^+\bar{\nu}$ and $p \to K^0\mu^+$ may be dominant [58]. Thus, interestingly, the observation of nucleon decay may distinguish string or higher-dimensional GUTs from 4d ones.

In orbifold GUTs or string theory, new discrete symmetries consistent with SUSY GUTs can forbid all dimension-three and four baryon- and lepton-number-violating operators. Even the μ term and dimension-five baryon- and lepton-number-violating operators can be forbidden to all orders in perturbation theory [32]. The μ term and dimension-five baryon- and lepton-number-violating operators may then be generated, albeit sufficiently suppressed, via non-perturbative effects. The simplest example of this is a Z_4^R symmetry which is the unique discrete R symmetry consistent with $SO(10)$ [32]. Even though it forbids the dimension-five proton decay operator to the desired level, it allows the required dimension-five neutrino mass term. In this case, proton decay is dominated by dimension-six operators, leading to decays such as $p \to e^+\pi^0$.

114.7. Yukawa coupling unification

In the SM, masses and mixings for quarks and leptons come from the Yukawa couplings with the Higgs doublet, but the values of these couplings remain a mystery. GUTs provide at least a partial understanding since each generation is embedded in unified multiplet(s). Specifically, since quarks and leptons are two sides of the same coin, the GUT symmetry relates the Yukawa couplings (and hence the masses) of quarks and leptons.

In $SU(5)$, there are two types of independent renormalizable Yukawa interactions given by $\lambda_{ij}\ (\mathbf{10}_i\ \mathbf{10}_j\ \mathbf{5_H}) + \lambda'_{ij}\ (\mathbf{10}_i\ \bar{\mathbf{5}}_j\ \bar{\mathbf{5}}_\mathbf{H})$. These contain the SM interactions $\lambda_{ij}\ (Q_i\ u^c_j\ H_u)$ $+\ \lambda'_{ij}\ (Q_i\ d^c_j\ H_d\ +\ e^c_i\ L_j\ H_d)$. Here $i, j\,(=1\text{–}3)$ are, as before, family indices. Hence, at the GUT scale we have tree-level relations between Yukawa coupling constants for charged lepton and down quark masses, such as $\lambda_b = \lambda_\tau$ in which $\lambda_{b/\tau}$ are the bottom quark $/\tau$ lepton Yukawa coupling constants [102,103]. In $SO(10)$, there is only one type of independent renormalizable Yukawa interaction given by $\lambda_{ij}\ (\mathbf{16}_i\ \mathbf{16}_j\ \mathbf{10_H})$, leading to relations among all Yukawa coupling constants and quark and lepton masses within one generation [104,105] (such as $\lambda_t = \lambda_b = \lambda_\tau$, with λ_t the top quark Yukawa coupling constant).

114.7.1. *The third generation, b–τ or t–b–τ unification* :

Third generation Yukawa couplings are larger than those of the first two generations. Hence, the fermion mass relations predicted from renormalizable GUT interactions which we introduced above are expected to be more reliable. In order to compare them with data, we have to include the radiative correction to these relations from the RG evolution between GUT and fermion mass scale, from integrating out heavy particles at the GUT scale, and from weak scale thresholds.

Since testing Yukawa coupling unification is only possible in models with successful gauge coupling unification, we here focus on SUSY GUTs. In the MSSM, top and bottom quark and τ lepton masses are related to the Yukawa coupling constants at the scale m_Z as

$$m_t(m_Z) = \lambda_t(m_Z)\, v_u(1 + \delta m_t/m_t), \quad m_{b/\tau}(m_Z) = \lambda_{b/\tau}(m_Z)\, v_d(1 + \delta m_{b/\tau}/m_{b/\tau}),$$

where $\langle H_u^0\rangle \equiv v_u = \sin\beta\, v/\sqrt{2}$, $\langle H_d^0\rangle \equiv v_d = \cos\beta\, v/\sqrt{2}$, $v_u/v_d \equiv \tan\beta$ and $v \sim 246$ GeV is fixed by the Fermi constant, G_μ. Here, $\delta m_f/m_f$ $(f = t, b, \tau)$ represents the threshold correction due to integrating out SUSY partners. For the bottom quark mass, it is found [106] that the dominant corrections come from the gluino-sbottom and from the Higgsino-stop loops,

$$\left(\frac{\delta m_b}{m_b}\right)_{g_3} \sim \frac{g_3^2}{6\pi^2}\frac{m_{\widetilde{g}}\mu}{m_{\rm SUSY}^2}\tan\beta \qquad \text{and} \qquad \left(\frac{\delta m_b}{m_b}\right)_{\lambda_t} \sim \frac{\lambda_t^2}{16\pi^2}\frac{A_t\mu}{m_{\rm SUSY}^2}\tan\beta\,, \tag{114.15}$$

where $m_{\widetilde{g}}$, μ, and A_t stand for gluino and Higgsino masses and trilinear stop coupling, respectively. Note that Eq. (114.15) only illustrates the structure of the corrections – non-trivial functional dependences on several soft parameters $\sim m_{\rm SUSY}$ have been suppressed. For the full one-loop correction to the bottom quark mass see, for example, Ref. [107].

Note also that the corrections do not go to zero as SUSY particles become much heavier than m_Z. They may change the bottom quark mass at the $\mathcal{O}(10)\%$ level for $\tan\beta = \mathcal{O}(10)$. The total effect is sensitive to the relative phase between gluino and Higgsino masses since $A_t \sim -m_{\widetilde{g}}$ due to the infrared fixed point nature of the RG equation for A_t [108] in settings where SUSY breaking terms come from Planck scale dynamics, such as gravity mediation. The τ lepton mass also receives a similar correction, though only at the few % level. The top quark mass correction, not being proportional to $\tan\beta$, is at most 10% [109].

Including one loop threshold corrections at m_Z and additional RG running, one finds the top, bottom and τ pole masses. In SUSY GUTs, b–τ unification has two possible solutions with $\tan\beta \sim 1$ or $\mathcal{O}(10)$. The small $\tan\beta$ solution may be realized in the MSSM if superpartner masses are $\mathcal{O}(10)$ TeV, as suggested by the observed Higgs mass [92]. The large $\tan\beta$ limit such as $\tan\beta \sim 40$–50 overlaps the $SO(10)$ symmetry relation [109]. When $\tan\beta$ is large, there are significant threshold corrections to down quark masses as mentioned above, and Yukawa unification is only consistent with low-energy data in a restricted region of SUSY parameter space, with important consequences for SUSY searches [109,110]. More recent analyses of Yukawa unification after LHC Run-I are found in Ref. [111].

Gauge coupling unification is also successful in the scenario of split supersymmetry [74], in which squarks and sleptons have mass at a scale $\tilde{m} \gg m_Z$, while gauginos and/or Higgsinos have masses of order the weak scale. Unification of b–τ Yukawa couplings requires $\tan\beta$ to be fine-tuned close to 1 [92]. If by contrast, $\tan\beta \gtrsim 1.5$, b–τ Yukawa unification only works for $\tilde{m} \lesssim 10^4$ GeV. This is because the effective theory between the gaugino mass scale and $\tilde{m}$ includes only one Higgs doublet, as in the standard model. As a result, the large top quark Yukawa coupling tends to increase the ratio λ_b/λ_τ due to the vertex correction, which is absent in supersymmetric theories, as one runs down in energy below $\tilde{m}$. This is opposite to what happens in the MSSM where the large top quark Yukawa coupling lowers the ratio λ_b/λ_τ [103].

114.7.2. *Beyond leading order: three-family models* :

Simple Yukawa unification is not possible for the first two generations. Indeed, $SU(5)$ implies $\lambda_s = \lambda_\mu$, $\lambda_d = \lambda_e$ and hence $\lambda_s/\lambda_d = \lambda_\mu/\lambda_e$. This is an RG-invariant relation which extrapolates to $m_s/m_d = m_\mu/m_e$ at the weak scale, in serious disagreement with data ($m_s/m_d \sim 20$ and $m_\mu/m_e \sim 200$). An elegant solution to this problem was given by Georgi and Jarlskog [112] (for a recent analysis in the SUSY context see [113]) .

More generally, we have to recall that in all of the previous discussion of Yukawa couplings, we assumed renormalizable interactions as well as the minimal matter and Higgs content. Since the GUT scale is close to the Planck scale, higher-dimension operators involving the GUT-breaking Higgs may modify the predictions, especially for lower generations. An example is provided by the operators $\mathbf{10}\ \bar{\mathbf{5}}\ \bar{\mathbf{5}}_{\mathbf{H}}\ \mathbf{24}_{\mathbf{H}}$ with $\mathbf{24}_{\mathbf{H}}$ the GUT-breaking Higgs of $SU(5)$. We can fit parameters to the observed fermion masses with these operators, though some fine-tuning is introduced in doing so. The SM Higgs doublet may come in part from higher representations of the GUT group. For example, the **45** of $SU(5)$ includes an $SU(2)_L$ doublet with appropriate $U(1)_Y$ charge [112]. This **45** can, in turn, come from the **120** or **126** of $SO(10)$ after its breaking to $SU(5)$ [114]. These fields may also have renormalizable couplings with quarks and leptons. The relations among the Yukawa coupling constants in the SM are modified if the SM Higgs doublet is a linear combination of several such doublets from different $SU(5)$ multiplets. Finally, the SM fermions may not be embedded in GUT multiplets in the minimal way. Indeed, if all quarks and leptons are embedded in **16**s of $SO(10)$, the renormalizable interactions with $\mathbf{10}_{\mathbf{H}}$ cannot explain the observed CKM mixing angles. This situation improves when extra matter multiplets, such as **10**, are introduced: After $U(1)_X$, which distinguishes the $\bar{\mathbf{5}}$s coming from the **16** and the **10** of $SO(10)$, is broken (e.g. by a VEV of $\mathbf{16}_{\mathbf{H}}$ or $\mathbf{126}_{\mathbf{H}}$), the r.h. down quarks and l.h. leptons in the SM can be linear combinations of components in **16**s and **10**s. As a result, $\lambda \neq \lambda'$ in $SU(5)$ [115].

To construct realistic three-family models, some or all of the above effects can be used. Even so, to achieve significant predictions for fermion masses and mixing angles grand unification alone is not sufficient. Other ingredients, for example additional global family symmetries are needed (in particular, non-abelian symmetries can strongly reduce the number of free parameters). These family symmetries constrain the set of effective higher-dimensional fermion mass operators discussed above. In addition, sequential breaking of the family symmetry can be correlated with the hierarchy of fermion masses. One simple, widely known idea in this context is to ensure that each $\mathbf{10}_i$ enters Yukawa interactions together with a suppression factor ϵ^{3-i} (ϵ being a small parameter). This way one automatically generates a stronger hierarchy in up-type quark Yukawas as compared to down-type quark and lepton Yukawas and no hierarchy for neutrinos, which agrees with observations at the $\mathcal{O}(1)$-level. Three-family models exist which fit all the data, including neutrino masses and mixing [26,116].

Finally, a particularly ambitious variant of unification is to require that the fermions of all three generations come from a single representation of a large gauge group. A somewhat weaker assumption is that the flavor group (e.g. $SU(3)$) unifies with the SM gauge group in a simple gauge group at some energy scale $M \geq M_G$. Early work on such 'flavor-unified GUTs', see e.g. [117], has been reviewed in [118,119]. For a selection of more recent papers see [120]. In such settings, Yukawa couplings are generally determined by gauge couplings together with symmetry breaking VEVs. This is reminiscent of heterotic string GUTs, where all couplings come from the 10d gauge coupling. However, while the $\mathrm{E}_8 \to SU(3)\times\mathrm{E}_6$ branching rule $\mathbf{248} = (\mathbf{8},\mathbf{1}) + (\mathbf{1},\mathbf{78}) + (\mathbf{3},\mathbf{27}) + (\overline{\mathbf{3}},\overline{\mathbf{27}})$ looks very suggestive in this context, the way in which most modern heterotic models arrive at three generations is actually more complicated.

114.7.3. *Flavor violation* :

Yukawa interactions of GUT-scale particles with quarks and leptons may leave imprints on the flavor violation induced by SUSY breaking parameters [121]. To understand this, focus first on the MSSM with universal Planck-scale boundary conditions (as e.g. in gravity mediation). Working in a basis where up-quark and lepton Yukawas are diagonal, one finds that the large top-quark Yukawa coupling reduces the l.h. squark mass squareds in the third generation radiatively. It turns out that only the l.h. down-type squark mass matrix has sizable off-diagonal terms in the flavor basis after CKM-rotation. However, in GUTs the color-triplet Higgs has flavor violating interactions from the Yukawa coupling λ_{ij} ($\mathbf{10}_i$ $\mathbf{10}_j$ $\mathbf{5_H}$), such that flavor-violating r.h. slepton mass terms are radiatively generated in addition [122]. If r.h. neutrinos are introduced as $SU(5)$ singlets with interactions λ''_{ij} ($\mathbf{1}_i$ $\bar{\mathbf{5}}_j$ $\mathbf{5_H}$), the doublet and color-triplet Higgses acquire another type of Yukawa coupling, respectively. They then radiatively generate flavor-violating l.h. slepton [123] and r.h. down squark masses [124]. These flavor-violating SUSY breaking terms induce new contributions to FCNC processes in quark and lepton sectors, such as $\mu \to e\gamma$ and K^0–$\bar{K}^0$ and B^0–$\bar{B}^0$ mixing. EDMs are also induced when both l.h. and r.h. squarks/sleptons have flavor-violating mass terms with relative phases, as discussed for $SO(10)$ in [125] or for $SU(5)$ with r.h. neutrinos in [126]. Thus, such low-energy observables constrain GUT-scale interactions.

114.8. Neutrino masses

We see from atmospheric and solar neutrino oscillation observations, along with long baseline accelerator and reactor experiments, that neutrinos have finite masses. By adding three "sterile" neutrinos ν_i^c with Yukawa couplings $\lambda_{\nu,ij}$ $(\nu_i^c\ L_j\ H_u)$ $(i,j = 1\text{–}3)$, one easily obtains three massive Dirac neutrinos with mass $m_\nu = \lambda_\nu\ v_u$, analogously to quark and charged lepton masses. However, in order to obtain a τ neutrino with mass of order 0.1 eV, one requires the exceedingly small coupling ratio $\lambda_{\nu_\tau}/\lambda_\tau \lesssim 10^{-10}$. By contrast, the seesaw mechanism *naturally* explains such tiny neutrino masses as follows [1,2,3]: The sterile neutrinos have no SM gauge quantum numbers so that there is no symmetry other than global lepton number which forbids the Majorana mass term $\frac{1}{2}M_{ij}\ \nu_i^c\ \nu_j^c$. Note also that sterile neutrinos can be identified with the r.h. neutrinos necessarily contained in complete families of $SO(10)$ or Pati-Salam models. Since the Majorana mass term violates $U(1)_X$ in $SO(10)$, one might expect $M_{ij} \sim M_G$. The heavy sterile neutrinos can be integrated out, defining an effective low-energy theory with only three light active Majorana neutrinos with the effective dimension-five operator

$$-\mathcal{L}_{eff} = \frac{1}{2}c_{ij}\ (L_i\ H_u)\ (L_j\ H_u)\,, \tag{114.16}$$

where $c = \lambda_\nu^T\ M^{-1}\ \lambda_\nu$. This then leads to a 3×3 Majorana neutrino mass matrix $m = m_\nu^T\ M^{-1}\ m_\nu$.

Atmospheric neutrino oscillations require neutrino masses with $\Delta m_\nu^2 \sim 2.5\times 10^{-3}$ eV^2 with maximal mixing, in the simplest scenario of two neutrino dominance. With hierarchical neutrino masses this implies $m_{\nu_\tau} = \sqrt{\Delta m_\nu^2} \sim 0.05$ eV. Next, we can try to relate the neutrino Yukawa coupling to the top quark Yukawa coupling, $\lambda_{\nu_\tau} = \lambda_t$ at the GUT scale, as in $SO(10)$ or $SU(4)\times SU(2)_L\times SU(2)_R$ models. This gives $M \sim 10^{14}$ GeV, which is remarkably close to the GUT scale.

Neutrinos pose a special problem for SUSY GUTs. The question is why the quark mixing angles in the CKM matrix are small while there are two large lepton mixing angles in the PMNS matrix (cf. however the comment at the end of Sec. 114.7). Discussions of neutrino masses and mixing angles can, for example, be found in Refs. [127] and [128]. For SUSY GUT models which fit quark and lepton masses, see Ref. [25]. Finally, for a compilation of the range of SUSY GUT predictions for neutrino mixing, see [129].

The seesaw mechanism implemented by r.h. neutrinos is sometimes called the type-I seesaw model. There are variant models in which the dimension-five operator for neutrino masses is induced in different ways: In the type-II model, an $SU(2)_L$ triplet Higgs boson Σ is introduced to have couplings ΣL^2 and also ΣH_u^2 [130]. In the type-III model, an $SU(2)_L$ triplet of fermions $\tilde\Sigma$ with a Yukawa coupling $\tilde\Sigma L H_u$ is introduced [131]. In these models, the dimension-five operator is induced by integrating out the triplet Higgs boson or fermions. Such models can also be implemented in GUTs by introducing Higgs bosons in the **15** or fermions in the **24** in $SU(5)$ GUTs or the **126** in $SO(10)$ GUTs. Notice that the gauge non-singlet fields in the type-II and III models have masses at the intermediate scale. Thus, gauge coupling unification is not automatic if they are implemented in SUSY GUTs.

114.9. Selected topics

114.9.1. *Magnetic monopoles* :

In the broken phase of a GUT there are typically localized classical solutions carrying magnetic charge under an unbroken $U(1)$ symmetry [132]. These magnetic monopoles with mass of order M_G/α_G can be produced during a possible GUT phase transition in the early universe. The flux of magnetic monopoles is experimentally found to be less than $\sim 10^{-16}$ cm^{-2} s^{-1} sr^{-1} [133]. Many more are however predicted, hence the GUT monopole problem. In fact, one of the original motivations for inflation was to solve the monopole problem by exponential expansion after the GUT phase transition [134] and hence dilution of the monopole density. Other possible solutions to the monopole problem include: sweeping them away by domain walls [135], $U(1)$ electromagnetic symmetry breaking at high temperature [136] or GUT symmetry non-restoration [137]. Parenthetically, it was also shown that GUT monopoles can catalyze nucleon decay [138]. A significantly stronger bound on the monopole flux can then be obtained by considering X-ray emission from radio pulsars due to monopole capture and the subsequent nucleon decay catalysis [139].

Note that the present upper bound on the inflationary vacuum energy density is very close to the GUT scale, $V_{inf}^{1/4} = (1.88 \times 10^{16}\,\mathrm{GeV}) \times (r/0.10)^{1/4}$, with the scalar-to-tensor ratio constraint to $r < 0.11$ [140]. This guarantees that reheating does not lead to temperatures above M_G and hence the monopole problem is solved by inflation (unless M_G is unexpectedly low).

114.9.2. *Anomaly constraints vs. GUT paradigm* :

As emphasized at the very beginning, the fact that the SM fermions of one generation fill out the $\mathbf{10} + \overline{\mathbf{5}}$ of $SU(5)$ appears to provide overwhelming evidence for some form of GUT embedding. However, one should be aware that a counterargument can be made which is related to the issue of 'charge quantization by anomaly cancellation' (see [141,142] for some early papers and [143] for a more detailed reference list): Imagine we only knew that the low-energy gauge group were G_{SM} and the matter content included the $(\mathbf{3}, \mathbf{2})_Y$, i.e. a 'quark doublet' with $U(1)$-charge Y. One can then ask which possibilities exist of adding further matter to ensure the cancellation of all triangle anomalies. It turns out that this problem has only three different, minimal* solutions [142]. One of those is precisely a single SM generation, with the apparent '$SU(5)$-ness' emerging accidentally. Thus, if one randomly picks models from the set of consistent gauge theories, preconditioning on G_{SM} and $(\mathbf{3}, \mathbf{2})_Y$, one may easily end up with '$\mathbf{10} + \overline{\mathbf{5}}$' of an $SU(5)$ that is in no way dynamically present. This is precisely what happens in the context of non-GUT string model building [144].

* Adding extra vector-like sets of fields, e.g. two fermions which only transform under $U(1)$ and have charges Y and $-Y$, is considered to violate minimality.

114.9.3. ***GUT baryogenesis and leptogenesis*** :

During inflation, any conserved quantum number is extremely diluted. Thus, one expects the observed baryon asymmetry of the universe to originate at reheating or in the subsequent cosmological evolution. In detail, the situation is slightly more involved: Both baryon number B and lepton number L are global symmetries of the SM. However, $(B+L)$ is anomalous and violated by thermal fluctuations in the early universe, via so-called sphaleron processes. Moreover, it is violated in GUT models, as is most apparent in proton decay. By contrast, $(B-L)$ is anomaly free and preserved by both the SM as well as $SU(5)$ or $SO(10)$ gauge interactions.

Now, the old idea of GUT baryogenesis [145,146] is to generate a $(B+L)$ and hence a baryon asymmetry by the out-of-equilibrium decay of the color-triplet Higgs. However such an asymmetry, generated at GUT temperatures, is washed out by sphalerons. This can be overcome [147] using lepton-number violating interaction of neutrinos to create a $(B-L)$ from the $(B+L)$ asymmetry, before sphaleron processes become sufficiently fast at $T < 10^{12}$ GeV. This $(B-L)$ asymmetry can then survive the subsequent sphaleron dominated phase. Note that this does not work in the minimal SUSY GUT setting, with the triplet Higgs above the GUT scale. The reason is that a correspondingly high reheating temperature would be required which, as explained above, is ruled out by Planck data.

However, the most widely accepted simple way out of the dilemma is to directly generate a net $(B-L)$ asymmetry dynamically in the early universe, also using r.h. neutrinos. Indeed, we have seen that neutrino oscillations suggest a new scale of physics of order 10^{14} GeV. This scale is associated with heavy Majorana neutrinos in the seesaw mechanism. 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 [148]. This mechanism is called leptogenesis.

If the three heavy Majorana neutrino masses are hierarchical, the net lepton number is produced by decay of the lightest one, and it is proportional to the CP asymmetry in the decay, ϵ_1. The CP asymmetry is bounded from above, and the lightest neutrino mass is required to be larger than 10^9 GeV in order to explain the observed baryon asymmetry [149]. This implies that the reheating temperature after inflation should be larger than 10^9 GeV so that the heavy neutrinos are thermally produced. In supersymmetric models, there is a tension between leptogenesis and Big Bang Nucleosynthesis (BBN) if gravitinos decay in the BBN era. The gravitino problem gives a constraint on the reheating temperature $\lesssim 10^{6-10}$ GeV though the precise value depends on the SUSY breaking parameters [150]. Recent reviews of leptogenesis can be found in Ref. [151].

114.10. Conclusion

Most conservatively, grand unification means that (some of) the SM gauge interactions of $U(1)_Y$, $SU(2)_L$ and $SU(3)_C$ become part of a larger, unifying gauge symmetry at a high energy scale. In most models, especially in the simplest and most appealing variants of $SU(5)$ and $SO(10)$ unification, the statement is much stronger: One expects the three gauge couplings to unify (up to small threshold corrections) at a unique scale, M_G, and the proton to be unstable due to exchange of gauge bosons of the larger symmetry group. Supersymmetric grand unified theories provide, by far, the most predictive and economical framework allowing for perturbative unification. For a selection of reviews, with many more details than could be discussed in the present article, see [118,152].

Thus, the three classical pillars of GUTs are gauge coupling unification at $M_G \sim 2\times10^{16}$ GeV, low-energy supersymmetry (with a large SUSY desert), and nucleon decay. The first of these may be viewed as predicting the value of the strong coupling – a prediction which has already been verified (see Fig. 114.1). Numerically, this prediction remains intact even if SUSY partner masses are somewhat above the weak scale. However, at the conceptual level a continuously increasing lower bound on the SUSY scale is nevertheless problematic for the GUT paradigm: Indeed, if the independent, gauge-hierarchy-based motivation for SUSY is completely abandoned, the SUSY scale and hence α_3 become simply free parameters and the first two pillars crumble. Thus, it is important to keep pushing bounds on proton decay which, although again not completely universal in all GUT constructions, is arguably a more generic part of the GUT paradigm than low-energy SUSY.

Whether or not Yukawa couplings unify is more model dependent. However, irrespective of possible (partial) Yukawa unification, there certainly exists a very interesting and potentially fruitful interplay between flavor model building and grand unification. Especially in the neutrino sector this is strongly influenced by the developing experimental situation.

Another phenomenological signature of grand unification is the strength of the direct coupling of the QCD axion to photons, relative to its coupling to gluons. It is quantified by the predicted anomaly ratio $E/N = 8/3$ (see [153,154]). This arises in field-theoretic axion models consistent with GUT symmetry (such as DFSZ [155]) and in string-theoretic GUTs [154]. In the latter, the axion does not come from the phase of a complex scalar but is a fundamental shift-symmetric real field, coupling through a higher-dimension operator directly to the product of the GUT field-strength and its dual.

It is probably fair to say that, due to limitations of the 4d approach, including especially remaining ambiguities (free parameters or ad hoc assumptions) in models of flavor and GUT breaking, the string theoretic approach has become more important in GUT model building. In this framework, challenges include learning how to deal with the many vacua of the 'landscape' as well as, for each vacuum, developing the tools for reliably calculating detailed, phenomenological observables. Finally, due to limitations of space, the present article has barely touched on the interesting cosmological implications of GUTs. They may become more important in the future, especially in the case that a high inflationary energy scale is established observationally.

References:

1. P. Minkowski, Phys. Lett. **B67**, 421 (1977).
2. 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;
S. Glashow, Quarks and leptons, published in *Proceedings of the Carg'ese Lectures*, M. Levy (ed.), Plenum Press, New York, (1980);
M. Gell-Mann, P. Ramond and R. Slansky, in *Supergravity*, ed. P. van Nieuwenhuizen *et al.*, North-Holland, Amsterdam, (1979), p. 315 [arXiv:1306.4669].
3. F. Wilczek, eConf C **790823**, 437 (1979);
E. Witten, Phys. Lett. **B91**, 81 (1980);
R.N. Mohapatra and G. Senjanovic, Phys. Rev. Lett. **44**, 912 (1980).
4. U. Amaldi, W. de Boer and H. Furstenau, Phys. Lett. **B260**, 447 (1991);
J.R. Ellis, S. Kelley and D.V. Nanopoulos, Phys. Lett. **B260**, 131 (1991);
P. Langacker and M. Luo, Phys. Rev. **D44**, 817 (1991);
C. Giunti, C.W. Kim and U.W. Lee, Mod. Phys. Lett. **A6**, 1745 (1991);
P. Langacker and N. Polonsky, Phys. Rev. **D47**, 4028 (1993), [hep-ph/9210235];
M. Carena, S. Pokorski and C.E.M. Wagner, Nucl. Phys. **B406**, 59 (1993) [hep-ph/9303202];
See also the review by S. Dimopoulos, S.A. Raby and F. Wilczek, Physics Today, p. 25 October (1991).
5. Y. Okada, M. Yamaguchi and T. Yanagida, Prog. Theor. Phys. **85**, 1 (1991);
J.R. Ellis, G. Ridolfi and F. Zwirner, Phys. Lett. **B257**, 83 (1991);
H.E. Haber and R. Hempfling, Phys. Rev. Lett. **66**, 1815 (1991).
6. M.S. Carena *et al.*, Phys. Lett. **B355**, 209 (1995) [hep-ph/9504316];
G. Degrassi *et al.*, Eur. Phys. J. **C28**, 133 (2003) [arXiv:hep-ph/0212020];
P. Kant *et al.*, JHEP **1008**, 104 (2010) [arXiv:1005.5709].
7. M.J.G. Veltman, Acta Phys. Pol. **B12**, 437 (1981);
L. Maiani, Gif-sur-Yvette Summer School on Particle Physics, 11th, Gif-sur-Yvette, France, 1979 (Inst. Nat. Phys. Nucl. Phys. Particules, Paris, 1979);
E. Witten, Nucl. Phys. **B188**, 513 (1981).
8. J.L. Feng *et al.*, Phys. Rev. Lett. **111**, 131802 (2013) [arXiv:1306.2318];
S. Heinemeyer, Lecture at 42nd ITEP winter school, Moscow, arXiv:1405.3781;
P. Draper and H. Rzehak, Phys. Rept. **619**, 1 (2016) [arXiv:1601.01890].
9. 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); see also J. C. Pati, Int. J. Mod. Phys. A **32**, 1741013 (2017), arXiv:1706.09531 for a recent account.
10. H. Georgi and S.L Glashow, Phys. Rev. Lett. **32**, 438 (1974).
11. J.R. Ellis, M.K. Gaillard and D.V. Nanopoulos, Phys. Lett. **B80**, 360 (1979) [Phys. Lett. **B82**, 464 (1979)];
E. Golowich, Phys. Rev. **D24**, 2899 (1981).

12. 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. **B105**, 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).
13. S. Weinberg, Phys. Rev. **D26**, 287 (1982);
N. Sakai and T. Yanagida, Nucl. Phys. **B197**, 533 (1982).
14. 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).
15. R. Slansky, Phys. Reports **79**, 1 (1981);
H. Georgi, Front. Phys. **54**, 1 (1982);
R. Feger and T.W. Kephart, Comp. Phys. Comm. **192**, 166 (2015) [arXiv:1206.6379];

N. Yamatsu, PTEP, no.4, 043B02 (2016) [arXiv:1512.05559].
16. S.M. Barr, Phys. Lett. **B112**, 219 (1982);
J.P. Derendinger, J.E. Kim and D.V. Nanopoulos, Phys. Lett. **B139**, 170 (1984);
I. Antoniadis *et al.*, Phys. Lett. **B194**, 231 (1987) and Phys. Lett. **B231**, 65 (1989).
17. F. Gursey, P. Ramond and P. Sikivie, Phys. Lett. **B60**, 177 (1976).
18. A. de Rujula, H. Georgi and S. L. Glashow, *5th Workshop on Grand Unification*, ed. K. Kang, H. Fried and P. Frampton, World Scientific, Singapore (1984), p. 88; 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).
19. 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);
J.E. Kim, Phys. Lett. **B591**, 119 (2004) [hep-ph/0403196].
20. E. Witten, Phys. Lett. **B105**, 267 (1981).
21. A. Masiero *et al.*, Phys. Lett. **B115**, 380 (1982);
B. Grinstein, Nucl. Phys. **B206**, 387 (1982).
22. S. Dimopoulos and F. Wilczek, *Proceedings Erice Summer School*, ed. A. Zichichi (1981);
M. Srednicki, Nucl. Phys. **B202**, 327 (1982).
23. K. Inoue, A. Kakuto and H. Takano, Prog. Theor. Phys. **75**, 664 (1986).
24. Y. Yamada, Z. Phys. **C60**, 83 (1993);
J. Hisano *et al.*, Phys. Lett. **B342**, 138 (1995) [hep-ph/9406417];
G. Altarelli, F. Feruglio and I. Masina, JHEP **0011**, 040 (2000) [hep-ph/0007254].
25. K.S. Babu and S.M. Barr, Phys. Rev. **D48**, 5354 (1993) [hep-ph/9306242];
K.S. Babu and S.M. Barr, Phys. Rev. **D50**, 3529 (1994) [hep-ph/9402291];
K.S. Babu, J.C. Pati and Z. Tavartkiladze, JHEP **1006**, 084 (2010) [arXiv:1003.2625].
26. K.S. Babu and R.N. Mohapatra, Phys. Rev. Lett. **74**, 2418 (1995) [hep-ph/9410326];

V. Lucas and S. Raby, Phys. Rev. **D54**, 2261 (1996) [hep-ph/9601303];
T. Blažek *et al.*, Phys. Rev. **D56**, 6919 (1997) [hep-ph/9611217];
S.M. Barr and S. Raby, Phys. Rev. Lett. **79**, 4748 (1997) [hep-ph/9705366];
K.S. Babu, J.C. Pati and F. Wilczek, Nucl. Phys. **B566**, 33 (2000) [hep-ph/9812538];
R. Dermíšek, A. Mafi and S. Raby, Phys. Rev. **D63**, 035001 (2001) [hep-ph/0007213].

27. 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) [hep-ph/9501336];
Q. Shafi and Z. Tavartkiladze, Phys. Lett. **B522**, 102 (2001) [hep-ph/0105140].
28. G.F. Giudice and A. Masiero, Phys. Lett. **B206**, 480 (1988);
J.E. Kim and H.P. Nilles, Mod. Phys. Lett. **A9**, 3575 (1994) [hep-ph/9406296].
29. L. Randall and C. Csaki, in *Palaiseau 1995, SUSY 95* 99-109 [hep-ph/9508208].
30. E. Witten, in *Hamburg 2002, SUSY 02* [hep-ph/0201018];
M. Dine, Y. Nir and Y. Shadmi, Phys. Rev. **D66**, 115001 (2002) [hep-ph/0206268].
31. A. Hebecker, J. March-Russell and R. Ziegler, JHEP **0908**, 064 (2009) [arXiv:0801.4101];
F. Brümmer *et al.*, JHEP **0908**, 011 (2009) [arXiv:0906.2957];
F. Brümmer *et al.*, JHEP **1004**, 006 (2010) [arXiv:1003.0084].
32. H.M. Lee *et al.*, Phys. Lett. **B694**, 491 (2011) [arXiv:1009.0905];
R. Kappl *et al.*, Nucl. Phys. **B847**, 325 (2011) [arXiv:1012.4574];
H.M. Lee *et al.*, Nucl. Phys. **B850**, 1 (2011) [arXiv:1102.3595].
33. Y. Hosotani, Phys. Lett. **B126**, 309 (1983).
34. L.J. Dixon *et al.*, Nucl. Phys. **B261**, 678 (1985) and Nucl. Phys. **B274**, 285 (1986);
L.E. Ibanez, H.P. Nilles and F. Quevedo, Phys. Lett. **B187**, 25 (1987);
L.E. Ibanez *et al.*, Phys. Lett. **B191**, 282 (1987).
35. P. Candelas *et al.*, Nucl. Phys. **B258**, 46 (1985).
36. Y. Kawamura, Prog. Theor. Phys. **103**, 613 (2000) [hep-ph/9902423] and Prog. Theor. Phys. **105**, 999 (2001) [hep-ph/0012125];
G. Altarelli and F. Feruglio, Phys. Lett. **B511**, 257 (2001) [hep-ph/0102301];
L.J. Hall and Y. Nomura, Phys. Rev. **D64**, 055003 (2001) [hep-ph/0103125];
A. Hebecker and J. March-Russell, Nucl. Phys. **B613**, 3 (2001) [hep-ph/0106166];
T. Asaka, W. Buchmuller and L. Covi, Phys. Lett. **B523**, 199 (2001) [hep-ph/0108021];
L.J. Hall *et al.*, Phys. Rev. **D65**, 035008 (2002) [hep-ph/0108071];
R. Dermisek and A. Mafi, Phys. Rev. **D65**, 055002 (2002) [hep-ph/0108139];
H.D. Kim and S. Raby, JHEP **0301**, 056 (2003) [hep-ph/0212348].
37. L.E. Ibanez and A.M. Uranga, “String theory and particle physics: An introduction to string phenomenology,” Cambridge University Press 2012;
K.S. Choi and J.E. Kim, Lect. Notes Phys. **696**, 1 (2006);
R. Blumenhagen *et al.*, Phys. Reports **445**, 1 (2007) [hep-th/0610327].
38. V. Braun *et al.*, JHEP **0605**, 043 (2006) [hep-th/0512177];
V. Bouchard and R. Donagi, Phys. Lett. **B633**, 783 (2006) [hep-th/0512149];

L.B. Anderson *et al.*, Phys. Rev. **D84**, 106005 (2011) [arXiv:1106.4804].
39. L.J. Dixon, V. Kaplunovsky and J. Louis, Nucl. Phys. **B355**, 649 (1991).
40. G. Aldazabal *et al.*, Nucl. Phys. **B452**, 3 (1995) [hep-th/9410206];
Z. Kakushadze *et al.*, Int. J. Mod. Phys. **A13**, 2551 (1998) [hep-th/9710149].
41. E. Witten, Nucl. Phys. **B258**, 75 (1985).
42. S. Gukov *et al.*, Phys. Rev. **D69**, 086008 (2004) [hep-th/0310159];
G. Curio, A. Krause and D. Lust, Fortsch. Phys. **54**, 225 (2006) [hep-th/0502168];
L.B. Anderson *et al.*, Phys. Rev. **D83**, 106011 (2011) [arXiv:1102.0011].
43. R. Friedman, J. Morgan and E. Witten, Commun. Math. Phys. **187**, 679 (1997) [hep-th/9701162].
44. R. Blumenhagen, G. Honecker and T. Weigand, JHEP **0508**, 009 (2005) [hep-th/0507041];
L. B. Anderson *et al.*, JHEP **1401** (2014) 047 [arXiv:1307.4787].
45. T. Kobayashi, S. Raby and R.J. Zhang, Phys. Lett. **B593**, 262 (2004) [hep-ph/0403065].
46. S. Forste *et al.*, Phys. Rev. **D70**, 106008 (2004) [hep-th/0406208];
T. Kobayashi, S. Raby and R.J. Zhang, Nucl. Phys. **B704**, 3 (2005) [hep-ph/0409098];
W. Buchmuller *et al.*, Nucl. Phys. **B712**, 139 (2005) [hep-ph/0412318];
W. Buchmuller *et al.*, Phys. Rev. Lett. **96**, 121602 (2006) [hep-ph/0511035], and Nucl. Phys. **B785**, 149 (2007) [hep-th/0606187];
O. Lebedev, *et al.*, Phys. Lett. **B645**, 88 (2007) [hep-th/0611095];
J.E. Kim, J.H. Kim and B. Kyae, JHEP **0706**, 034 (2007) [hep-ph/0702278];
O. Lebedev, *et al.*, Phys. Rev. **D77**, 046013 (2008) [arXiv:0708.2691].
47. S.B. Giddings, S. Kachru and J. Polchinski, Phys. Rev. **D66**, 106006 (2002) [hep-th/0105097];
S. Kachru *et al.*, Phys. Rev. **D68**, 046005 (2003) [hep-th/0301240].
48. R. Blumenhagen *et al.*, Nucl. Phys. **B815**, 1 (2009) [arXiv:0811.2936].
49. R. Blumenhagen *et al.*, Nucl. Phys. **B616**, 3 (2001) [hep-th/0107138].
50. R. Donagi and M. Wijnholt, Adv. Theor. Math. Phys. **15**, 1237 (2011) [arXiv:0802.2969].
51. C. Beasley, J.J. Heckman and C. Vafa, JHEP **0901**, 058 (2009) [arXiv:0802.3391], and JHEP **0901**, 059 (2009) [arXiv:0806.0102].
52. T. Weigand, Class. Quantum Grav. **27**, 214004 (2010) [arXiv:1009.3497];
J.J. Heckman, Ann. Rev. Nucl. and Part. Sci. **160**, 237 (2010) [arXiv:1001.0577];
M. Cvetic, I. Garcia-Etxebarria and J. Halverson, JHEP **1101**, 073 (2011) [arXiv:1003.5337];
A. Maharana and E. Palti, Int. J. Mod. Phys. **A28**, 1330005 (2013) [arXiv:1212.0555];
S. Krippendorf, S. Schafer-Nameki and J.M. Wong, JHEP **1511**, 008 (2015) [arXiv:1507.05961].
53. R. Donagi and M. Wijnholt, Adv. Theor. Math. Phys. **15**, 1523 (2011) [arXiv:0808.2223];
R. Blumenhagen, Phys. Rev. Lett. **102**, 071601 (2009) [arXiv:0812.0248];

K. S. Choi and J. E. Kim, Phys. Rev. **D83**, 065016 (2011) [arXiv:1012.0847];
C. Mayrhofer, E. Palti and T. Weigand, JHEP **1309**, 082 (2013) [arXiv:1303.3589];
G. K. Leontaris and Q. Shafi, Phys. Rev. **D96**, 066023 (2017) [arXiv:1706.08372].

54. E. Witten, Nucl. Phys. **B471**, 135 (1996) [hep-th/9602070].
55. A. Hebecker and M. Trapletti, Nucl. Phys. **B713**, 173 (2005) [hep-th/0411131].
56. H. Georgi, H. R. Quinn and S. Weinberg, Phys. Rev. Lett. **33**, 451 (1974);
See also the definition of effective field theories by S. Weinberg, Phys. Lett. **B91**, 51 (1980).
57. C. Patrignani *et al.* [Particle Data Group], Chin. Phys. C **40**, 100001 (2016).
58. M.L. Alciati *et al.*, JHEP **0503**, 054 (2005) [hep-ph/0501086].
59. 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 *et al.*, Birkhäuser, Boston (1981).
60. W.J. Marciano, p. 190, *Eighth Workshop on Grand Unification*, ed. K. Wali, World Scientific Publishing Co., Singapore (1987).
61. M.S. Carena *et al.*, in Ref. [4].
62. B.C. Allanach, Comp. Phys. Comm. **143**, 305 (2002) [hep-ph/0104145].
63. A. Djouadi, J. L. Kneur and G. Moultaka, Comp. Phys. Comm. **176**, 426 (2007) [hep-ph/0211331].
64. W. Porod and F. Staub, Comp. Phys. Comm. **183**, 2458 (2012) [arXiv:1104.1573].
65. W. Siegel, Phys. Lett. **B94**, 37 (1980).
66. I. Antoniadis, C. Kounnas, and R. Lacaze, Nucl. Phys. **B221**, 377 (1983).
67. M.A. Shifman, Int. J. Mod. Phys. **A11**, 5761 (1996) [arXiv:hep-ph/9606281];
N. Arkani-Hamed and H. Murayama, JHEP **0006**, 030 (2000) [arXiv:hep-th/9707133];
I. Jack, D.R.T. Jones and A. Pickering, Phys. Lett. **B435**, 61 (1998) [arXiv:hep-ph/9805482].
68. M.B. Einhorn and D.R.T. Jones in Ref. [12];
I. Antoniadis, C. Kounnas, and K. Tamvakis, Phys. Lett. **B119**, 377 (1982).
69. J. Hisano, H. Murayama and T. Yanagida, Phys. Rev. Lett. **69**, 1014 (1992) and Nucl. Phys. **B402**, 46 (1993) [hep-ph/9207279].
70. H. Murayama and A. Pierce, Phys. Rev. **D65**, 055009 (2002) [hep-ph/0108104].
71. G. Anderson *et al.*, eConf **C960625**, SUP107 (1996) [hep-ph/9609457].
72. S. Raby, M. Ratz and K. Schmidt-Hoberg, Phys. Lett. **B687**, 342 (2010) [arXiv:0911.4249].
73. J. Hisano, T. Kuwahara and N. Nagata, Phys. Lett. **B723**, 324 (2013) [arXiv:1304.0343].
74. N. Arkani-Hamed, A. Delgado and G.F. Giudice, Nucl. Phys. **B741**, 108 (2006) [hep-ph/0601041].
75. M. Ibe, T. Moroi and T.T. Yanagida, Phys. Lett. **B644**, 355 (2007) [hep-ph/0610277].
76. G.F. Giudice *et al.*, JHEP **9812**, 027 (1998) [hep-ph/9810442];
L. Randall and R. Sundrum, Nucl. Phys. **B557**, 79 (1999) [hep-th/9810155].

77. R.N. Mohapatra and M.K. Parida, Phys. Rev. **D47**, 264 (1993) [hep-ph/9204234].
78. D.G. Lee *et al.*, Phys. Rev. **D51**, 229 (1995) [hep-ph/9404238].
79. L.E. Ibanez *et al.*, JHEP **1207**, 195 (2012) [arXiv:1206.2655].
80. K.R. Dienes, E. Dudas and T. Gherghetta, Phys. Rev. Lett. **91**, 061601 (2003) [hep-th/0210294].
81. K. Agashe, R. Contino and R. Sundrum, Phys. Rev. Lett. **95**, 171804 (2005) [hep-ph/0502222].
82. S. Weinberg, Phys. Rev. Lett. **43**, 1566 (1979);
F. Wilczek and A. Zee, Phys. Rev. Lett. **43**, 1571 (1979).
83. R. Barbier *et al.*, Phys. Rev. **420**, 1 (2005) [hep-ph/0406039].
84. G. Farrar and P. Fayet, Phys. Lett. **B76**, 575 (1978).
85. S. Dimopoulos, S. Raby and F. Wilczek, Phys. Lett. **B112**, 133 (1982);
J. Ellis, D.V. Nanopoulos and S. Rudaz, Nucl. Phys. **B202**, 43 (1982).
86. For a recent discussion, see C.S. Aulakh *et al.*, Nucl. Phys. **B597**, 89 (2001) [hep-ph/0004031].
87. K. Abe *et al.* [Super-Kamiokande Collaboration], Phys. Rev. **D95**, 012004 (2017) [arXiv:1610.03597];
S. Mine for the Super Kamiokande Collaboration, J. Phys. Conf. Ser. **718**, 062044 (2016).
88. Y. Aoki, T. Izubuchi, E. Shintani and A. Soni, Phys. Rev. **D96**, 014506 (2017) [arXiv:1705.01338 [hep-lat]].
89. For recent analysis, see J. Hisano, T. Kuwahara and Y. Omura, Nucl. Phys. **B898**, 1 (2015) [arXiv:1503.08561].
90. T. Goto and T. Nihei, Phys. Rev. **D59**, 115009 (1999) [hep-ph/9808255].
91. J.L. Chkareuli and I.G. Gogoladze, Phys. Rev. **D58**, 055011 (1998) [hep-ph/9803335].
92. G.F. Giudice and A. Romanino, Nucl. Phys. **B699**, 65 (2004) [hep-ph/0406088], [Erratum: *ibid.*, Nucl. Phys. **B706**, 65 (2005)].
93. J. Hisano *et al.*, JHEP **1307**, 038 (2013) [arXiv:1304.3651].
94. B. Bajc, P. Fileviez Perez and G. Senjanovic, Phys. Rev. **D66**, 075005 (2002) [hep-ph/0204311].
95. K.S. Choi, Phys. Lett. **B668**, 392 (2008) [arXiv:0807.2766].
96. F. An *et al.* [JUNO Collab.], J. Phys. G **43**, 030401 (2016) [arXiv:1507.05613].
97. Hyper-Kamiokande Collab., 'Hyper-Kamiokande Design Report', KEK-PREPRINT-2016-21, ICRR-REPORT-701-2016-1 (2016).
98. R. Acciarri *et al.* [DUNE Collaboration], 'The Physics Program for DUNE at LBNF', arXiv:1512.06148.
99. R. Kallosh *et al.*, Phys. Rev. **D52**, 912 (1995) [hep-th/9502069].
100. L.E. Ibanez and G.G. Ross, Nucl. Phys. **B368**, 3 (1992);
M. Berasaluce-Gonzalez *et al.*, JHEP **1112**, 113 (2011) [arXiv:1106.4169 [hep-th]].
101. L.J. Hall and Y. Nomura, Phys. Rev. **D66**, 075004 (2002) [hep-ph/0205067];
H.D. Kim, S. Raby and L. Schradin, JHEP **0505**, 036 (2005) [hep-ph/0411328].
102. M.S. Chanowitz, J.R. Ellis and M.K. Gaillard, Nucl. Phys. **B128**, 506 (1977);
A.J. Buras *et al.*, Nucl. Phys. **B135**, 66 (1978).

103. K. Inoue *et al.*, Prog. Theor. Phys. **67**, 1889 (1982);
L.E. Ibanez and C. Lopez, Phys. Lett. **B126**, 54 (1983) and Nucl. Phys. **B233**, 511 (1984).
104. H. Georgi and D.V. Nanopoulos, Nucl. Phys. **B159**, 16 (1979);
J. Harvey, P. Ramond and D.B. Reiss, Phys. Lett. **B92**, 309 (1980) and Nucl. Phys. **B199**, 223 (1982).
105. 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) [hep-ph/9209250];
B. Ananthanarayan, G. Lazarides and Q. Shafi, Phys. Lett. **B300**, 245 (1993);
G. Anderson *et al.*, Phys. Rev. **D49**, 3660 (1994) [hep-ph/9308333];
B. Ananthanarayan, Q. Shafi and X.M. Wang, Phys. Rev. **D50**, 5980 (1994) [hep-ph/9311225].
106. L.J. Hall, R. Rattazzi and U. Sarid, Phys. Rev. **D50**, 7048 (1994) [hep-ph/9306309, hep-ph/9306309];
M. Carena *et al.*, Nucl. Phys. **B426**, 269 (1994) [hep-ph/9402253].
107. A. Anandakrishnan, B.C. Bryant and S. Raby, JHEP **1505**, 088 (2015) [arXiv:1411.7035]..
108. M. Lanzagorta and G.G. Ross, Phys. Lett. **B364**, 163 (1995) [hep-ph/9507366].
109. K. Tobe and J.D. Wells, Nucl. Phys. **B663**, 123 (2003) [hep-ph/0301015].
110. T. Blazek, R. Dermisek and S. Raby, Phys. Rev. Lett. **88**, 111804 (2002) and Phys. Rev. **D65**, 115004 (2002) [hep-ph/0107097];
D. Auto, *et al.*, JHEP **0306**, 023 (2003) [hep-ph/0302155];
R. Dermisek *et al.*, JHEP **0304**, 037 (2003) and JHEP **0509**, 029 (2005) [hep-ph/0304101].
111. A. Anandakrishnan, S. Raby and A. Wingerter, Phys. Rev. **D87**, 055005 (2013) [arXiv:1212.0542];
M. Adeel Ajaib *et al.*, JHEP **1307**, 139 (2013) [arXiv:1303.6964];
Z. Poh and S. Raby, Phys. Rev. **D92**, 015017 (2015) [arXiv:1505.00264];
M. Badziak, M. Olechowski and S. Pokorski, JHEP **1310**, 088 (2013) [arXiv:1307.7999].
112. H. Georgi and C. Jarlskog, Phys. Lett. **B86**, 297 (1979).
113. S. Antusch and M. Spinrath, Phys. Rev. **D79**, 095004 (2009) [arXiv:0902.4644].
114. G. Lazarides, Q. Shafi and C. Wetterich, Nucl. Phys. **B181**, 287 (1981);
T.E. Clark, T.K. Kuo and N. Nakagawa, Phys. Lett. **B115**, 26 (1982);
K.S. Babu and R.N. Mohapatra, Phys. Rev. Lett. **70**, 2845 (1993) [hep-ph/9209215].
115. R. Barbieri and D.V. Nanopoulos, Phys. Lett. **B91**, 369 (1980).
116. R. Barbieri and A. Strumia, Nucl. Phys. **B493**, 3 (1997) [hep-ph/9704402];
T. Blazek, S. Raby and K. Tobe, Phys. Rev. **D60**, 113001 (1999) and Phys. Rev.

D62, 055001 (2000) [hep-ph/9903340];
Q. Shafi and Z. Tavartkiladze, Phys. Lett. **B487**, 145 (2000) [hep-ph/0002150];
C.H. Albright and S.M. Barr, Phys. Rev. Lett. **85**, 244 (2000) [hep-ph/0002155];
Z. Berezhiani and A. Rossi, Nucl. Phys. **B594**, 113 (2001) [hep-ph/0003084];
C.H. Albright and S.M. Barr, Phys. Rev. **D64**, 073010 (2001) [hep-ph/0104294];
M.-C. Chen, K.T. Mahanthappa, Int. J. Mod. Phys. **A18**, 5819 (2003) [hep-ph/0305088];
R. Dermisek and S. Raby, Phys. Lett. **B622**, 327 (2005) [hep-ph/0507045].

117. H. Georgi, Nucl. Phys. **B156**, 126 (1979);
P.H. Frampton, Phys. Lett. **B88**, 299 (1979);
P. Frampton and S. Nandi, Phys. Rev. Lett. **43**, 1460 (1979);
J.E. Kim, Phys. Rev. Lett. **45**, 1916 (1980), Phys. Rev. **D23**, 2706 (1981) and Phys. Rev. **D26**, 674 (1982);
R. Barbieri and D.V. Nanopoulos, Phys. Lett. **B91**, 369 (1980);
Y. Fujimoto, Phys. Rev. **D26**, 3183 (1982).
118. P. Langacker, Phys. Reports **72**, 185 (1981).
119. H. Georgi, Conf. Proc. C **820726**, 705 (1982).
120. S.M. Barr, Phys. Rev. **D78**, 055008 (2008) [arXiv:0805.4808];
Y. Goto, Y. Kawamura and T. Miura, Phys. Rev. **D88**, 055016 (2013) [arXiv:1307.2631];
J.E. Kim, JHEP **1506**, 114 (2015) [arXiv:1503.03104];
C.H. Albright, R.P. Feger and T.W. Kephart, arXiv:1601.07523;
M. Reig, J. W. F. Valle, C. A. Vaquera-Araujo and F. Wilczek, arXiv:1706.03116.
121. L.J. Hall, V.A. Kostelecky and S. Raby, Nucl. Phys. **B267**, 415 (1986).
122. R. Barbieri and L.J. Hall, Phys. Lett. **B338**, 212 (1994) [hep-ph/9408406];
R. Barbieri, L.J. Hall and A. Strumia, Nucl. Phys. **B445**, 219 (1995) [hep-ph/9501334].
123. F. Borzumati and A. Masiero, Phys. Rev. Lett. **57**, 961 (1986);
J. Hisano *et al.*, Phys. Lett. **B357**, 579 (1995) [hep-ph/9501407];
J. Hisano *et al.*, Phys. Rev. **D53**, 2442 (1996) [hep-ph/9501407];
J. Hisano and D. Nomura, Phys. Rev. **D59**, 116005 (1999) [hep-ph/9810479].
124. T. Moroi, JHEP **0003**, 019 (2000) [hep-ph/0002208];
D. Chang, A. Masiero and H. Murayama, Phys. Rev. **D67**, 075013 (2003) [hep-ph/0205111].
125. S. Dimopoulos and L.J. Hall, Phys. Lett. **B344**, 185 (1995) [hep-ph/9411273].
126. J. Hisano *et al.*, Phys. Lett. **B604**, 216 (2004) [hep-ph/0407169].
127. G.L. Fogli *et al.*, Phys. Rev. **D84**, 053007 (2011) [arXiv:1106.6028].
128. T. Schwetz, M. Tortola, and J.W.F. Valle, New J. Phys. **13**, 109401 (2011) [arXiv:1108.1376].
129. C.H. Albright and M.-C. Chen, Phys. Rev. **D74**, 113006 (2006) [hep-ph/0608137].
130. M. Magg and C. Wetterich, Phys. Lett. **B94**, 61 (1980);
J. Schechter and J.W.F. Valle, Phys. Rev. **D22**, 2227 (1980);
G. Lazarides, Q. Shafi and C. Wetterich, Nucl. Phys. **B181**, 287 (1981);

R.N. Mohapatra and G. Senjanovic, Phys. Rev. **D23**, 165 (1981);
G.B. Gelmini and M. Roncadelli, Phys. Lett. **B99**, 411 (1981).

131. R. Foot, *et al.*, Z. Phys. **C44**, 441 (1989).
132. G. 't Hooft, Nucl. Phys. **B79**, 276 (1974) A.M. Polyakov, Pis'ma Zh. Eksp. Teor. Fiz. **20**, 430 (1974) [Sov. Phys. 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, Rept. on Prog. in Phys. **41**, 1357 (1978).
133. M. Ambrosio *et al.* [MACRO Collaboration], Eur. Phys. J. C **25** (2002) 511 [hep-ex/0207020];
S. Balestra *et al.* 'Magnetic Monopole Bibliography-II,' arXiv:1105.5587;
L. Patrizii and M. Spurio, Ann. Rev. Nucl. Part. Sci. **65** (2015) 279 [arXiv:1510.07125].
134. For a review, see A.D. Linde, *Particle Physics and Inflationary Cosmology*, Harwood Academic, Switzerland (1990).
135. G.R. Dvali, H. Liu and T. Vachaspati, Phys. Rev. Lett. **80**, 2281 (1998) [hep-ph/9710301].
136. P. Langacker and S.Y. Pi, Phys. Rev. Lett. **45**, 1 (1980).
137. G.R. Dvali, A. Melfo and G. Senjanovic, Phys. Rev. Lett. **75**, 4559 (1995) [hep-ph/9507230].
138. V. Rubakov, Nucl. Phys. **B203**, 311 (1982) and 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).
139. K. Freese, M. S. Turner and D. N. Schramm, Phys. Rev. Lett. **51**, 1625 (1983).
140. P.A.R. Ade *et al.* [BICEP2 and Planck Collabs.], Phys. Rev. Lett. **114**, 101301 (2015) [arXiv:1502.00612];
P.A.R. Ade *et al.* [Planck Collab.], Astron. Astrophys. **594**, A20 (2016) [arXiv:1502.02114].
141. N.G. Deshpande, OITS-107;
C.Q. Geng and R.E. Marshak, Phys. Rev. **D39**, 693 (1989);
A. Font, L.E. Ibanez and F. Quevedo, Phys. Lett. **B228**, 79 (1989);
K.S. Babu and R.N. Mohapatra, Phys. Rev. Lett. **63**, 938 (1989).
142. R. Foot *et al.*, Phys. Rev. **D39**, 3411 (1989).
143. M. Nowakowski and A. Pilaftsis, Phys. Rev. **D48**, 259 (1993) [hep-ph/9304312].
144. T.P. T. Dijkstra, L.R. Huiszoon and A.N. Schellekens, Phys. Lett. **B609**, 408 (2005) [hep-th/0403196];
F. Gmeiner *et al.*, JHEP **0601**, 004 (2006) [hep-th/0510170];
B. Gato-Rivera and A.N. Schellekens, Nucl. Phys. **B883**, 529 (2014) [arXiv:1401.1782].
145. A.Y. Ignatiev *et al.*, Phys. Lett. **B76**, 436 (1978);
M. Yoshimura, Phys. Rev. Lett. **41**, 281 (1978) [Phys. Rev. Lett. **42**, 746 (1979)].

146. D. Toussaint *et al.*, Phys. Rev. **D19**, 1036 (1979);
S. Weinberg, Phys. Rev. Lett. **42**, 850 (1979);
M. Yoshimura, Phys. Lett. **B88**, 294 (1979);
S.M. Barr, G. Segre and H.A. Weldon, Phys. Rev. **D20**, 2494 (1979);
D.V. Nanopoulos and S. Weinberg, Phys. Rev. **D20**, 2484 (1979);
A. Yildiz and P.H. Cox, Phys. Rev. **D21**, 906 (1980).
147. M. Fukugita and T. Yanagida, Phys. Rev. Lett. **89**, 131602 (2002) [hep-ph/0203194].
148. M. Fukugita and T. Yanagida, Phys. Lett. **B174**, 45 (1986).
149. S. Davidson and A. Ibarra, Phys. Lett. **B535**, 25 (2002) [hep-ph/0202239];
K. Hamaguchi, H. Murayama and T. Yanagida, Phys. Rev. **D65**, 043512 (2002) [hep-ph/0109030].
150. M. Kawasaki *et al.*, Phys. Rev. **D78**, 065011 (2008) [arXiv:0804.3745].
151. W. Buchmuller, R.D. Peccei and T. Yanagida, Ann. Rev. Nucl. and Part. Sci. **55**, 311 (2005) [hep-ph/0502169];
C.S. Fong, E. Nardi and A. Riotto, Adv. High Energy Phys. **2012**, 158303 (2012) [arXiv:1301.3062].
152. G.G. Ross, "Grand Unified Theories", Benjamin/Cummings, 1984.;
K.R. Dienes, Phys. Reports **287**, 447 (1997) [hep-th/9602045];
P. Nath and P. Fileviez Perez, Phys. Reports **441**, 191 (2007) [hep-ph/0601023];
S. Raby, "Supersymmetric Grand Unified Theories" Lect. Notes Phys. **939** 1 (2017).
153. A. Ringwald, L.J. Rosenberg and G. Rybka, Review of 'Axions and other Very Light Bosons' in Ref. [57].
154. P. Svrcek and E. Witten, JHEP **0606** 051 (2006) [hep-th/0605206].
155. M. Dine, W. Fischler and M. Srednicki, Phys. Lett. **104B** 199 (1981);
A. R. Zhitnitsky, Sov. J. Nucl. Phys. **31** 260 (1980).