Ref.TH.2981-CERN

GRAND UNIFIED THEORIES - A FRESH LOOK

D.V. Nanopoulos

CERN - Geneva

and

Department of Theoretical Physics *)
University of Athens
Athens 621, Greece

A B S T R A C T

The "basics" of Grand Unified Theories (GUTs) are discussed. Particular emphasis is given to a proposed novel approach to the fermion mass spectrum.

Invited talk given at the Symposium
"Perspectives in Modern Field Theories"
Stockholm, 23 - 26 September 1980

*) Address until January 1981.

1. QUANTUM FIELD THEORY

Let us accept that quantum field theory, an appropriate and successful mixture of quantum mechanics and special relativity, is the right framework to describe particles and their interactions. Since perturbation theory, at least for now, is the only way to make sense out of quantum field theory, let us concentrate on renormalizable field theories. Then, gauge theories seem to be the most outstanding candidates for providing the correct description of the physical laws. Gauge theories have such a rich structure that it is possible to describe and interrelate a whole lot of different phenomena, which seem at first sight quite unrelated.

Here I will try to present a somewhat unconventional approach to some exciting recent developments in the area of quantum field theory devoted to the unification of all elementary particle forces.

2. "THE FERMI-SYNDROME"

Let us suppose that we are living in a world where only strong (QCD) and electromagnetic (EM) interactions are present acting on the "standard" three generations of fermions

$$(u, d, e, \nu_e) ; (c, s, \mu, \nu_\mu) ; (t, b, \tau, \nu_\tau) \quad (1)$$

As usual, strong and electromagnetic interactions are described by a renormalizable gauge theory based on the unbroken groups $SU(3)_{\text{COLOUR}}$ and $U(1)_{\text{EM}}$, respectively. So the relevant group is $G_0 \equiv SU(3)_C \times U(1)_{\text{EM}}$. Then, one can prove¹⁾ in perturbation theory that, under such circumstances, there is a whole class of quantum numbers that are conserved, like the baryon number (B), lepton number (L), strangeness, charm, etc., corresponding to some unbroken global symmetries. It should be stressed right away that the restriction of renormalizability, i.e., allowing only couplings of dimension equal to or less than four, plays a rather essential role. Suppose that we find experimentally that the neutron is unstable or that strangeness is violated, what will we do then? Well, we have to put by hand terms in our Lagrangian that violate these quantum numbers, i.e., we are bound to introduce new couplings of dimension five or more between our "known" fields, i.e., non-renormalizable interactions.

For example, one may consider terms like (generically written)

$$\frac{g^2}{M_W^2} (\bar{s}u)(\bar{\mu}\nu_\mu) \quad (2)$$

where g is some gauge coupling constant $O(e)$ and M_W is some scale, determined experimentally, but definitely much bigger than the energy scale at our disposal.

Actually, we have just introduced a new kind of interactions, which are characterized for energies $E_{CM} \ll M_W$ by an effective coupling constant $(E_{CM}^2/M_W^2)g^2$ much smaller than g^2 and thus they are properly called "weak interactions".

For the bulk of the phenomena in our "low energy" world, weak interactions play essentially a very minor role, though Eq. (2) cannot be the whole story. As we increase our energy capacities and specifically as E_{CM} exceeds M_W , terms like the one of Eq. (2) cannot be neglected anymore $(E_{CM}^2/M_W^2)g^2 \sim 1$ which is unfortunate because they lead to violation of unitarity and other horrible and unpleasant things. We have to do a bit better. If gauge theories work for strong and electromagnetic interactions why not use them also for weak interactions? The basic form of weak interactions is then

$$g(\bar{\psi}_1 \gamma_\mu (1 - \gamma_5) \psi_2) W_\mu, \quad (3)$$

where Eq. (2) is now translated as an "effective" interaction emerging from (3) at energies much smaller than M_W , the mass of the gauge bosons (W) mediating weak interactions.

An important point to stress here is the fact that, while looking for a renormalizable theory of weak interactions, it transpires that the unification of weak and electromagnetic interactions is a necessity. Such an outcome is difficult to overestimate. We have quite strong experimental evidence²⁾ that the group that unites weak and electromagnetic interactions is $SU(2) \times U(1)$.

So we reach the point where strong, electromagnetic and weak interactions are described by a renormalizable gauge theory based on the group $G_1 \equiv SU(3) \times SU(2) \times U(1)$. The main difference between strong/electromagnetic and weak interactions is that $SU(2) \times U(1)$ is spontaneously broken $[SU(2) \times U(1) \rightarrow U(1)_{EM}$ at $M_W \approx 100 \text{ GeV}]$ while $SU(3)_C$ and $U(1)_{EM}$ remain unbroken (exact symmetries). Again, we can prove^{1),3)} that under such circumstances $[SU(3) \times SU(2) \times U(1)$ gauge theory, "standard" fermions, one Higgs doublet and only renormalizable couplings] the only unbroken global symmetries are those that correspond to B and L conservation. The fact that we restrict ourselves to renormalizable couplings again plays an essential role in the proof.

Suppose now that we find experimentally that protons decay, and/or that neutrinos have a Majorana mass ($m_{\nu\nu}$), what then? Well, we have now learned the trick. Mutatis mutandis, replace (2) by

$$\frac{g^2}{M_X^2} (uu)(\bar{e}^+ d) \quad (2')$$

and (3) by

$$g(\psi_1 \gamma_\mu (1 \pm \gamma_5) \psi_2) X_\mu \quad (3')$$

where Eq. (2') should now be considered as an effective interaction emerging from (3') at energies much smaller than M_X , the mass of the gauge bosons (X) mediating these "exotic" new interactions, sometimes called grand unified interactions. No surprises, dull repetition of history^{4),5)}: B or L number violation would signal the existence of new interactions beyond the $SU(3) \times SU(2) \times U(1)$ and by analogy with the previous case, i.e.,

strangeness violation $\rightarrow$ weak-electromagnetic unification

We would like to argue that

B and/or L number violation $\rightarrow$ strong-electromagnetic-weak unification,

in other words, there is some big group G such that $G \supset SU(3) \times SU(2) \times U(1)$, i.e., G , is the grand unification group. Lower bounds⁶⁾ on the proton lifetime ($\tau_p \geq 10^{30}$ years) indicate that $M_X \geq O(10^{14} \text{ GeV})$, indeed a superhigh value, at least twelve orders of magnitude bigger than M_W and at most five orders of magnitude smaller than the Planck mass, $M_{Pl} \approx 1.2 \cdot 10^{19} \text{ GeV}$. The effects of such a superlarge mass scale will be discussed later on.

The examples used up to now show us the general rules of discovering new interactions. Since Fermi taught us all, I call them the "Fermi-Silver rules":

- 1) Find some conserved quantum number under the "established" (known) interactions.
- 2) Find some "reason" as to why this "conserved" quantum number is "slightly" violated.
- 3) Find some experimentalists to look for such "tiny" violations. If they find it, we have made it.

Nowadays, there is a frenzied activity of looking for proton decay deep in the mines and justifying it by employing theoretical arguments from the very very

early hot Universe. Grand unified theories have indeed changed experimentalists into mine engineers and theorists into cosmologists. Our present approach to physics is certainly dominated by the "Fermi-Syndrome", the introduction of more and more new "mass scales" which signal the existence of more and more new interactions.

3. GRAND UNIFICATION

Grand unified theories (GUTs), theories that unify^{4),5)} weak, electromagnetic and strong interactions beyond their aesthetical appeal, answer a lot of questions and resolve many of the mysteries contained in the "standard" ($SU(3) \times SU(2) \times U(1)$) model. Furthermore, by making some spectacular and dramatic predictions like the instability of the proton, or the massiveness of the neutrino, it will, hopefully soon be possible, to test such grand ideas. Since many reviews⁷⁾ exist by now on the "classical" applications of grand unified theories (GUTs), let me briefly mention the highlights of these grand applications. In grand unified theories, there is a big "simple" group G that contains $G_1 (\equiv SU(3) \times SU(2) \times U(1))$, characterized by a unique gauge coupling constant g , and where quarks and leptons are sitting in the same representation(s) of the group G .

Then, because the group G is "simple", the following facts are immediate consequences.

- 1) The electric charge is quantized. No more mystery as to why $Q_{\text{electron}} + Q_{\text{proton}} = 0 (< 10^{-20})$. Incidentally, grand unification implies that α , the EM fine structure satisfies the inequality⁸⁾

$$\alpha(0) < \frac{1}{25} \quad (4)$$

if the proton decays ! Not a trivial bound.

- 2) At superhigh energies, there are "simple relations" (involving at most Clebsch-Gordan coefficients) between the different coupling constants (g_1, g_2, g_3) because they are all trivially related to g .

For example, at superhigh energies, we get⁵⁾

$$\sin^2 \theta_W \equiv \frac{e^2}{g_2^2} = \frac{3}{8} \quad (5)$$

Then at "lower" energies each gauge coupling constant is renormalized in a different way⁹⁾, and so eventually at our "low energy regime" we find the big differences in "strength" between strong, electromagnetic and weak interactions.

For example, at "low energies", we get^{9),10)}

$$\sin^2 \theta_W(M_W) = \frac{1}{6} + \frac{5}{9} \frac{\alpha(M_W)}{\alpha_3(M_W)} \approx 0.21 \quad (6)$$

to be compared with the "experimental" value of $\sin^2 \theta_W \approx 0.230 \pm 0.015$ - a rather successful prediction of GUTs. The fact that quarks and leptons are in the same multiplet(s) has some very important consequences.

1) There are relations between quark and lepton quantum numbers. For example, one finds that $Q_d = (1/3) Q_{e-}$, where the factor of three comes in because there are three colours. Also, if for example the u and d quarks are thrown in a left-handed doublet of $SU(2)_{\text{weak}}$ $((u,d)_L)$, then it necessarily follows that ν_e and e have also to be in a left-handed doublet $((\nu_e, e)_L)$, etc. An extensive analysis of these phenomena is given in Ref. 11). In other words, we get a satisfactory explanation of the startlingly similar behaviour of quarks and leptons under weak and electromagnetic interactions.

2) At superhigh energies there are "simple" relations (involving at most Clebsch-Gordan coefficients) between quark and lepton masses.

For example, a "natural" relation of the form^{5),10)}

$$m_b = m_\tau \quad (7)$$

at superhigh energies is often found in many grand unified models (GUMs). Again at "lower energies" quark masses and lepton masses are renormalized in a different way¹⁰⁾ (basically quarks feel the strong interactions, while the leptons are insensitive to them), with a result that, at "present energies"¹⁰⁾

$$\frac{m_b}{m_\tau} = \left[\frac{\alpha_3(2m_b)}{\alpha_3(M_X)} \right]^{\frac{12}{33-2f}} \quad (8)$$

where Eq. (7) has been used and f indicates the number of quark flavours. Using $f = 6$ and presently accepted values for the strong coupling constant α_3 ($\equiv \frac{g_3^2}{4\pi}$) we find¹⁰⁾

$$\frac{m_b}{m_\tau} \approx (2.7 - 3) \quad (9)$$

just what we see experimentally ! Another rather successful prediction of GUTs. A simple quark-lepton mass relation, like the one of Eq. (7), seems also to work (?) for the second generation ($m_s = m_\mu$), but it is definitely wrong for the first generation ($m_d \neq m_e$). Some modification is needed. We will come to this point later.

The fact that we used $f = 6$ in computing the renormalization is very crucial. If we start increasing the number of flavours, as it is apparent from (9), the m_b/m_τ ratio will increase unchecked¹²⁾. Thus we claim that in GUTs the number of flavours is more or less determined to be^{10),12)}

$$f = 6 \quad (10)$$

This is a rather drastic prediction and not difficult to test experimentally.

3) Since quarks and leptons coexist in the same multiplet(s), there are gauge interactions (mediated by the X bosons) that transform one to another and thus B and L numbers are violated. Protons are not forever. The predicted proton lifetime¹³⁾

$$\tau_P \approx 10^{31 \pm 2} \text{ years} \quad (11)$$

is not far above the experimental lower bound⁶⁾ $\tau_P > 10^{30}$ years. Thanks to the presently intense experimental search¹⁴⁾ for proton decay, hopefully we will know soon.

With the form of L violation generally found in GUTs, we find that neutrinos do indeed acquire a naturally small Majorana mass¹⁵⁾ $(m_{\nu\nu})$ of the type discussed earlier, where

$$m_\nu \approx \left(\frac{M_W}{M}\right)m \approx (10^{-5} \text{ eV} \rightarrow \text{few eV}) \quad (12)$$

and, as usual, m represents a "low energy" charged fermion mass (most likely the corresponding $q = 2/3$ quark mass), and M represents a "superheavy" mass scale. Again experiments will tell, if we do not already know¹⁶⁾.

4. CONSEQUENCES OF THE EXISTENCE OF SUPERHEAVY MASS SCALE $M(\sim 10^{15} \text{ GeV})$

It is important to notice that the proton quasistability, the fact that the strong interactions are strong ($\alpha_s \approx 1$) around 1 GeV, the observed value of $\sin^2\theta_W$ and the observed ratio $\frac{m_b}{m_\tau}$, all indicate independently and uniquely that the energy scale (M) where the grand unification occurs is extremely (exotically) large^{5),9),10),13)}

$$M \approx 10^{15} \text{ GeV} \quad (13)$$

Actually it is a triumph of grand unification that, given Eq. (13), all the above mentioned experimental facts, completely unrelated at first sight, are subject to a common explanation in the guts of which lies the mere fact of the existence

of a new superheavy mass scale, M . Here are some more implications of the existence of a new superheavy mass scale.

4.1 The Survival Hypothesis^{17),18)}

Up to now we have accepted the existence of at least two energy scales related with new phenomena according to the following line of symmetry breaking

$$G \xrightarrow{M \sim 10^{15} \text{ GeV}} G_1 (\equiv SU(3) \times SU(2) \times U(1)) \xrightarrow{M_W \sim 100 \text{ GeV}} G_0 (\equiv SU(3) \times U(1)) \quad (14)$$

It is only natural to assume that anybody who "feels" the mechanism that triggers the $G \rightarrow G_1$ breaking will also get a superheavy mass of order M . So we better be careful that the observed "low energy" particles are "insensitive" to the superheavy symmetry breaking. This is not difficult to arrange for the "low energy" gauge bosons. Choose particular directions of the spontaneous symmetry breaking and then gauge invariance guarantees the masslessness, at this stage of breaking, of the "observed" gauge bosons. There is also no difficulty in arranging things for the fermions. Put the "observed" fermions in such representations so that it is impossible to get invariant mass terms under G_1 . This is the so-called¹⁸⁾ "Survival Hypothesis": "Low mass fermions are those that cannot receive G_1 invariant masses".

In other words, the mere fact of the existence of "low mass" fermions puts constraints on their representation content. Actually, since nature chooses the exact (unbroken) groups of strong ($SU(3)$) and electromagnetic ($U(1)_{EM}$) interactions to be vectorlike (non-chiral), i.e., L and R fermionic components transform identically under these groups, our only hope left is that $SU(2)$ is a chiral group, i.e., the weak interactions violate P and C both in the charged and neutral currents. Indeed, this is the case experimentally. What a strange connection, the existence of "low mass" fermions implies parity violation in weak interactions. If Fermi only knew ! A detailed analysis of the implications of the survival hypothesis on the "low mass" fermion spectrum may be found in Ref. 11).

4.2 The Gauge Hierarchy Problem

We have seen that some gauge bosons remain massless because a certain gauge symmetry remains unbroken, while the existence of some unbroken chiral symmetries keeps some fermions massless. Then we say that the masslessness of these (gauge and fermion) particles is a natural effect as emerging from some symmetry. Unfortunately, there is no symmetry able to keep the Higgs bosons needed to trigger the $G_1 \rightarrow G_0$ spontaneous symmetry breaking, massless after the superheavy symmetry breaking. Scale invariance may come to mind but, as is well known

(remember the "dilatation anomaly")¹⁹⁾, it is explicitly broken in higher orders. Here lies the heart of the gauge hierarchy problem: how is it possible to understand naturally the huge gap between M_W and M_X

$$\frac{M_W}{M_X} \approx 10^{-12} \quad (15)$$

This is, for the time being, a big puzzle. Maybe it is related to our ignorance of how to handle quantum gravity, or maybe it is a rather technically complicated problem and it needs more hard thought and calculations before it is resolved. Whatever the answer to the gauge hierarchy problem may be, I believe that the spectacular successes of GUTs make us feel that we are on the right track. Just remember that all the successes of the "old-fashioned" Fermi-type (non-renormalizable) weak interactions remained intact after the invasion of the real thing, i.e., gauge theories.

4.3 Renormalizability

One may wonder why a renormalizable gauge theory, such as "standard" $SU(3) \times SU(2) \times U(1)$, gives such a successful description of the world, at least inside the presently explored energy range? That is to say, what is so deep about renormalizability? The argument that perturbation theory makes no sense for non-renormalizable interactions is not totally convincing, since one may imagine some other approach to field theory which bypasses perturbation theory. What then? The answer lies in the so-called "zero mass decoupling theorem"²⁰⁾. This theorem states that if a large mass (M in our case) occurs in field theory, and if a subset of fields or states remains massless as $M \rightarrow \infty$, then these states decouple from the rest in the limit, and their effective interactions at low energy are renormalizable (with additional non-renormalizable interactions inversely proportional to some power of M).

The renormalizability then of the "low energy" interactions of the "low mass" particles is due to the mere existence of a superheavy mass scale M . Many of the mysteries of the "low energy" world are suddenly dissolved. The set of "elementary fields" of low energies should not contain fields with spin higher than one, since their interaction will be necessarily non-renormalizable. Similarly, spin 1 fields must be²¹⁾ gauge fields, and spin 1/2 fields have to form an anomaly free set²²⁾. Unfortunately, no constraints are imposed on the scalar fields. Finally, one may replace M by M_{Planck} , the scale which sets in by gravity and ditto justifies the renormalizability of grand unified theories.

If the tacit assumption that nothing happens between a few hundred GeV and 10^{15} GeV (the "desert" or "grand plateau" hypothesis) is correct, as experiments may soon tell, then physics certainly takes a completely new turn. Building larger and larger accelerators may not provide the solution to our problems. Much harder thinking, presumably quite unconventional, experimentally and theoretically, will be needed if we would like to make real progress. It is not accidental that during the last few years quite an activity has developed in studying the interconnections between GUTs and cosmology. The very early and thus extremely hot Universe is a first-class "laboratory" to test a lot of "grand ideas". The GUTs-cosmology amalgamation²³⁾ has produced quite exciting and interesting results, so that I foresee more and more interaction between GUTs and cosmology if, after all, we would like to understand the finest details of the world around us. For a recent review on the GUTs cosmology front see Ref. 23).

Next I would like to discuss some new ideas concerning the fermion mass spectrum.

5. THE PROBLEM OF MASS

We have already seen that, in the framework of GUTs, we get a satisfactory explanation¹⁰⁾ of the quark lepton mass differences [see Eqs (8) and (9)] and we understand¹⁵⁾ in a natural way why the neutrino has a very tiny mass, if it is not massless [see Eq. (12)]. Our next task is then to try to understand the so-called generation-(mass)-gap. It is an experimental fact that the masses of the first generation (u, d, e, ν_e) are $O(\alpha^2 M_W)$, while for the second generation (c, s, μ, ν_μ) are $O(\alpha M_W)$ and finally some particles of the third one, (t, b, τ, ν_τ) is $O(M_W)$. Apparently, nature tries to tell us something. Let us try to listen! The first thought that immediately comes to mind is that the third generation gets a "tree-level" mass by directly "feeling" the mechanism that breaks $SU(2)_{\text{weak}}$, while the second generation gets a radiative "one-loop" mass, with the first generation getting a radiative "two-loop" mass, both generations being "insensitive" to the $SU(2)_{\text{weak}}$ symmetry breaking mechanism.

Here we have essentially two problems to solve. First, find the mechanism that allows the third generation to get a "tree-level" or "direct" mass, while at the same time prevents the other two generations from getting "direct" mass. In this case, the mass spectrum for the three generations should look like $(0, 0, M_W)$ or for short $(0, 0, 1)$ in units of M_W . I call it the $(0, 0, 1)$ problem.

Secondly, find the suitable higher order diagrams that create the appropriate masses for the first two generations. Here, the mass spectrum of the three generations should look like $(\alpha^2 M_W, \alpha M_W, M_W)$ or for short $(\alpha^2, \alpha, 1)$ in units of M_W . I call it the $(\alpha^2, \alpha, 1)$ problem. For the sake of clarity, let us discuss each problem separately.

5.1 The (0,0,1) Problem

Clearly, at the level of $SU(3) \times SU(2) \times U(1)$ or even at the level of conventional GUTs like $SU(5)$, $O(10)$, or E_6 , the appearance of more than one generation is redundant. Since all three generations have identical properties under such interactions, there is no hope to finding a solution to our problem inside the framework of conventional GUTs. Certainly, we should involve generation changing interactions in the game, and let us call G' the "simple" group that unifies all interactions (except gravity) including generation changing interactions. Then Eq. (14) generalizes to

$$G' \xrightarrow{M_{Pl} > M' > M} G \xrightarrow{M} G_1 \xrightarrow{M_W} G_0 \quad (16)$$

Now let us proceed by analogy. The basic reason for the presence of superheavy and low mass fermions in the conventional GUTs (G) is that these two classes of fermions transform differently under G ; that is the heart of the survival hypothesis. This piece of information is embodied in the structure of the Yukawa interactions invariant under G . For example, in the $SU(5)$ model⁵⁾ since $\bar{5} \times 10 (= \bar{5} + 45)$ and $10 \times 10 (= \bar{5} + \bar{45} + 50)$ do not contain the 24, the Higgs representation that causes the superheavy symmetry breaking, fermions belonging to the $\bar{5} + 10$ representation remain massless at the $G \rightarrow G_1$ breaking. $SU(5)$ invariance prevents the appearance of any $\bar{5} \times 10 \times 24$ or $10 \times 10 \times 24$ coupling. On the other hand, suppose that some fermions belong to the $5 + \bar{5}$ representations of $SU(5)$. Then clearly since $\bar{5} \times 5 (= 1 + 24)$ contain the 1 and 24, these fermions will become superheavy. In this case, $SU(5)$ invariance allows the appearance of a $\bar{5} \times 5 \times 24$ coupling. It should be stressed that the L and R components of the $(5 + \bar{5})$ superheavy fermions belong to equivalent representations of $SU(2)_{\text{weak}}$, while those of the $(\bar{5} + 10)$ "low mass" fermions belong to non-equivalent representations of $SU(2)_{\text{weak}}$. This is one way of creating naturally superlarge mass differences between fermions. "Low mass" and superheavy fermions belong to inequivalent representations of G and then simply G invariance forbids some of the fermions to be involved in Yukawa couplings containing superheavy Higgs, and thus G invariance protects some of the fermions from getting superheavy masses. This is not the whole story. It is possible to create naturally superlarge mass differences even for fermions belonging to the same representation. For example, in E_6 , a promising group for grand unification, we may put fermions in the fundamental 27 representation. But now, $27 \times 27 (= \bar{27} + 351 + 351')$ contains the 351, the Higgs representation that causes the superheavy symmetry breaking, and thus some of the fermions

contained in the 27 will become superheavy¹⁸⁾. Why do some and not all of the fermions become superheavy? It is the form of the 351 vacuum expectation value (VEV) that leaves some fermions massless. In other words, while E_6 invariance allows a $27 \times 27 \times \overline{351}$ coupling, the form of $\langle 351 \rangle$ is such that $27 \times 27 \times \langle \overline{351} \rangle$ contains some zero spots.

Let us now put the whole act together. "Low mass fermions are those that either G invariance prevents from participating in Yukawa interactions involving Higgs bosons (Φ) that cause the superheavy symmetry breaking, i.e., there is no $\psi\psi\Phi$ terms, or the particular form of the vacuum expectation values of these superheavy Higgs fields (ϕ) causes the vanishing of the $\psi\psi\langle\Phi\rangle$ terms".

This is a more technical form of the survival hypothesis. It is exactly this form of the survival hypothesis that, suitably twisted (translated), I want to use¹¹⁾ in order to solve the $(0,0,1)$ problem. Just re-read the aforementioned, deliberately detailed familiar analysis of natural ways to create superlarge fermion mass differences, using the following substitution rules:

Superheavy fermions	$\longrightarrow$	Third generation fermions
Low mass fermions	$\longrightarrow$	First and second generation fermions
G invariance and/ or particular form of VEVs	$\longrightarrow$	G' invariance and/or particular form of appropriate VEVs

This is my proposed solution to the $(0,0,1)$ problem. It is G' invariance and/or some particular forms of the VEVs of appropriate Higgs fields that are responsible for the $(0,0,1)$ zeroth order form of the fermion mass matrix. This is what I will call the "sidewise" or "horizontal" survival hypothesis.

It is the "standard" or "vertical" survival hypothesis that prevents all three generations from getting a superheavy mass, while it is the "sidewise" or "horizontal" survival hypothesis that protects the first and second generation from getting a "tree level" or "direct" low mass. A happy state of affairs? Almost. We have to find a way of producing naturally very small masses for the first two generations, as we will try to do next.

5.2 The $(\alpha^2, \alpha, 1)$ Problem

Let us proceed again by analogy. After the lengthy discussion in Section 2, it is apparent that if there are no post- G_1 ($\equiv SU(3) \times SU(2) \times U(1)$) interactions and given the observed particle content of the theory, baryon (B) and lepton number (L) conservation follows automatically. It is the existence of post G_1 interactions, i.e., grand unified interactions contained in G , that are responsible for the naturally "tiny" breaking of these global symmetries (B and L). We all know that this is an immediate consequence of the fact that

quarks and leptons belong to the same representation(s) of G . There are gauge (or Yukawa) interactions involving post G gauge or Higgs bosons that transform quarks to leptons and thus the very existence of these couplings breaks the global symmetries that were responsible for B and L conservation. The next step is clear, just make the substitution:

B, L quantum numbers	$\longrightarrow$	First and second generation masses
Global symmetry responsible for B and L conservation	$\longrightarrow$	Chiral symmetry responsible for $m_{1,2} = 0$
G interactions break B and L	$\longrightarrow$	G' (or sometimes even G) interactions, after G' breaking, are responsible for $m_{1,2} \neq 0$ by appropriately breaking the relevant chiral symmetries

and the way to solve the $(\alpha^2, \alpha, 1)$ problem becomes apparent. Presumably the fermions of the first two generations (call them "light") are sitting in the same representation(s) of G' together with other fermions of the superheavy and/or third generation type (call them "heavy"). The mere existence of gauge or Yukawa couplings involving a "light", a "heavy" and a "superheavy" gauge or Higgs boson that, after the G' breaking (some "heavy" fermions develop mass), breaks our naturally imposed zeroth order chiral symmetries. Then, the road leading to the creation of "light" fermion masses, necessarily in higher orders (involving G' interactions), is widely open. What a world! G' invariance prevents some of the fermions from getting "tree level" masses, while after G' breaking, G' interactions are responsible for the "tiny" masses of the first two generations. My claim is, that the "light" fermions would have been strictly massless if they were not sitting in the same G' representation(s) with "heavy" fermions. That is, in a strictly $SU(3) \times SU(2) \times U(1)$ world not only would we have B and L conservation, but also masslessness of the first two generations^{*)}: It is through the post- $SU(3) \times SU(2) \times U(1)$ interactions that the first two generations are getting naturally "tiny" masses. It is amusing to notice that we can carry this analogy even further. Just notice that $B-L$ violation in GUTs occurs in higher order compared to B, L violation. In amplitude²⁴⁾, $(B-L)_{\text{violat}} = (M_W/M)(B \text{ or } L)_{\text{violat}}$, which is exactly the same relation between the neutrino and charged fermion masses, Eq. (12).

According to the "Fermi-Syndrome" philosophy exposed in Section 2, at the $SU(3) \times SU(2) \times U(1)$ level the mass terms for the first two generations will have an "effective" non-renormalizable form, for example

^{*)} Conversely, the observed small masses of the first two generations may be taken as a signal of the existence of post $SU(3) \times SU(2) \times U(1)$ or even post- G interactions.

$$\left(\frac{\alpha}{\pi}\right)^n \frac{1}{(M)^k} (\bar{\Psi}\Psi) (\phi_1)^k \phi_2 \quad (17)$$

where ϕ_1 and ϕ_2 are Higgs fields (the subscript indicates their $SU(2)_{\text{weak}}$ multiplicity) able to get VEV of order $\langle\phi_1\rangle \sim M$ and $\langle\phi_2\rangle \sim M_W$, and n indicates the number of loops of the diagram from where this term emerges. Then (17) implies

$$m_{\text{"light"}} \simeq \left(\frac{\alpha}{\pi}\right)^n M_W \quad (18)$$

a highly desirable formula, since $n = 1$ corresponds to the second generation mass and $n = 2$ to the first generation mass. Then, applying the philosophy of Section 2, we should look for the corresponding diagrams that produce the "effective" mass terms (17) or (18).

i) $n = 1$

One may consider diagrams of the form^{18),25)}

FIG.1

where everything in the loop is superheavy. This diagram produces an "effective" mass term of the form (17) with $n = 1$ and $k = 2$. Thus, in this case, we get $m_2 \simeq (\alpha/\pi)M_W$. The physics here is very interesting because the crucial mixing of the two gauge bosons is a direct consequence of the fact that Q and q have necessarily different $SU(2)_{\text{weak}}$ transformation properties as a consequence of the survival hypothesis.

ii) $n = 2$

An example of a diagram belonging to this category looks like²⁶⁾

FIG.2

in an obvious notation.

This diagram produces an effective mass term of the form (17) with $n = 2$, $k = 1$ and thus $m_1 \approx (\alpha/\pi)^2 M_W$. Of course, it is tacitly assumed here that, again because of the "internal" differences (under G') of the first and second generation, it is impossible to give mass to the first generation through one-loop diagrams. For more explicit applications of these ideas the interested reader should consult Ref. 25), 26), 27). In a few words, the third generation gets its mass from the same source that breaks $SU(2)_{\text{weak}}$ (i.e., gives the W bosons their masses) while the second and first generation cannot "feel" this $SU(2)$ breaking source. It takes the involvement of post- $SU(3) \times SU(2) \times U(1)$ interactions to create masses for the first two generations, which then necessarily have to appear as "radiative" corrections and thus naturally small. In a strictly $SU(3) \times SU(2) \times U(1)$ world the first two generations would be massless. Actually, since the first two generations have "radiative" masses, we expect deviations from "direct" or "tree level" quark-lepton mass relations like Eq. (7). Presumably such deviations will be much more drastic for the first than for the second generation and thus we expect $m_d \neq m_e$ and $m_s \approx m_\mu$, as seen experimentally.

Finally, as a by-product of this way of thinking about the fermion masses, one gets in many GUTs a firm prediction¹⁸⁾ of the top quark mass

$$m_t = m_c \cdot \frac{m_\tau}{m_\mu} \quad (19)$$

which, appropriately renormalized at low energies [like Eq. (8)], gives¹⁸⁾

$$m_t \approx (18 - 22) \text{ GeV} \quad (20)$$

not outside of the energy range capacities of PETRA or PEP.

We have deliberately tried to present the "basics" of GUTs without referring to particular models. The interested reader should consult Ref. 28) for a recent review on grand unified models. Grand unified models (GUMs) are great !

ACKNOWLEDGEMENTS

It is a great pleasure to thank my colleagues at the University of Athens, F. Hadjioannou, A. Lahanas and E.G. Papantonopoulos for their warm hospitality, discussions and reading of the manuscript. Discussions with Riccardo Barbieri and John Ellis who also read the manuscript are deeply appreciated.

REFERENCES

- 1) D.V. Nanopoulos, Nuovo Cimento Letters 8 (1973) 873;
S. Weinberg, Phys. Rev. Letters 31 (1973) 494.
- 2) S.L. Glashow, A. Salam and S. Weinberg, Nobel Lectures, Stockholm,
December 1979.
- 3) S. Weinberg, Phys. Rev. Letters 43 (1979) 1566;
F. Wilczek and A. Zee, Phys. Rev. Letters 43 (1979) 1571.
- 4) J.C. Pati and A. Salam, Phys. Rev. Letters 31 (1973) 661; Phys. Rev. D8
(1973) 1240; Phys. Rev. D10 (1974) 275.
- 5) H. Georgi and S.L. Glashow, Phys. Rev. Letters 32 (1974) 438.
- 6) J. Learned, F. Reines and A. Soni, Phys. Rev. Letters 43 (1979) 907.
- 7) D.V. Nanopoulos, CERN preprint TH.2896 (1980);
J. Ellis, CERN preprint TH.2942 (1980);
R. Barbieri, CERN preprint TH.2935 (1980).
- 8) S.L. Glashow and D.V. Nanopoulos, Nature 281 (1979) 464.
- 9) H. Georgi, H.R. Quinn and S. Weinberg, Phys. Rev. Letters 33 (1974) 451.
- 10) A.J. Buras, J. Ellis, M.K. Gaillard and D.V. Nanopoulos, Nucl. Phys. B135
(1978) 66.
- 11) D.V. Nanopoulos, "The fermion spectrum according to GUTs", CERN preprint
TH.2866 (1980).
- 12) D.V. Nanopoulos and D.A. Ross, Nucl. Phys. B157 (1979) 273.
- 13) J. Ellis, M.K. Gaillard, D.V. Nanopoulos and S. Rudaz, CERN preprint
TH.2833 (1980), Nucl. Phys. B in press, and references therein.
- 14) Irvine-Michigan-Brookhaven Collaboration, M. Goldhaber et al., "Proposal
for a Nucleon Decay Detector" (1979);
Harvard-Purdue-Wisconsin Collaboration, J. Blandino et al., "A Decay Mode
Independent Search for Baryon Decay Using a Volume Cerenkov Detector"
(1979);
University of Minnesota, H. Courant et al., "A Dense Detector for Proton
Decay" (1979);
Frascati-Milano-Torino Collaboration, G. Battistoni et al., "Proposal for an
Experiment on Nucleon Stability with a Fine Grain Calorimeter" (1979);
V.S. Narasimhan, "Test of Conservation of Baryon Number, Proposal for an
Experiment at KGF" (1980).
- 15) H. Georgi and D.V. Nanopoulos, Nucl. Phys. B155 (1979) 52, and unpublished.
M. Gell-Mann, P. Ramond and R. Slansky, in "Supergravity", Proc. of the
Supergravity Workshop at Stony Brook, ed. by P. van Nieuwenhuizen and
D.Z. Friedman (North Holland, Amsterdam, 1979) p. 315;
R. Barbieri, D.V. Nanopoulos, G. Morchio and F. Strocchi, Phys. Letters 90B
(1980) 91.
- 16) F. Reines, H.W. Sobel and E. Pasierb, Phys. Rev. Letters 45 (1980) 1307;
V.A. Lubimov et al., Phys. Letters 94B (1980) 266.

- 17) H. Georgi, Nucl. Phys. B156 (1979) 126.
- 18) R. Barbieri and D.V. Nanopoulos, Phys. Letters 91B (1980) 369.
- 19) R.J. Crewther, Phys. Rev. Letters 28 (1972) 1421;
M.S. Chanowitz and J. Ellis, Phys. Letters 40B (1972) 397 and Phys. Rev. D7 (1973) 2490.
- 20) J. Ellis, M.K. Gaillard and B. Zumino, Phys. Letters 94B (1980) 343;
J. Ellis, M.K. Gaillard, L. Maiani and B. Zumino, LAPP preprint TH-15/CERN preprint TH.2481 (1980);
B. Zumino, CERN preprint TH.2954 (1980).
- 21) J.M. Cornwall, P.N. Levin and G. Tiktopoulos, Phys. Rev. Letters 30 (1973) 1268;
Phys. Rev. D10 (1974) 1145
- 22) C. Bouchiat, J. Iliopoulos and P. Meyer, Phys. Letters 38B (1972) 519.
- 23) For recent reviews see
J. Ellis, M.K. Gaillard and D.V. Nanopoulos. CERN preprint TH.2858 (1980).
D.V. Nanopoulos, CERN preprint TH.2871 (1980).
- 24) D.V. Nanopoulos, D. Sutherland and A. Yildiz, Nuovo Cimento Letters 28 (1980) 205.
- 25) R. Barbieri and D.V. Nanopoulos, Phys. Letters 95B (1980) 43.
- 26) D.V. Nanopoulos, "A Solution to the Problem of the Fermion Mass", to be published in the proceedings of the XXth International Conference on High Energy Physics, Madison (July 1980).
- 27) R. Barbieri, D.V. Nanopoulos and D. Wyler, CERN preprint in preparation.
- 28) D.V. Nanopoulos, see Ref. 7).