Unification: Coupling Constants

Frank Wilczek

April 9, 2014

Abstract

This is the second of two notes on unification of fundamental interactions. It describes the quantitative unification of couplings.

In the preceding note we saw how the messy multiplet structure of quarks and leptons in the standard model, including their hypercharges, comes to look much nicer when viewed within the context of $SU(5)$ or especially $SO(10)$ unification.

Those higher symmetries realize the separate $SU(3) \times SU(2) \times U(1)$ gauge symmetries of the standard model as different aspects of single, more encompassing symmetries. The glory of gauge symmetry, however, is that it controls not only bookkeeping, but also dynamics. For a simple (in the technical sense) gauge group such as $SU(5)$ or $SO(10)$, symmetry dictates all the couplings of the gauge bosons, up to a single overall coupling constant. So unification predicts relationships among the strong, weak, and hypercharge couplings. Basically – up to the group-theoretic task of normalization – it predicts that the three couplings for $SU(3) \times SU(2) \times U(1)$ must be equal.

As observed, of course, they are not. But the two great dynamical lessons of the standard model – symmetry breaking through field condensation (Higgs mechanism) and running couplings (asymptotic freedom) – suggest a way out. We can imagine that the symmetry breaking $G \rightarrow SU(3) \times SU(2) \times U(1)$ occurs through a big condensation, at a high mass scale. In the symmetric theory, appropriate to larger mass scales, there was only one unified coupling, so the couplings of $SU(3) \times SU(2) \times U(1)$ start out, at that scale, equal. But we make our observations at a much lower mass scale. To get to the unified coupling, we must evolve the observed couplings up to high energy, taking into account vacuum polarization. Note that throughout this evolution the unified symmetry is violated, so the three $SU(3) \times SU(2) \times U(1)$ couplings evolve differently.

Before entering the details, let's take a soft-focus view of what we can expect from this sort of calculation. Our input will be the observed couplings, plus some hypothesis $\mathcal{H}$ about the spectrum of virtual particles we need to include in the vacuum polarization. Our output should be the unified coupling strength, and the scale of unification. For any given $\mathcal{H}$, we have three inputs and two outputs, so there will be a consistency condition. If it works, we will have reduced the number of free parameters in the core of the standard model by one, from three to two.

There are additional *physical* consistency conditions, concerning the value of the unification scale, which are quite significant. Also, in case of success, we will need to discuss the plausibility of our hypothesis $\mathcal{H}$. We'll return to these important points later, after we've done the central calculation.

1 Normalization of Hypercharge

The value of nonabelian couplings can be fixed absolutely, because the generators obey nonlinear commutation relations. It is common practice, for $SU(n)$ or $SO(n)$ groups, to choose the coupling constant to multiply, in the fundamental representation, generators the trace of whose square $\frac{1}{2}$. Thus for $SU(2)$ we have the covariant derivative

$$\nabla_\mu \ = \ \partial_\mu + ig_2 \frac{\sigma_a}{2} A^a_\mu \qquad (1)$$

for isospinors, and so forth. (Of course, we understand here that the A^a appear in a Maxwell-like (Yang-Mills) action, in such a way that the plane wave "generalized photons" are canonically normalized.)

This fixes the normalizations for g_3, g_2, the couplings associated with $SU(3)$ and $SU(2)$ respectively, and also for g_5 in $SU(5)$[1]. The normalization of the hypercharge generator in $SU(5)$ is also thereby fixed, but its numerical relationship to the conventional normalization of the hypercharge of $U(1)_Y$, as it appears in electroweak phenomenology, requires discussion.

We have seen that in $SU(5)$ the fundamental representation (actually its conjugate, but that makes no difference here) is implemented on $(\bar{d}, \bar{d}, \bar{d}, L)$. The trace of the square of the hypercharge generator, times the square of the coupling constant, acting on this is therefore $g_5^2/2$. On the other hand, if we evaluate the same thing in the electroweak notation, we get

$$(g')^2(3 \times (\frac{1}{3})^2 \ + \ 2 \times (\frac{1}{2})^2) \ = \ (g')^2 \ \times \ \frac{5}{6} \qquad (2)$$

[1] There is no need, here, to discuss $SO(10)$ separately.

Equating these two evaluations of the same thing, we have

$$
\begin{aligned}
g_5^2 \times \frac{1}{2} &= (g')^2 \times \frac{5}{6} \\
g_5^2 \times \frac{3}{5} &= (g')^2
\end{aligned}
\tag{3}
$$

It will be convenient, for later purposes, to express this as the definition

$$
g_1^2 \equiv \frac{5}{3}(g')^2 \tag{4}
$$

– for it is g_1 that, in the unified theory, should evolve to become equal to g_2 and g_3 (and g_5).

Before any running of the couplings, we get the two "predictions"

$$
\begin{aligned}
g_3^2 &= g_2^2 \ (= g_5^2 \ ) \\
\sin^2 \theta_W &\equiv \frac{(g')^2}{g_2^2 + (g')^2} = \frac{\frac{3}{5}}{1 + \frac{3}{5}} = \frac{3}{8}
\end{aligned}
\tag{5}
$$

They're way off!

2 Structure of Couplings Renormalization

Each of the couplings runs logarithmically, due to vacuum polarization. As you see from Figure 1, the logarithmically divergent (before regularization) terms are proportional, in lowest order, to the cube of the coupling[2]. So we get equations for the running couplings of the form

$$
\frac{d\bar{g}_j(Q)}{d \ln Q} \approx b_j \bar{g}_j^3 \tag{6}
$$

where b is a number that depends on the spectrum of virtual particles that contribute.

To expose the logic of coupling unification, it is helpful to re-write Eqn. (6) as

$$
\frac{d\, 1/\bar{g}_j^2}{d \ln Q} = -\frac{2}{\bar{g}_j^3} \frac{d\bar{g}_j(Q)}{d \ln Q} = -2b_j \tag{7}
$$

[2]This emerges most clearly if we consider the gluon self-coupling. When we consider couplings of gluons to fermions, there appears to be cross-talk between the different interactions, as you see in Figure 2. These cross terms turn out not to contribute, due to Ward's identity, which is a diagrammatic manifestation of gauge symmetry. They'd better not, because the renormalized non-abelian coupling should be universal – the same for gluons, fermions, scalars, ghosts, ...)

with the solution

$$\frac{1}{\bar{g}_j(Q)^2} \;=\; -\,2b_j\,\ln\frac{Q}{Q_0} + \frac{1}{\bar{g}_j(Q_0)^2} \tag{8}$$

Here we take Q_0 to be an accessible laboratory scale, where we do empirical measurements of the $\bar{g}_j(Q)^2$.

Now the unification condition is that for some Q the effective couplings $\bar{g}_3^2(Q), \bar{g}_2^2(Q), \bar{g}_2^2(Q)$ become equal to a common value, call it $g_5(Q)$. By subtracting the solutions Eqn. (8) for $j = 2, 3$ we derive an equation determining the unification scale:

$$2\,(b_3 - b_2)\ln\frac{Q}{Q_0} \;=\; \frac{1}{\bar{g}_3(Q_0)^2} - \frac{1}{\bar{g}_2(Q_0)^2} \tag{9}$$

Since we must derive the same scale from other pairs of couplings, we have the consistency condition

$$\frac{b_3 - b_2}{b_2 - b_1} \;=\; \frac{\frac{1}{\bar{g}_3(Q_0)^2} - \frac{1}{\bar{g}_2(Q_0)^2}}{\frac{1}{\bar{g}_2(Q_0)^2} - \frac{1}{\bar{g}_1(Q_0)^2}} \tag{10}$$

This is the anticipated prediction, supplying a numerical relation among the observed couplings.

Of course, once we have the unification scale, we can go back to determine the value of the unified coupling using Eqn. (8).

3 Numerics of Couplings Renormalization

3.1 Renormalization Group Coefficients

The renormalization group coefficients b_j can be calculated perturbatively, for any combination of spin 0, $\frac{1}{2}$, and 1 (gauge) fields. The result is

$$16\pi^2\, b \;=\; -\,\frac{11}{3}C_A + \frac{4}{3}T(R_{\frac{1}{2}}) + \frac{2}{3}T(R_0) \tag{11}$$

Here C_A is the value of the Casimir operator for the adjoint representation of the gauge group in question; explicitly, we have

$$C_A(SU(n)) \;=\; n \tag{12}$$

$T(R_{\frac{1}{2}})$ is the trace of the square of a normalized generator acting on the spin-$\frac{1}{2}$ fields in the theory, which may of course include several multiplets.

We will only need the basic (defining result) $T = \frac{1}{2}$ for the fundamental representation, and, when we come to consider supersymmetry, $T = C = n$ for the adjoint of $SU(n)$. The coefficient $\frac{4}{3}$ for fermions holds for a complex Dirac fermions. For Weyl fermions (i.e., fermion fields with definite helicity) we get half that, as we do for Majorana (real) fermions. $T(R_0)$, unsurprisingly, is the trace of the square of a normalized generator acting on the spin-0 fields in the theory, which may of course include several multiplets. The coefficient $\frac{1}{3}$ holds for complex scalars; for real scalars we would get half that (but that case will not arise below).

3.2 Minimal Extrapolation

Taking the particles of the standard model, but allowing (why not?) for n_f families and n_s Higgs doublets, we have

$$16\pi^2\, b_3 \;=\; -\,11 \,+\, \frac{4}{3}\, n_f \tag{13}$$

since there are two fundamentals of Dirac fermions per family. Similarly

$$16\pi^2\, b_2 \;=\; -\,\frac{22}{3} \,+\, \frac{4}{3}\, n_f \,+\, \frac{1}{3} n_s \tag{14}$$

(four fundamentals of Weyl fermions!) and finally

$$16\pi^2\, b_1 \;=\; \frac{4}{3}\, n_f \,+\, \frac{1}{5} n_s \tag{15}$$

The b_1 equation can be obtained painlessly by noting that the contribution of complete families must respect the $SU(5)$ symmetry – that remark governs the fermions directly, while the Higgs particle should be fleshed out with a color triplet that isn't there, so its contribution is reduced by the factor

$$\frac{2 \times (\frac{1}{2})^2}{2 \times (\frac{1}{2})^2 \,+\, 3 \times (\frac{1}{3})^2} \;=\; \frac{3}{5} \tag{16}$$

If we run the couplings using these coefficients, with $n_f = 3$, $n_s = 1$, we get the unsatisfactory result shown in the first panel of Figure 15.1 (taken from [1]).

3.3 Extrapolation With Supersymmetry

To implement low-energy supersymmetry, in a minimal fashion, we must expand the standard model in several ways:

1. We have spin-$\frac{1}{2}$ Majorana fermion partners of the gauge fields, in the adjoint representation. Grouping their contribution with the gluons, has the effect of changing the $-\frac{11}{3}$ in Eqn. (11) to -3.

2. For each chiral fermion in the theory, we get a complex scalar superpartner. Grouping their contribution with the fermions, this has the effect of changing the $\frac{4}{3}$ in Eqn. (11) to 2 – still understanding, of course, that we halve this for chiral fermions!

3. Conversely, for each ordinary Higgs doublet we need a chiral fermion with the same quantum numbers. This changes the $\frac{1}{3}$ in Eqn. (11) to 1

4. Supersymmetry requires, at a minimum, $n_s = 2$ Higgs doublets.

Putting all this together, we find that when the contribution of virtual supersymmetry is included, we have

$$\begin{aligned} 16\pi^2\, b_3 &= \; -9 + 2\, n_f \\ 16\pi^2\, b_2 &= \; -6 + 2\, n_f + \frac{1}{2} n_s \\ 16\pi^2\, b_1 &= \; 0 + 2\, n_f + \frac{3}{10} n_s \end{aligned} \tag{17}$$

If we run the couplings using these coefficients, with $n_f = 3$, $n_s = 2$, we get the very satisfactory result shown in the first panel of Figure 15.1 (taken from [1]). The unification scale is computed to be $\approx 2 \times 10^{16}$ GeV.

As a measure of the delicacy and resolving power of the calculation, let us note that taking $n_s = 4$ leads to a 15% error in the prediction of the Weinberg angle, if we use strong-weak unification to fix the scale.

3.4 Observations on the Generality

A very important general observation: To the order we have been working, *complete $SU(5)$ multiplets affect neither the predicted relation among observed couplings, nor the predicted scale of unification*! That striking conclusion follows because complete multiplets contribute equally to b_1, b_2, and b_3, and in Eqns.(10, 9) only differences among those coefficients occur. So our successful "minimal supersymmetry" hypothesis is not so special as might appear at first sight. For example, one need not postulate a complete desert (apart from supersymmetry!) in the mass spectrum between current observations and the unification scale; one can populate it with any number of singlets, or with a modest number of complete families – it is only broken families that are worrisome. We can also allow, within supersymmetry, the masses of squarks and sleptons – the partners of quarks and leptons – to float up to a high scale together, since they form complete $SU(5)$ multiplets. In principle, there could even be different large masses for the different families of superpartners. (The partners of gauge bosons, on the other hand, *do not* form a complete $SU(5)$ multiplet.) These sorts of possibilities have been explored in speculative phenomenology, first under the epithet "focus point" and more recently also as "split" and "mini-split" supersymmetry. Raising the squark and slepton masses is an attractive option phenomenologically, because it relieves difficulties with proton decay and flavor violation – processes that otherwise tend to be over-predicted in supersymmetric models. (On the other hand, we probably do not want to raise those masses > 10 TeV, as this leads to difficulties with another attractive consequence of unified theories, namely their quantitative explanation of the mass ratio m_b/m_τ [1].)

Effects of complete multiplets will show up in more accurate calculations, taken to higher order. They also affect the value of the unified coupling. By increasing the value of b_j in Eqn. (8), they make the unified coupling larger. We probably don't want to have too many of them, therefore.

4 Loose Ends

Above I have outlined the lowest-order calculation. There is a vast technical literature on corrections, both those due to additional couplings and those due to masses [1]. I think it is fair to say that the situation is generally satisfactory, although unfortunately there are many more potentially parameters than experimental constraints, when one attempts precision work.

The Higgs doublet of the standard model, or the two Higgs doublets of its supersymmetric extension, do not fill out unified multiplets. Indeed, there are powerful bounds on the mass of the possible color triplet partners, since they make it difficult to maintain baryon number conservation as a good approximation. There are several ideas to address this doublet-triplet splitting embarrassment, but no consensus on which (if any) is correct [1].

5 Prospect

5.1 Significance of the Scale

5.1.1 Relation to Planck Scale

The Planck energy

$$\mathcal{E}_{\text{Planck}} = \sqrt{\frac{\hbar c^5}{8\pi G_N}} \approx 2.4 \times 10^{18}\,\text{GeV} \tag{18}$$

is another famous energy scale that can be constructed from fundamental constants[3]. Here the construction is simple dimensional analysis, based on Newton's gravitational constant G_N together with $\hbar, c$. On the face of it, Planck units set the scale for effects of quantum gravity; thus when we consider basic (technically: hard) processes whose typical energies are of order E, we expect gravitational effects of order $(E/\mathcal{E}_{\text{Planck}})^2$.

Our scale $\mathcal{E}_{\text{unification}} \approx 2 \times 10^{16}$ GeV is significantly, but not grotesquely, smaller than the Planck scale. This means that at the unification scale the strength of gravity, heuristically and semi-quantitatively, is of order

$$(\mathcal{E}_{\text{unification}}/\mathcal{E}_{\text{Planck}})^2 \sim 10^{-4} \tag{19}$$

to be compared with the strength $g_5^2/4\pi \sim 10^{-2}$ for the other interactions. The relative smallness of gravity, thus estimated, which of course is more severe at lower energies, suggests that in our neglect of quantum gravity in the preceding calculations may be justified.

On the other hand, it seems to me remarkable that the comparison is so close. A classic challenge in fundamental physics is to understand the grotesque smallness of the observed force of gravity, compared to other interactions, as it operates between fundamental particles . Famously, the gravitational interaction is $\sim 10^{42}$ times smaller than any of the other forces. Again, however, proper comparison requires that we specify the energy scale at which the comparison is made. Since the strength of gravity, in general relativity, depends on energy directly, it appears hugely enhanced when observed with high-energy probes. At the scale of unification $\mathcal{E}_{\text{unified}} \sim 2 \times 10^{16}$GeV the discrepant factor 10^{42} is reduced to $\sim 10^4$, or even a bit less. While this does not meet the challenge fully, evidently it marks a big step in the right direction.

---

[3]We have quoted the so-called rationalized Planck scale, including the factor 8π that naturally appears with G_N in the Lagrangian of general relativity.

5.1.2 Neutrino Masses and Proton Decay

By expanding our theory, unification along the lines we have been discussing brings in additional interactions. Since the unified multiplets combine particles that normally (i.e., with the standard model itself!) don't transform into one another. The two classic predictions for "beyond the standard model" interactions are small neutrino masses and proton decay. The first has been been vindicated; the second not (yet?). In both cases, the large scale $\mathcal{E}_{\text{unification}}$ is crucial for explaining the smallness of the new effects. For an authoritative review of these and other aspects of unification, emphasizing the phenomenological issues, with many further references, see [1].

5.1.3 Inflation Scale?

Finally, it is timely to mention a possible connection to the recent (probable) observation of B modes in the cosmic microwave background radiation, indicative of primordial gravity waves. The amplitude of the gravity wave emission is set by the rate of inflationary expansion, which in turn is set by the vacuum energy density during inflation. This leads to

$$\mathcal{E}_{\text{inflation}} \; = \; 1.06 \times 10^{16}\text{GeV} \left(\frac{r}{0.01}\right)^{1/4} \; = \; V^{1/4} \tag{20}$$

where V is the vacuum energy density. The tentative observation $r \approx .2$ gives

$$\mathcal{E}_{\text{inflation}} \approx \mathcal{E}_{\text{unification}} \tag{21}$$

This seems to me to be an impressive numerological "coincidence". As such, it raises the challenge, to incorporate it as an organic feature of attractive models.

5.2 Conclusion

The unification of quark and lepton quantum numbers in $SU(5)$, and especially $SO(10)$, is unforced and strikingly beautiful. The unification of coupling strengths fails quantitatively if one makes a minimal extrapolation of the standard model, but under the hypothesis of low-energy supersymmetry its success is likewise unforced and strikingly beautiful. It brings a new scale into physics, which has several attractive features. The discovery (or not) of some superpartners at the LHC will bring this line of thought to fulfillment (or not).

References

[1] S. Raby http://pdg.lbl.gov/2013/reviews/rpp2013-rev-guts.pdf. The particle data group updates its review regularly, so the later parts of this link will eventually fail; but pdg.lbl.gov should remain a reliable, easy-to-use portal.

Figure 1

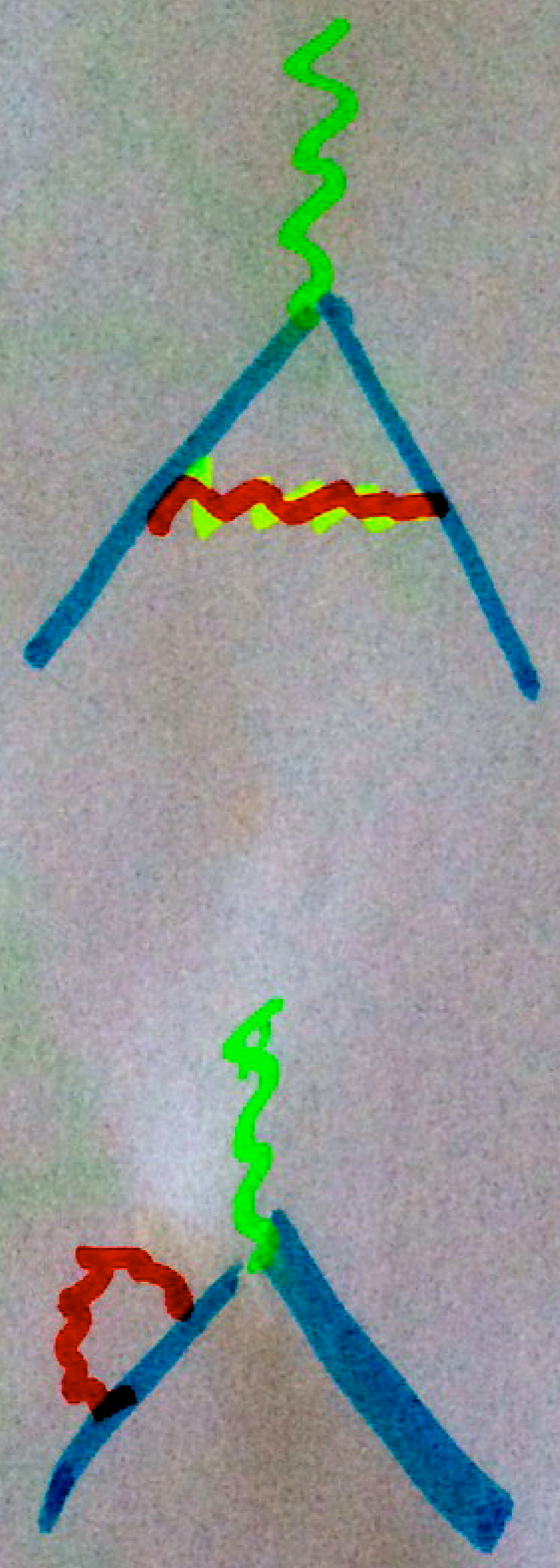

Figure 2

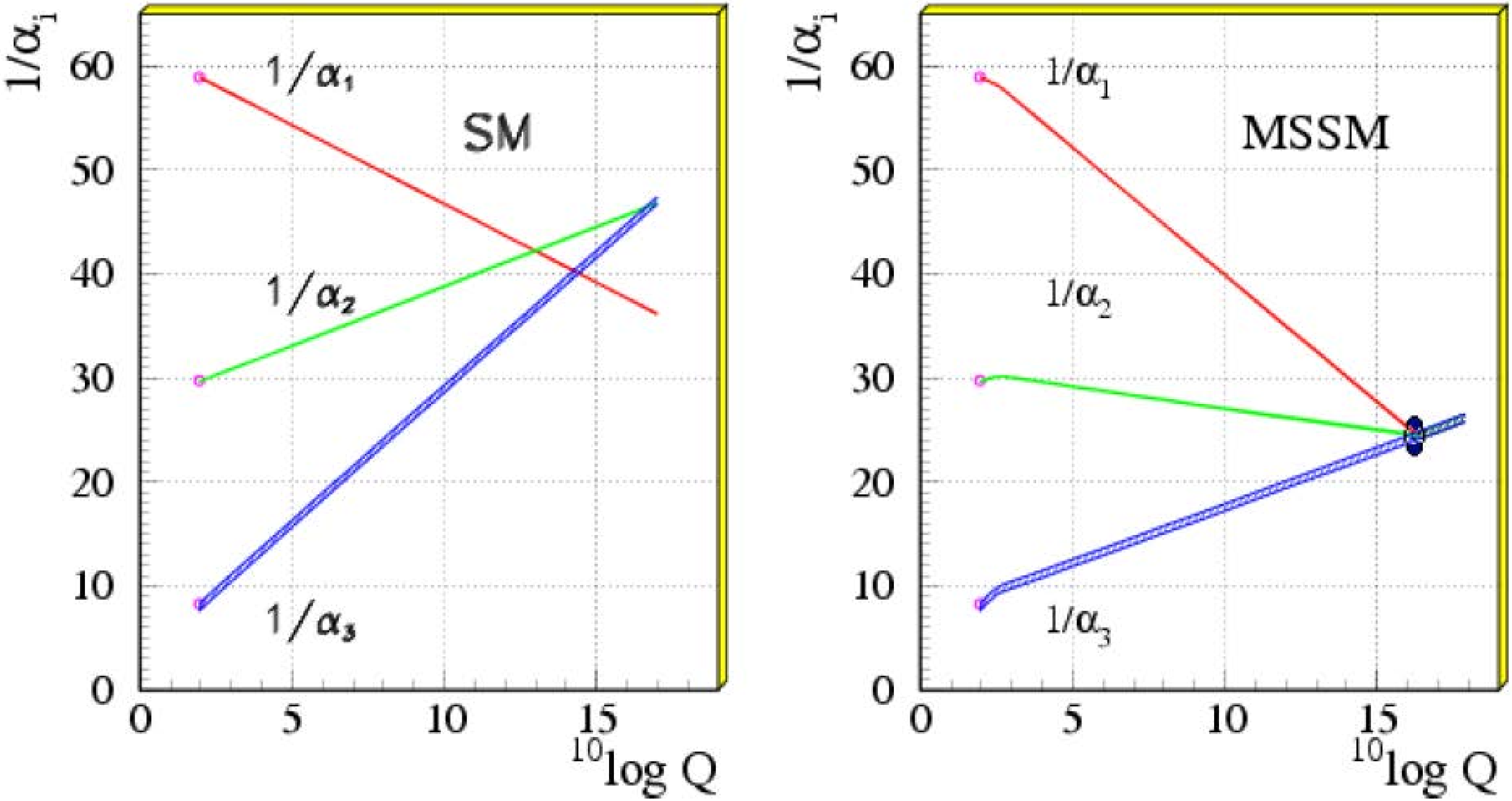

Figure 15.1: Gauge coupling unification in non-SUSY GUTs on the left vs. SUSY GUTs on the right using the LEP data as of 1991. Note, the difference in the running for SUSY is the inclusion of supersymmetric partners of standard model particles at scales of order a TeV (Fig. taken from Ref. 24). Given the present accurate measurements of the three low energy couplings, in particular $\alpha_s(M_Z)$, GUT scale threshold corrections are now needed to precisely fit the low energy data. The dark blob in the plot on the right represents these model dependent corrections.