INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE CENTRATOM TRIESTE

SMR.626 - 7
Lect. I

SUMMER SCHOOL IN HIGH ENERGY PHYSICS AND COSMOLOGY

15 June - 31 July 1992

GRAND UNIFICATION: CURRENT STATUS AND A FUTURE PERSPECTIVE

Jogesh C. Pati
Department of Physics and Astronomy
University of Maryland
College Park, MD 20742
USA

Please note: These are preliminary notes intended for internal distribution only.

GRAND UNIFICATION IN A BROADER PERSPECTIVE*

(In Honor of Andre Sakharov)

Jogesh C. Pati
Department of Physics and Astronomy
University of Maryland
College Park, Maryland 20742

UMDHEP 92-210
(December, 1991)

Talk presented at the A.D. Sakharov International Conference on Physics, held at the Lebedev Physical Institute, Moscow (May 27-31, 1991), to appear in the Proceedings.

*Research supported in part by a grant from the National Science Foundation, U.S.A.

ABSTRACT

The current status of grand unification is reviewed and a point of view regarding its "ultimate fate" is presented. As expressed at the time of its initiation, grand unification brings certain desirable features to physics. These include the inter-relationships between the fundamental particles and their forces and the associated consequences of quantization of electric charge and the non-conservations of B and/or L. *It is emphasized that these aspects of grand unification are so basic that they should survive, whatever be the ultimate picture.* This should be the case, even if grand unification, as proposed in the traditional sense, may well undergo *substantial modifications* within a more comprehensive theory, such as a superstring theory. Such a comprehensive theory should retain the successes of the traditional framework, while removing its shortcomings. Although superstring theories have yet to make contact with the low-energy world in a reliable way, there is every indication that they will lead to major changes in (a) the effective gauge symmetry and the field-content below the Planck scale, and (b) the unification-scale and thereby also in the predictions of proton life-time and $\sin^2\theta_W$, compared to those of the minimal and other traditional models (with or without supersymmetry). They would, of course, still retain all the desirable and conceptual aspects of grand unification - in particular the inter-relationships between the fundamental particles and their forces.

Whatever be the final picture, it should be devoid of the arbitrariness that is present in the traditional framework with respect to the choice of the multiplets and that of the associated masses and the coupling parameters. In particular, it should explain (a) the origin of three families, (b) the inter-family mass-hierarchy and (c) the origin of mass-scales spanning the range from M_{Planck} to m_W, and then to m_e and m_ν and thereby the origin of small numbers such as $(m_W/M_{Pl}) \sim 10^{-17}, (m_e/M_{Pl}) \sim 10^{-22}$ and $(m_{\nu_e}/M_{Pl}) \lesssim 10^{-27}$. The prospects of two alternative routes which may link superstring theories to low-energy physics–one conventional and the other rather unconventional–in addressing some of these issues are reviewed and certain crucial tests are mentioned. These can help distinguish between the two alternatives.

GRAND UNIFICATION IN A BROADER PERSPECTIVE

(In Honor of Andre Sakharov)

I. INTRODUCTION

Andre Sakharov was a great scientist and above all a great humanitarian-a combination that is rare to find anywhere at anytime. I regard it a special privilege to speak at this meeting, in honor of his memory.

In this talk I will present a perspective of grand unification and discuss some of its consequences, including in particular the non-conservation of baryon number. This seems fitting, because the topic of baryon non-conservation was close to Sakharov's heart. It is also very dear to me.

While the nature of baryon non-conservation pertains to de p fundamental physics, it still remains an unsettled issue experimentally. One thus wishes that forthcoming searches that would be possible through the Superkamiokande, Gran Sasso and other planned experiments would provide us with much desired information on the question of conservations of the two related fundamental entities: B and L.

In the context of grand unification I will emphasize the roles of $SU(4)$-color and left-right symmetry, because these two concepts appear to be among the central features of almost all grand unified schemes except $SU(5)$. As one consequence of left-right symmetry, I will discuss the question of non-vanishing neutrino masses, since neutrinos may well provide a major window for probing into physics beyond the standard model. In particular, the solar neutrino puzzle may already be pointing to massive neutrinos and in turn to the existence of an intermediate scale of order 10^{11} GeV, which would characterize the onset of fundamentally new physics.

At the end, I will present a personal view on what is likely to be the ultimate fate of grand unification, and thereby of B and L, in the light of recent theoretical developments which attempt to address certain basic issues and go beyond grand unification.

II. SAKHAROV AND BARYON ASYMMETRY

It was indeed a remarkable insight which Sakharov had, when he realised in 1967 [1] that three ingredients are necessary to understand how the universe could develop a baryon-asymmetry in the early stages of its evolution, starting with a state that is completely symmetric between matter and anti-matter. These are: (i) non-conservation of baryon number, (ii) C and CP non-conservation and (iii) thermodynamic non-equibrium of baryon non-conserving processes during the relevant epoch in the evolution of the universe. This suggestion of Sakharov, together with the subsequent motivation for baryon non-conservation (see below), is now regarded as the most natural resolution of a long-standing puzzle in cosmology: the origin of an excess of baryons over antibaryons.

I will return to the problem of baryon asymmetry later. For the present, let me say that it appears rather remarkable to me that this beautiful contribution of Sakharov remained unknown to most physicists in the world (at least to those living outside the Soviet Union) for almost a decade. This reflects the poor state of communication that existed until recently between physicists in the Soviet Union and those in the rest of the world. I myself came to know about Sakharov's work only in 1978, when several authors discussed the issue of baryogenesis in the context of grand unification.

III. PHYSICS BEYOND THE STANDARD MODEL

Meanwhile, a completely independent motivation for non-conservation of baryon number had arisen starting in 1972, on purely aesthetic grounds, based on one's desire to unify basic particles (quarks and leptons) and their forces. Let me recall briefly the circumstances which led to this development.

The renormalizability of the spontaneously broken $SU(2)_L \times U(1)$ theory [2] was shown by 't Hooft in 1971 [3]. This drew much attention to the $SU(2)_L \times U(1)$ theory in the theoretical as well as experimental worlds. Long before its successes were revealed through the discoveries of the neutral current phenomena and charm during the years 1974 and 1975, however, it was noticed by Salam and myself in the summer of 1972 [4] that the $SU(2)_L \times U(1)_Y$-theory cannot be a fundamental theory by itself, even if it is eventually borne out to a high degree of accuracy by low-energy experiments. This is because it possesses much arbitrariness, first, in its gauge-sector. In particular, the choice of the $U(1)_Y$-quantum numbers of the various fields is rather arbitrary, which is made just to fit the known charges, subject to $Q_{em} = I_{3L} + Y/2$. So also is the choice of

the weak angle tan θ_W denoting the ratio of the $U(1)_Y$ to the $SU(2)_L$ coupling constant. Moreover, the theory can not provide satisfactory answers to questions such as:

(1) Why do there exist quarks with strong interactions and leptons without such interactions?

(2) Why do there exist weak, electromagnetic *and* strong interactions with their effective strengths at low energies as observed?

(3) Why is electric charge quantised with $Q_{e^-} + Q_p = 0$?

(4) Why are the weak interactions *universal* with regard to quarks and leptons even though the strong interactions are not?

(5) Why do the negatively charged electron and the positively charged proton, rather than the positron and the proton, exhibit the same sign of longitudinal polarization in weak decays?

Furthermore the $SU(2)_L \times U(1)_Y$-theory possesses a number of arbitrary parameters in the Higgs-sector as well, involving the Higgs-mass, the Higgs-quartic and in particular the Higgs-Yukawa couplings, which respectively determine the masses of the gauge bosons and of the fermions. The arbitrariness of the standard model is reflected by the fact that it has some *nineteen arbitrary parameters* comprising the 3 gauge coupling constants (g_1 , g_2 , g_3), the hierarchical 9 fermion masses (i.e. m_u, m_d, m_s, m_c, m_e, m_μ, m_t, m_b and m_τ), the three Cabibbo-Kobayashi-Masakawa mixing angles ($\theta_{1,2,3}$), the two masses m_W and m_ϕ, the CP-violating phase δ and the angle $\bar{\theta} = \theta_{QCD} - \theta_{weak}$ associated with strong CP-violation. These are in addition to the arbitrariness in the choice of the hypercharge quantum numbers Y.

Believing that a fundamental theory must be devoid of such arbitrariness, one is thus led to believe firmly that there exist truly new physics beyond the standard model which should ideally be capable of removing the conceptual shortcomings as well as the arbitrariness mentioned above. In particular this new physics ought to provide natural answers to questions such as (1) – (5) and should be devoid of arbitrary parameters. How much of this new physics should show, how and at what energies depend, of course, on its nature.

Possible Resolutions

There are a few basic suggestions listed below which seem promising in removing some or perhaps all of these shortcomings although the task is far from being accomplished at present.

(1) <u>Grand Unification:</u>

The first suggestion in this regard was that of grand unification [4,5,6]. It comprises the unification of quarks and leptons and simultaneously the unity of the weak, the electromagnetic and the strong forces. One major motivation for such a unification is that it assures quantisation of electric charge. It also provided answers to all the conceptual questions (1)-(5) raised above. The price one had to pay, however, if it is a price, for putting quarks and leptons in one multiplet is that B and L can no longer be absolute [4]. Owing to limits from the *Eötvos* - type experiments, it turns out that B and/or L must be violated either explicitly as in $SU(5)$, or spontaneously as in a "maximal" symmetry like $SU(16)$, see later. In symmetries like $SO(10)$, B and B + L are violated explicitly, but B-L is is violated spontaneously. Invariably, such violations lead to proton- decay at some level, which is thus regarded as a hall mark of quark-lepton unification.

(2) Supersymmetry:

An important idea which was proposed subsequently is that of Supersymmetry [7]. It comprises the idea of a symmetry between fermions and bosons. As a local rather than a global symmetry, supersymmetry requires the existence of gravity and is called supergravity [8]. As such, local supersymmetry is likely to be utilised by nature at a fundamental level to bring in unification of gravity with the other forces. Supersymmetry has also the additional virtue of taming quantum corrections and thereby maintaining a large hierarchy in mass-ratios such as $(M_{Pl}/m_W) \simeq 10^{17}$ in spite of quantum corrections. By itself, it does not, however, explain the origin of such large hierarchies.

(3) Compositeness:

The proliferation of quarks, leptons and Higgs bosons on the one hand and that of the parameters associated with their masses and couplings on the other hand have led to the idea that all these particles may be composites of a more elementary set of objects which should bring economy and elegance at a fundamental level [9]. This idea seems to be especially promising when it is combined with the idea of local supersymmetry - i.e. supergravity [10]. In this latter form, it seems to be capable of providing a natural reason for the origin of hierarchical mass scales from M_{Planck} to m_ν -as well as an explanation for the origin of family replication. The peculiarity of supersymmetry plays a crucial role in achieving both features. One major virtue of this approach is that it leads to a host of predictions which can be tested by existing and forthcoming experiments including LEP II, LHC and SSC.

(4) Superstrings:

Last but not least, the most recent idea is that of superstrings [11] which presumes that the fundamental constituents of all matter are not truly pointlike but are rather extended stringlike objects with sizes $\sim (M_{Planck})^{-1} \sim 10^{-33}\ cm$. Consistency often demands that these theories be defined in dimensions higher than four and possess local supersymmetry. It turns out that this simple extension from pointlike to stringlike constituents requires the existence of gravitational as well as Yang-Mills forces. The superstring theories do not exhibit, at present, the manifest uniqueness in their formulations and in their solutions which was believed to be the case in the early stages. Yet, they are the most promising theories at a fundamental level in that they almost certainly yield a well-behaved theory of quantum gravity and seem capable of unifying all the forces of nature including gravity. In this last sense, they comprise the basic idea of grand unification [12] and go beyond it.

The ideas mentioned above are clearly not mutually exclusive. The "final" picture, which could address successfully to all the issues raised above may well involve a combination of some of these ideas. In fact, as mentioned above, the superstring theories, comprise local supersymmetry as well as the concept of conventional grand unification if they yield elementary quarks, leptons and Higgs bosons near the Planck Scale. Alternatively, they could comprise local supersymmetry and the idea of compositeness of quarks, leptons and Higgs bosons, if they yield elementary preonic superfields rather than quarks and leptons near the Planck scale [13,10]. Even in this case, the spirit of grand unification would be retained at the preonic level. I shall mention briefly these possibilities at the end. For now, I will focus on conventional grand unification and present some of its motivations and consequences.

IV. GRAND UNIFICATION

4.1 Unity of Quarks and Leptons

Salam and I observed [4] that at least the arbitrariness of the $SU(2)_L \times U(1)_Y$-theory in its gauge-sector can be removed and rather satisfactory answers to the questions raised above can be obtained provided we overcome the age-old *psychological barrier* which existed between quarks and leptons by assuming that these two sets of particles are members of one multiplet of a big nonabelian symmetry G (semisimple or simple) and provided we generate weak, electromagnetic as well as strong interactions by gauging this symmetry-structure G.

The question arises how best to implement these two related ideas in the context of a viable model. One idea which appeared to me most natural and simple then and still does is to unite three quark-colors of a given flavor and a lepton of the same flavor into a quartet of a vectorial gauge symmetry $SU(4)$-color. *In other words, lepton-number is simply the fourth color* [4].

• To implement this idea, we must assume first of all that there exists a right-handed neutrino ν_R (in addition to ν_L) to provide the fourth $SU(4)$-color-partner of the three right-handed up quarks $\left(u_{red} u_{yellow} \text{ and } u_{blue}\right)_R$.

• Second, we must assume that $SU(4)$-color breaks spontaneously at some heavy scale $M_X \gg m_W$ into $SU(3)^{col} \times U(1)_{B-L}$, and $SU(3)^{col}$-local gauge symmetry associated with the three quark-colors must be used to generate the basic strong interactions of quarks. *Thus, as a by-product of the desire to unify quarks and leptons through $SU(4)$-color, one was naturally led to suggest that $SU(3)^{col}$ gauge symmetry should accompany the commuting electroweak symmetry $SU(2)_L \times U(1)_Y$ and that the corresponding gauge force is the sole origin of the strong nuclear force* [4,14]. The subsequent discovery of the property of asymptotic freedom of non-abelian gauge forces [15] provided strong support to this idea about the origin of the nuclear force.

• While the basic symmetry $SU(4)$-color treats quarks and leptons universally, the subgroup $SU(3)$-color, does not. It operates only on quarks, but not on leptons. Thus, the existence of $SU(4)$-color together with its breaking into $SU(3)^{col} \times U(1)_{B-L}$ explains why there exist quarks *and* leptons with quarks exhibiting strong interactions but not leptons.

• Furthermore, if one associates the familiar weak interactions with the forces generated by the flavor gauge-symmetry (see below), such a symmetry must commute with the full $SU(4)$-color for renormalizability; hence *weak interactions must be universal with regard to all four colors, i.e. with respect to quarks and leptons.*

• Finally the non-abelian nature of the full gauge structure G (simple or semisimple) explains quantization of electric charge (see below). And, it turns out that in simple realizations of the idea of quark-lepton unification one obtains $Q_{e^-} + Q_p = 0$ (rather than $Q_{e^+} + Q_p = 0$). In turn this explains why the negatively charged electron and the positively charged proton exhibit the same sign of polarization in β-decay, rather than the positron and the proton.

Thus the idea of quark-lepton unification in the context of a simple or semi-simple group provides answers to some of the fundamental questions (1)-(5) raised in the introduction which could not be answered within $SU(2)_L \times U(1)_Y$ or within $SU(2)_L \times U(1)_Y \times SU(3)^{col}$.

4.2 Parity As an Intrinsic Symmetry of Nature

Another important by-product of $SU(4)$ - Color is left-right symmetry [4,16,17]. First of all, being a vectorial gauge symmetry, it forces the existence of all the eight right-handed fermions (F_R^e) accompanying their left-handed counterparts (F_L^e) in a given family. This implies that there must exist a right-handed neutrino ν_R^e accompanying ν_L^e. *In other words, basic matter must be left-right symmetric in contrast to the standard model.*

What about the gauge interactions in the flavor-sector? The answer to this question turns out to be rather unique in the sense that once we gauge $SU(4)^{col}$ and also demand that quantization of electric charge should be a compelling feature of the theory, the minimal gauge symmetry satisfying these properties turns out to be [4]

$$\mathcal{G}_o = SU(2)_L \times SU(2)_R \times SU(4)^{col} . \quad (1)$$

The representations of the left and right-handed fermions of the electron-family with respect to the symmetry group $\mathcal{G}_o$ are shown below:

$$F_L^e = \begin{pmatrix} u_r & u_y & u_b & u_\ell & = & \nu^e \\ d_r & d_y & d_b & d_\ell & = & e^- \end{pmatrix}_L = (2,1,4)$$

$$F_R^e = \begin{pmatrix} u_r & u_y & u_b & u_\ell & = & \nu^e \\ d_r & d_y & d_b & d_\ell & = & e^- \end{pmatrix}_R = (1,2,4) \quad (2)$$

The μ and τ-family fermions have identical representations. Here (r, y, b) denotes the three quark-colors (red, yellow and blue) and ℓ denotes lepton-number or leptonic color, which is named "lilac".

In other words, $SU(4)$ - color and the demand of quantization of electric charge forces that we must gauge $SU(2)_R$ (not just I_{3R}) which introduces three new gauges bosons $W_R^\pm$ and W_R^3. While the standard gauge bosons ($W_L^\pm$ and W_L^3) couple to the left-handed fermions F_L^e, the new ones $W_R^\pm$ and W_R^3 couple in an exactly parallel manner to the right-handed fermions F_R^e.

Parity is conserved in the basic interactions, if we now demand that the bare $SU(2)_L$ and $SU(2)_R$ coupling constants are equal and that the full lagrangian of the fermions, the gauge bosons and the Higgs scalars possesses a discrete left-right symmetry P which is broken only spontaneously [4,16,17]. Observed left-handedness and parity violation at low energies is attributed to spontaneous breaking of $L \leftrightarrow R$ symmetry which makes W_R's heavier than W_L's. Such a view of the origin of parity violation, in spirit, is the same as the one of unification of gauge forces which also breaks spontaneously. In this sense, conservation of parity ($L \leftrightarrow R$ *symmetry*) in the basic equations is aesthetically more appealing than the case where parity violation is put in by hand — as in the standard model.

From the discussions presented above, it is apparent that *the ideas of $SU(4)$-color and left-right symmetry, although independent, are essentially inseparable from each other.* The assumptions of $SU(4)$-color and quantization of electric charge enforces left-right symmetry. Conversely, the assumptions of left-right symmetry and quantization of electric charge imply $SU(4)$-color.

The electric charge now has an elegant composition in terms of the diagonal generators I_{3L}, I_{3R} and F_{15} belonging to $SU(2)_L, SU(2)_R$ and $SU(4)$-color. Here F_{15} is the fifteenth generator of $SU(4)$-color given by $\left(\frac{1}{2\sqrt{6}}\right)(1,1,1,-3)$ for the four colors (r, y, b, ℓ). Noting that F_{15} is proportional to $(B-L)$ [18], the electric charge takes the form

$$Q_{em} = I_{3L} + I_{3R} + \sqrt{\frac{2}{3}}\, F_{15} = I_{3L} + I_{3R} + \frac{B-L}{2} . \quad (3)$$

Note that $I_{3R} + \sqrt{\frac{2}{3}} F_{15} = I_{3R} + (B-L)/2$ plays the role of the weak hypercharge Y of $SU(2)_L \times U(1)_Y$. Since I_{3R} and F_{15} are quantised and have well defined values for a given representation, the arbitrariness in the choice of Y is removed.

4.3 Spontaneous Breaking of Parity and of SU(4)-Color

One presumes that there is a two-stage spontaneous breaking of $\mathcal{G}_o = SU(2)_L \times SU(2)_R \times SU(4)^C$ into $U(1)_{em} \times SU(3)^C$. One desirable scheme for such a breaking is provided by the VEV's of two Higgs multiplets — $\Delta_R \sim (1, 3_R, 10^C)$ and $\phi \sim (2_L, 2_R, 1^c)$. The role of Δ_R turns out to be especially important to account for the light neutrino masses [19]. The actions of $\langle \Delta_R \rangle$ and $\langle \phi \rangle$ are outlined below:

$$\begin{array}{c} \mathcal{G}_o = SU(2)_L \times SU(2)_R \times SU(4)^C_{L+R} \\ \downarrow \ \langle \Delta_R \sim \left(1, 3_R, 10^C\right) \rangle >> 1 \text{ TeV} \\ SU(2)_L \times U(1)_Y \times SU(3)^C \\ \downarrow \ \langle \phi \sim \left(2_L, 2_R, 1^C\right) \rangle \sim 250 \ GeV \\ U(1)_{em} \times SU(3)^C \end{array} \tag{4}$$

Consistent with the demand of left-right symmetry of the basic lagrangian, the multiplet $\Delta_R \sim \left(1, 3_R, 10^C\right)$ is introduced together with a multiplet $\Delta_L \sim (3_L, 1, 10^C)$ with identical mass and coupling parameters so that the lagrangian respects the basic symmetry ($F_L \leftrightarrow F_R, W_L \leftrightarrow W_R, \Delta_L \leftarrow \Delta_R$). It can be shown [17] that despite the left-right symmetry of the basic lagrangian, absolute minimum of the potential involving the Higgs field can lead to a left-right symmetric pattern of VEV such that either Δ_R or Δ_L has a non-zero VEV while the other has a zero VEV in the limit $\langle \phi \rangle = 0$. [This happens when the mutual coupling of Δ_L with Δ_R is sufficiently strong compared to their quartic self couplings]. Although no simple criterion exists as to whether $\langle \Delta_R \rangle$ or $\langle \Delta_L \rangle$ should be non-zero, let us assume $\langle \Delta_R \rangle \neq 0$ and $\langle \Delta_L \rangle = 0$, which is one of the two allowed solutions.

It is easy to convince oneself that there is just one component of $\Delta_R \sim (1, 3_R, 10^C)$ which is electrically neutral. This is the component which has $I_{3R} = +1$ and is like a dilepton in the $SU(4)$-color space, having $B - L = -2$. In other words, it is the component having the quantum numbers of two right-handed neutrinos "$\nu_R \nu_R$".

$$\langle (\Delta_R)^{\ell\ell}_{I_{3R}=+1} \rangle \neq 0 \Longrightarrow \Delta I_{3R} = \pm 1 \, ; \quad \Delta(B-L) = \mp 2 \, . \tag{5}$$

Thus, $\langle \Delta_R \rangle$ breaks $SU(2)_R$ and $(B - L)$, but preserves $SU(3)$-color.

The second stage is more standard. The Higgs potential may be arranged so that the multiplet ϕ transforming as $(2_L, 2_R, 1)$ under $\mathcal{G}_o$ has the VEV:

$$\langle \phi \rangle = \begin{array}{c} I_{3_R} \backslash I_{3_L} \rightarrow \\ -1/2 \\ +1/2 \end{array} \begin{array}{cc} +1/2 & -1/2 \\ \multicolumn{2}{c}{\begin{pmatrix} \kappa & 0 \\ 0 & \kappa' e^{i\delta} \end{pmatrix}} \end{array} \tag{6}$$

where κ and κ' are less than or of order 250 GeV. They break $SU(2)_L \times U(1)_Y$ into $U(1)_{em}$ giving masses to $W_L^{\pm}$ and Z^o while photon remains massless. The phase δ induces spontaneous CP violation [20]. The parameter $\kappa\kappa'$ induces $W_L - W_R$ mixing. The Yukawa couplings of ϕ and $\tilde{\phi} \equiv \tau_2 \phi^* \tau_2$ with the fermions of the type $h_{ij} \bar{F}^i_L F^j_R \phi + \tilde{h}_{ij} \bar{F}^i_L F^j_R \tilde{\phi} + hc$, where i, j are family indices, give masses to the fermions and induce Cabibbo-type mixings.

It turns out that a consistent description of quark-lepton masses requires the introduction of an additional Higgs multiplet $\xi = (2, 2, 15)$ which is non-trivial under $SU(4)$-color and, thus, its VEV distinguishes between quarks and leptons. It too can have Yukawa couplings with the fermions of the type $\eta_{ij} \bar{F}^i_L F^j_R \xi + h.c.$ While the introduction of ξ would influence the fermion masses, it would not, of course, alter the pattern of gauge symmetry breaking exhibited in (4).

4.4 Massive Neutrinos:

In left-right symmetric theories, since ν_R and ν_L co-exist, neutrinos are expected to be massive. First, the neutrino in a given family acquires a Dirac mass through the VEV of ϕ which is comparable, within a factor of two to three (say) [21], to the mass of the up-quark member in the same family:

$$\begin{aligned} & m(\nu^i)_D \bar{\nu}^i_R \nu^i_L + hc \\ & m(\nu^i)_D \sim h^i \kappa \, . \end{aligned} \tag{7}$$

In other words, one might expect $m(\nu_e)_D \sim 1$ to few MeV, $m(\nu_\mu)_D \sim 300$ to 500 MeV (say) and $m(\nu_\tau)_D \sim$ (20-30) GeV (say). The subscript D denotes Dirac mass, while i is the family index.

In addition, the right-handed neutrinos can acquire heavy Majorana masses through the VEV of $\langle \Delta_R \rangle = v_R \gg 1 TeV$. One can write (symbolically) a Majorana type invariant Yukawa coupling:

$$h_M \left(F^T_R C^{-1} F_R \Delta_R + F^T_L C^{-1} F_L \Delta_L \right) + hc \, . \tag{8}$$

Now, with $\langle \Delta_R \rangle^{\ell\ell}_{I_{3_R}=+1} = v_R \neq 0$, but $\langle \Delta_L \rangle = 0$, the right-handed neutrino acquire a heavy Majorana mass (through a term $m_R \nu^T_R C^{-1} \nu + hc$) which violates lepton number by two units. The majorana mass is given by:

$$m_R = h_M \, v_R \, . \tag{9}$$

In the presence of a Dirac-mass, one thus has the so-called *see-saw type mass-matrix* given by:

$$\begin{array}{c} \\ \nu_L^i \\ \bar{\nu}_R^i \end{array} \begin{array}{c} \nu_L^i \quad \bar{\nu}_R^i \\ \begin{pmatrix} m_L \simeq 0 & m_D^i \\ m_D^i & m_R \end{pmatrix} \end{array} i = e, \mu, \tau; . \tag{10}$$

For simplicity, we are assuming that the majorana mass m_R is family-independent in contrast to the Dirac mass m_D^i. We have also ignored family-mixing. Given that $m_D^i << m_R$ and m_L is smaller than $(m_D^i)^2/m_R$, the mass-matrix (10) diagonalizes approximately to the form:

$$\begin{bmatrix} -(m_D^i)^2/m_R & 0 \\ 0 & m_R \end{bmatrix} . \tag{11}$$

Thus the right-handed neutrino has a heavy Majorana mass m_R while the observed neutrino which is essentially ν_L has a tiny Majorana mass $(m_D^i)^2/m_R$ which is much lighter than the Dirac mass m_D^i. In other words, the smallness of the mass of the observed neutrino is related to the heaviness of W_R and, thereby, to the suppression of $V + A$ interactions at low energies. From the quark-masses, we expect (as mentioned before) $m_D^i \simeq (1-2)$ MeV, (300-500) MeV and (20-30) GeV for ν_e, ν_μ and ν_τ respectively. Allowing for some representative choices for the heavy scale m_R (motivations for some of these choices will be seen later in the context of grand unification), we obtain the following sets of values for the masses of the three left-handed neutrinos.

$$\begin{array}{llll} & m(\nu_{eL}) & m(\nu_{\mu L}) & m(\nu_{\tau L}) \\ m_R \sim 10^4\ GeV & (1-4)\times 10^{-1}\ eV & (1-4)\times 10^4\ eV & (4-9)\times 10\ MeV \\ m_R \simeq 10^{10.5}\ GeV & (3-12)\times 10^{-8}\ eV & (3-12)\times 10^{-3}\ eV & (2-27)\ eV \\ m_R \simeq 10^{14}\ GeV & (1-4)\times 10^{-11}\ eV & (1-4)\times 10^{-6}\ eV & (4-9)\times 10^{-3}\ eV . \end{array} \tag{12}$$

These should be regarded approximate to within a factor of 2 to 5 (say) owing to uncertainties in m_D^i. These masses are typically too small — especially for the second and the third lines — to be measured directly in the laboratory from either end-point electron spectra in β-decay experiments or neutrino-oscillation experiments. But the second and the third lines can be relevant for a possible resolution of the solar neutrino puzzle. The classic experiments of Davis which detects $\nu_e's$ from $\nu_e + cl^{37} \to e^- + A^{37}$ have been reporting for sometime a reduction in the solar neutrino flux which is about three to four times lower than that expected from the standard solar model (SSM). Recently, the Kamiokande experiment in Japan, which detects $\nu_e's$ by $\nu_e + e^- \to \nu_e + e^-$ by observing the recoil electrons in a water Cerenkov detector, also reports a reduction in the solar neutrino flux, by about a factor of two compared to SSM. The two experiments can be

compatible with each other if one takes into account certain intrinsic differences between them such as differing thresholds for neutrino energies above which they are sensitive. A reduction in the solar neutrino flux compared to SSM is the so-called solar neutrino puzzle. It was noted by Mikhayev, Smirnov and Wolfenstein [22,23] that matter-induced enhancement of the conversion of ν_e to ν_μ or ν_τ, even for small mixing angles, can lead to a significant reduction of solar neutrinos of the electron type on earth owing to a resonance-effect.

For example, both the chlorine and the Kamiokande experiments can be explained by the MSW effect for reasonable mixing angles, with $\Delta m_{ie}^2 \equiv m_{\nu_i}^2 - m_{\nu_e}^2 \simeq 10^{-6}$ to $10^{-7} eV^2$, where $\nu_i = \nu_\mu$ or ν_τ. More precisely, one requires $(\Delta m_{ie}^2) sin^2\theta \approx 10^{-8} eV^2$, where θ is the $\nu_i - \nu_e$ mixing angle in vacuum [24] (In theoretical models, $sin^2\theta \leq 0.1$. Thus, for $sin^2\theta \approx 10^{-2}$ to 10^{-1}, one needs $\Delta m_{ie}^2 \approx 10^{-6}$ to $10^{-7}\ eV^2$. Allowing for some uncertainty in m_D^i and m_R, such values of Δm^2 may be obtainable for the intermediate m_R (*e.g.* $m_R \simeq 10^{11}$ to $10^{11.5} GeV$ gives $\Delta m_{e\mu}^2 \equiv m_{\nu_\mu}^2 - m_{\nu_e}^2 \simeq 10^{-6}$ to $10^{-7} eV^2$) and also for a superheavy m_R (*e.g.* $m_R \simeq 10^{14.5} GeV$ gives $\Delta m_{e\tau}^2 \simeq 10^{-7} eV^2$).

Most recently, the Soviet-American Gallium Experiment (SAGE), located in the Baksan Laboratory in the caucuses, and the GALLEX-experiment, located in the Gran Sasso tunnel in Italy, have started taking data on solar neutrinos. These can detect *relatively low-energy neutrinos* by the reaction $\nu_e + Ga^{71} \to e^- + Ge^{71}$. As such, these experiments are sensitive to neutrinos coming from the pp-reaction in the sun (i.e. pp $\to de^+\nu_e; E\nu_e \geq 420 KeV$), which are the predominant source of the solar neutrinos and whose flux can be calculated without much uncertainty. These experiments can, therefore, provide in the near future definitive information on the reality of the solar neutrino puzzle and the relevance of the MSW mechanism.

Establishment of non-vanishing neutrino masses by whatever means, i.e. solar-neutrino studies, neutrino-oscillation experiments and/or direct mass-measurements, would provide an important clue to the nature of the basic theory. *In particular, it would be a strong signal for left-right symmetry since that implies the co-existence of left and right-handed neutrinos. This in turn provides a compelling reason for neutrinos to have at least a Dirac mass.* By contrast, for a left-right asymmetric theory like $SU(2)_L \times U(1)_Y$ or its extension $SU(5)$ (see later), there is no natural reason for the existence of ν_R; nor is there any good reason for an artificially small Majorana mass for ν_L [25].

If the observed neutrinos turn out to have non-vanishing masses (rather than zero), one would naturally ask what is so special about them so that they are the only fermions which are so light compared to all the other fermions? In retrospect, the mechanism [19] based on heavy majorana mass of ν_R's provides the aesthetically most pleasing answer in this regard, because the neutrinos are the only fermions which are electrically neutral and thus the only ones which can acquire Majorana masses without violating the conservation of electric charge.

Furthermore, note that the smallness of neutrino masses gets related, within this picture, to the heaviness of W_R's and, thereby, to the degree of parity violation, which is represented by the ratio $\left(m^2_{W_R} - m^2_{W_L}\right)/(m^2_{W_R})$. By the same token, in as much as the only known natural mechanism for ensuring light neutrino masses is the one of see-saw, *observation of such light masses, in particular the establishment of the MSW explanation of the solar neutrino puzzle, would be a clear indication for the existence of new physics at a superheavy scale (most probably around $10^{11} - 10^{14}$ GeV).*

4.5 The Advantages of $SU(2)_L \times SU(2)_R \times SU(4)^c$: A Summary

To summarise, the symmetry group $\mathcal{G}_o = SU(2)_L \times SU(2)_R \times SU(4)^c$ introduces a number of new concepts which are aesthetically desirable:

(i) Quark-lepton unification through $SU(4)$-color.

(ii) Left-right symmetry and, thereby, the view that parity is an exact symmetry of nature at a basic level, which is violated only spontaneously.

(iii) Quantisation of electric charge.

(iv) Massive neutrinos which may be relevant to a resolution of the solar neutrino puzzle.

(v) Last but not least, through $SU(4)$-color, $\mathcal{G}_o$ introduces $B-L$ as a local symmetry and thereby a massless gauge particle coupled to $B-L$. Following arguments based on *Eötvos*-type experiments (see later), it follows that the corresponding gauge particle must acquire a mass through spontaneous symmetry breaking. This in turn means that $B-L$ must be violated spontaneously at some level (For example, $< \Delta_R >$ violates lepton number L and thus $B-L$ by two units, see eq.(5)). *In short, owing to quark-lepton unification, $SU(4)$-color implies that B and/or L can not be absolute* [4]. This is, of course, welcome to implement baryogenesis, following the suggestion of Sakharov [1].

Because of these special and most desirable features, I believe that the symmetry group $\mathcal{G}_o = SU(2)_L \times SU(2)_R \times SU(4)^c$ is likely to be part of a fundamental theory. In particular, the new concepts brought in by it are likely to survive one way or another. In essence, it is distinguished from the standard model and its extension such as $SU(5)$ (see later) by virtue of possessing left-right symmetry and thereby having massive neutrinos as a compelling feature. Any group G (simple or semisimple) which contains $\mathcal{G}_o$ as a subgroup (see below) would, of course, retain all the advantages of $\mathcal{G}_o$, listed above.

4.6 Embedding of $SU(2)_L \times SU(2)_R \times SU(4)^c$ in a Simple Group: Unity of Forces

It is natural to extend $\mathcal{G}_o$ further so that one may relate the coupling constant g_2 of $SU(2)_{L,R}$ with that of $SU(4)^{col}$. Such a unification will embody quark-lepton unification

as well as unification of the Yang-Mills forces. These two unifications, together, constitute "Grand Unification." The known differences between the $SU(2)_L, U(1)_Y$ and $SU(3)^{col}$ coupling constants must then be interpreted as low energy effects arising from a spontaneous breaking of the grand unification symmetry G into $SU(2)_L \times U(1) \times SU(3)^{col}$, together with finite renormalisation effects (see below).

Several grand unification groups suggest themselves as candidates for extension of $\mathcal{G}_o$: $[SU(4)]^4, SU(8) \times SU(8), SO(10), SU(16)$ and E_6 [26]. Of these $SO(10)$ is probably the most direct and obvious extension, because $SU(2)_L \times SU(2)_R$ is isomorphic to $SO(4), SU(4)^{col}$ is isomorphic to $SO(6)$ and $SO(4) \times SO(6)$ can simply be extended to $SO(10)$. $SO(10)$ is also the maximal one family symmetry operating on 16 two component fermions which is anomaly-free like $\mathcal{G}_o$. On the other hand, $SU(16)$ is the "maximal" one family-symmetry which gauges all available degrees of freedom within a family, consisting of sixteen two-component fermions (i.e. F_L and $(F^c)_L$; see below. It contains $SO(10)$ and $SU(8) \times SU(8)$ as subgroups. For anomaly-cancellation, $SU(16)$, as also $SU(8) \times SU(8)$ and $[SU(4)]^4$, require new exotic fermions, for example mirror fermions. (While these ideas were consistent at the time they were originally proposed, it seems that mirror families with chiral coupling are on the verge of being in conflict with recent precision electroweak measurements). A few typical representations of $SO(10)$ are mentioned below:

SO(10): Since under charge conjugation right-handed fermions become left-handed anti-fermions, $((\psi_R)^C = (\psi^C)_L$, here C denotes charge conjugation), one can combine the eight two-component left-handed fermions F_L transforming as $(2_L, 1, 4^c)$ under $\mathcal{G}_o$ with the eight two-component left-handed anti-fermions $(F^c)_L = (F_R)^c \sim$ transforming as $(1, 2_R, 4^{*c}$ under $\mathcal{G}_o$ to make a $\underset{\sim}{16}$ of $SO(10)$, which is the spinor representation of $SO(10)$.

For the electron-family, the corresponding fermions are:

$$\underset{\sim}{16}^e_L = (F^e_L|(F^c)^{e_L}) = \left[(u)_{r,y,b}, (d)_{r,y,b}, \nu^e, e^- |(\bar{u})_{r,y,b}, (\bar{d})_{r,y,b}, \bar{\nu}^e, e^+\right]_L . \quad (13)$$

Note that the right-handed neutrino ν_R is included as $(\bar{\nu})_L$. A few representations of $SO(10)$ and their decompositions under $\mathcal{G}_o = SU(2)_L \times SU(2)_R \times SU(4)^C$ as well as decompositions of some product representations under $SO(10)$ are given below [27]:

$$\begin{aligned}
&\underline{10} = (2,2,1) + (1,1,6)\\
&\underline{16} = (2,1,4) + (1,2,4^*)\\
&\underline{45} = (3,1,1) + (1,3,1) + (1,1,15) + (2,2,6)\\
&\underline{54} = (1,1,1) + (2,2,6) + (3,3,1) + (1,1,20)\\
&\underline{120} = (2,2,15) + (3,1,6) + (1,3,6^*) + (2,2,1) + (1,1,20)\\
&\underline{126} = (3,1,10) + (1,3,10^*) + (2,2,15) + (1,1,6)\\
&\underline{10} \times \underline{10} = \underline{1} + \underline{45} + \underline{54}\\
&\underline{10} \times \underline{16} = \underline{16}^* + \underline{144}\\
&\underline{16} \times \underline{16} = \underline{10} + \underline{120} + \underline{126}\\
&\underline{16} \times \underline{16}^* = \underline{1} + \underline{45} + \underline{210}.
\end{aligned} \qquad (14)$$

The gauge multiplet of $SO(10)$ transforms as $\underline{45}$. Note that a scalar multiplet transforming as $\underline{126}$ of $SO(10)$ contains the $\Delta_R \sim (1,\underline{3},10^c)$ and its "left-handed" partner $\Delta_L \sim (3,1,10^{*c})$. VEV of Δ_R was utilized in section 4.3 to provide a desirable breaking pattern for $SU(2)_L \times SU(2)_R \times SU(4)^C$ and simultaneously to give a heavy Majorana mass to ν_R. The same purpose is now served by the VEV of $\underline{126}$. The Yukawa couplings of fermions with scalars "H" will have the symbolic form $h_{ij}(\Psi^{iT}_{\underline{16},L}C^{-1}\Psi^{j}_{\underline{16},L})H + hc$, where i,j refer to family indices. The scalars "H" would thus belong, in general, to the product $(\underline{16} \times \underline{16}) = \underline{126} + \underline{10} + \underline{120}$. The VEV of $\Delta_R \sim (1,3,10) \subset \underline{126}$ gives Majorana mass to ν_R. The $(2,2,1)$ component of $\underline{10}$ is assumed to have VEV's of the electroweak scale ($\sim$ 250 GeV), which breaks $SU(2)_L \times U(1)_Y$ to $U(1)_{em}$. Two complex 10's are used to give appropriate Dirac masses and mixings to all the fermions including neutrinos. Since the (2,2,1) is singlet of $SU(4)$-color, the $\underline{10}$ would lead to the mass equalities $m_e^{(o)} = m_d^{(o)}, m_\mu^{(o)} = m_s^{(o)}, m_\tau^{o} = m_b^{(o)}, m_u^{(o)} = (m_{\nu_e})_D, m_c^{(o)} = (m_{\nu_\mu})_D$ and $m_t^{(o)} = (m_{\nu_\tau})_D$ at the $SU(4)$-unification scale. As mentioned before, these relations are altered at low energies due to gluonic corrections, which would enhance quark masses but not the lepton-masses. A realistic pattern of fermion masses also needs VEV of $(2,2,15) \subset \underline{126}$, which would distinguish between quarks and leptons. The VEV of $\underline{120}$ can contribute to

the mixing between different families.

Two alternative patterns of spontaneous breaking of $SO(10)$ are shown below:

$$\text{(II)}\quad SO(10) \xrightarrow[(1,1,1)]{\underline{54}} SU(2)_L \times SU(2)_R \times SU(4)^C \xrightarrow{\langle\Delta_R\rangle \subset \underline{126}} SU(2)_L \times U(1)_Y \times SU(3)^C \xrightarrow{\langle\underline{10}\rangle} U(1)_{em} \times SU(3)^C \qquad (15)$$

$$\text{(II)}\quad SO(10) \xrightarrow[(1,1,15)]{\underline{45}} SU(2)_L \times SU(2)_R \times SU(3)^C \times U(1)_{B-L} \xrightarrow{\langle\Delta_R\rangle \subset \underline{126}} SU(2)_L \times U(1)_Y \times SU(3)^C \xrightarrow{\underline{10}} U(1)_{em} \times SU(3)^C . \qquad (16)$$

One may generally assume that the VEV of $\langle\underline{54}\rangle$ or of $\langle\underline{45}\rangle$ together with that of $\langle\underline{126}\rangle$ give a two or one-stage breaking of $SO(10)$ to $SU(2)_L \times U(1)_Y \times SU(3)^C$, their scales being far above the electroweak scale. Their scales will be determined in the context of renormalization group analysis (see later).

$SO(10)$ may be embedded in higher symmetry groups like the exceptional group E_6 [26] which unify the $\underline{16}$-plet of fermions of $SO(10)$ together with new fermions in a $\underline{27}$:

$$\left(\underline{27}\right)_{E_6} = \left(\underline{16} + \underline{10} + \underline{1}\right)_{SO(10)} . \qquad (17)$$

Note the existence of $SO(10)$-singlet and decouplet within the 27-plet of E_6. The symmetry group E_6 arises naturally from the $E_8 \times E_8$ heterotic superstring theory.

$SO(10)$ may alternatively be embedded within the "maximal" symmetry $SU(16)$ which gauges all degrees of freedom of the $\underline{16}$-plet [26]. The same $\underline{16}$-plet which is the spinor representation of $SO(10)$ serves as the fundamental representation of $SU(16)$. For

$SU(16)$, one would need new families of fermions, for example left-handed mirror families transforming as $\underset{\sim}{16}^*_L$, to cancel anomalies. The symmetry $SO(10)$ may alternatively be embedded in still bigger symmetries like $SO(14)$ or $SO(18)$ [29] which unify more than are family of fermions (These symmetries also introduce mirror families together with standard families in one irreducible representation). These alternative embeddings of $SU(2)_L \times SU(2)_R \times SU(4)^C$ via $SO(10)$ are shown below:

$$\begin{array}{ccccl} & & & \nearrow & E_6\left(\underset{\sim}{27}_L\right) \\ SU(2)_L \times SU(2)_R \times SU(4)^C & \longrightarrow & SO(10) & \longrightarrow & SU(16)\left(\underset{\sim}{16}_L + \underset{\sim}{16}^*_L\right) \\ (2,1,4)_L + (1,2,4^*)_L & & \underset{\sim}{16}_L & \searrow & SO(14) \text{ or } SO(18) \end{array}$$

4.7 The Minimal Grand Unification Symmetry *SU(5)*:

If one did not insist on either $SU(4)$-color or left-right symmetry, one could embed $SU(2)_L \times U(1)_Y \times SU(3)^C$ within $SU(5)$ [5], which is the minimal grand unifying symmetry. In this case, the 15 two-component fermions in a family (not including the right-handed neutrino ν_R or its charge conjugate $(\bar{\nu})_L$) are assigned to two different representations of $SU(5)$ — $\underset{\sim}{5}^* + \underset{\sim}{10}$. The compositions of $\underset{\sim}{5}^*$ and $\underset{\sim}{10}$ for the electron family are shown below:

$$\psi_{\underset{\sim}{5}^*L} = \left(d^c_r, d^c_y, d^c_b | e^-, -\nu_e\right)_L \tag{18}$$

$$\psi_{\underset{\sim}{10}L} = \begin{array}{|c|c|c|c|c|} \hline 0 & u^c_b & u^c_y & -u_r & -d_r \\ \hline -u^c_b & 0 & u^c_r & -u_y & -d_y \\ \hline -u^c_y & -u^c_r & 0 & -u_b & -d_b \\ \hline u_r & u_y & u_b & 0 & -e^+ \\ \hline d_r & d_y & d_b & e^+ & 0 \\ \hline \end{array} \tag{19}$$

Here the superscript c denotes charge conjugate. Thus $(u^c_b)_L$, for example, denotes left-handed anti-up quark with anti-blue color, which is the antiparticle of $(u_b)_R$. The multiplets $\underset{\sim}{5}^*$ and $\underset{\sim}{10}$ individually give rise to triangle anomalies, which mutually cancel each other.

$$A\left(\underset{\sim}{5}^*\right) + A\left(\underset{\sim}{10}\right) = 0\,. \tag{20}$$

The generator $(Y/2)$ of standard model with $Q_{em} = I_{3L} + Y/2$ to a suitably normalized $SU(5)$ generator Y' (which satisfies, $Tr\,(F_i^2) = \frac{1}{2}$) by

$$[Y']_{SU(5)} = \sqrt{3/5}\;[Y/2]_{Stand.\ Model} \tag{21}$$

$SU(5)$, being the minimal grand unification symmetry, is clearly the easiest to work with. It also serves as a *prototype* for all grand unification symmetries by illustrating some major features of grand unification. For this reason, it is the symmetry which is most widely studied. But it is not necessarily the most elegant of all grand unification symmetries in that it assigns members of even one family to two different multiplets 5^* and 10 and, in this sense, goes against one major motivation behind grand unification — *i.e.* unification of matter.

For this reason, I believe that $SU(5)$ is unlikely to be a fundamental symmetry. It may, of course, arise effectively as a subgroup of bigger symmetries like $SO(10)$ or $SU(16)$ or E_6 which do assign members of one family to one multiplet and in addition possess the elegant features of left-right symmetry and $SU(4)$-color. In fact the $\underset{\sim}{16}$-plet of $SO(10)$ (or $SU(16)$) decomposes under $SU(5) \subset SO(10)$ as follows:

$$\left[\underset{\sim}{16}\right]_{SO(10)} = \left[5^* + 10 + 1\right]_{SU(5)}\,. \tag{21}$$

Thus the $\underset{\sim}{16}$ two-component fermions of $SO(10)$ precisely contain the 5^* and 10 of $SU(5)$. In addition, it contains a singlet of $SU(5)$, which is $(\bar{\nu})_L$ or equivalently the right-handed neutrino ν_R. From the viewpoint of $SO(10)$, the existence of ν_R is essential. But from the view-point of $SU(5)$ regarded as a fundamental symmetry, the introduction of a singlet ν_R is not compelling. The gauge bosons of $SU(5)$ may be arranged in a traceless 5×5 matrix $V = \sum_{a=1}^{24} V^a \lambda^a/2$ given by:

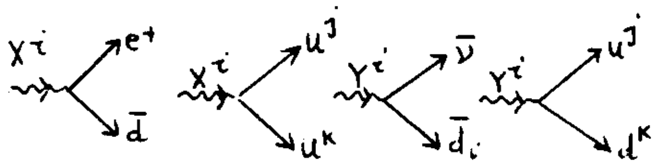
$$V = \left|\begin{array}{c|c} \underset{gluons}{V(\underset{\sim}{8})} \; - \sqrt{\tfrac{2}{15}}\, B^\circ\,(\,\mathbb{1})_{3\times 3} & \begin{array}{cc} X_1 & Y_1 \\ X_2 & Y_2 \\ X_3 & Y_3 \end{array} \\ \hline \begin{array}{ccc} \bar{X}_1 & \bar{X}_2 & \bar{X}_3 \\ \bar{Y}_1 & \bar{Y}_2 & \bar{Y}_3 \end{array} & W(\underset{\sim}{3}) - \sqrt{\tfrac{3}{10}}\, B^\circ\,(\mathbb{1})_{2\times 2} \end{array}\right| \tag{22}$$

The $(V(\underset{\sim}{8}), W(\underset{\sim}{3})$ and B° represent the gauge bosons of $SU(3)$-color, $SU(2)_L$ and $U(1)_Y$ respectively. The X's and Y's couple to lepto-quark and diquark currents made of fermions in 5^* and 10 as shown below:

$$\mathcal{L}(X,Y) = \frac{g}{\sqrt{2}} \bar{X}^i_\mu \left[\bar{e}_L \gamma^\mu d^c_{iL} + \epsilon_{ijk} \bar{u}^{ck}_L \gamma^\mu u^j_L + \bar{d}^c_L \gamma^\mu e^+_L\right]$$

$$+ \frac{g}{\sqrt{2}} \bar{Y}^i_\mu \left[-\bar{\nu}_L \gamma^\mu d^c_{iL} + \epsilon_{ijk} \bar{u}^{cj}_L \gamma^\mu d^k_L - \bar{u}_{iL} \gamma^\mu e^+_L\right] + h.c. \tag{23}$$

The indices i,j,k refer to (r,y,b) color indices of quarks. These give rise to vertices drawn below:



B	$-1/3$	$2/3$	$-1/3$	$2/3$
L	-1	0	-1	0
$B-L$	$2/3$	$2/3$	$2/3$	$2/3$

We see that X's as well as Y's couple to currents with differing B and L but the same $B-L$. As a result they would mediate B and L non-conserving four-fermion interactions of the type $uu \to \bar{d}e^+$ and $ud \to \bar{d}\bar{\nu}$ which would induce proton decays of the type $p \to e^+\pi^o$ and $p \to \bar{\nu}\pi^+$. These violate B and L but conserve $B-L$. As we will see later, B and L violations would turn to be a generic feature of almost all grand unification symmetries having their roots in quark-lepton unification [4]. The amplitudes for proton-decay interactions are proportional to $\left(g^2/m^2_{X,Y}\right)$. These would give proton-lifetimes of order $\left[m^4_{X,Y}/\left(g^4 m^5_p\right)\right]$. The fact that the proton is so long-lived ($\tau_p > 10^{31}$ years) clearly implies that X,Y need to be superheavy ($m_{X,Y} > (10^{14} - 10^{15}$ GeV)). It will turn out that renormalization group analysis will imply (X,Y) to have masses in this range. The important fundamental feature is that grand unification has brought a new complexion to particle physics –i.e. non-conservations of baryon and lepton numbers with a long-lived but unstable proton. I will return to a more detailed discussion of this point.

Breaking SU(5)

$SU(5)$ can be broken spontaneously to the standard model symmetry $SU(2)_L \times U(1)_Y \times SU(3)^C$ by using the VEV of a scalar multiplet Σ transforming as a $\underset{\sim}{24}$ of $SU(5)$,

which can be broken further to $U(1)_{em} \times SU(3)^C$ by the VEV of a scalar multiplet H transforming as a $\underset{\sim}{5}$ of $SU(5)$.

$$SU(5) \sim \underset{\underset{\sim}{24}}{\xrightarrow{\langle\Sigma\rangle}} SU(2)_L \times U(1)_Y \times SU(3)^C \sim \underset{\underset{\sim}{5}}{\xrightarrow{\langle H\rangle}} U(1)_{em} \times SU(3)^C \tag{24}$$

$$\langle\Sigma\rangle = v_\Sigma \; diagonal \left(1,1,1,-\frac{3}{2},-\frac{3}{2}\right),$$

$$\langle H\rangle = v_H \,(0,0,0,0,1)\,. \tag{25}$$

It is easy to see that $\langle\Sigma\rangle$ will give masses to (X,Y)-gauge bosons: $m_X = m_Y \sim g v_\Sigma$ leaving the 12 gauge bosons of $SU(2)_L \times U(1)_Y \times SU(3)^C$ massless, while $\langle H\rangle$ will give masses to $W^\pm$ and Z^o leaving the photon and the 8 gluons massless.

On phenomenological grounds, such as the longevity of the proton and the renormalization group analysis (see later), one needs $v_\Sigma \gg v_H$. Typically, one would need $v_\Sigma \sim 10^{12} v_H$. Such a large disparity between the VEV's of two different Higgs multiplets is, on the one hand, a *generic feature* of all grand unification models. On the other hand, it is difficult to preserve such a hierarchy naturally in the presence of radiative corrections to the masses and the quartic couplings of the Higgs Scalars. This is the standard *gauge hierarchy problem* of grand unification [30]. It turns out that at least a technical resolution of this problem is obtained through supersymmetry which avoids the necessity for fine tuning (see remarks later).

Fermion Masses:

Since the left-handed fermions belong to $\bar{5}+10$ of $SU(5)$ and $\bar{5}\times\bar{5} = \overline{10}+\overline{15}$, $\bar{5}\times 10 = 5+\overline{45}$, and $10\times 10 = \bar{5}+45+50$, it is clear that the scalar multiplets Φ_K, which can have gauge invariant Yukawa couplings with the fermions of the form $h^{ab}_{ijk}\left(\Psi^{(i,a)}_L\right)^T C^{-1} \Psi^{(j,b)}_L \Phi_K + hc$, must transform as either $\bar{5}, \overline{10}, \overline{15}, \overline{45}$, or 50 of $SU(5)$ or their hermitean conjugates. Here i,j,k denote $SU(5)$ representations and a,b denote family-indices (e,μ,τ). Of these, only $\bar{5}$ and $\overline{45}$ have $SU(3)$-color singlet neutral member transforming as a member of $SU(2)_L$-doublet which can thus acquire VEV preserving $U(1)_{em}\times SU(3)^C$ and satisfying the $\Delta I_{weak} = 1/2$ rule for $SU(2)_L \times U(1)_Y$ breaking [31].

The *minimal* $SU(5)$-model, which has the most predictive power, introduces only $\underset{\sim}{24}$ and $\underset{\sim}{5}$ of Higgs. Since $\underset{\sim}{24}$ cannot have Yukawa coupling, only $\langle H_{\underset{\sim}{5}}\rangle$ contributes to fermion masses in the minimal model. This leads to the quark-lepton mass equalities:

$$m^{(o)}_d = m^{(o)}_e\,,\; m^{(o)}_s = m^{(o)}_\mu \;\text{ and }\; m^{(o)}_b = m^{(o)}_\tau \tag{26}$$

which hold at the grand unification scale. Recall that the same relations emerged in $SU(2)_L \times SU(2)_R \times SU(4)^C$ with only (2,2,1) of Higgs and in $SO(10)$ with $\underset{\sim}{10}$ of Higgs, as expected because $[\underset{\sim}{10}]_{SO(10)} = [5+5^*]_{SU(5)}$. Although these relations are modified at low energies due to gluonic radiative corrections [32] which yield the desired relation $m_b \simeq 2.7\ m_\tau$, for the physical masses, one still has the bad relation $(m_e/m_\mu) = (m_d/m_s)$, where as, "experimentally," $m_e/m_\mu \approx 1/200$ and $m_d/m_s \approx 1/18$. To cure this problem, one must introduce a $\underset{\sim}{45}$ of $SU(5)$ which, by itself gives $m_e^{\prime(o)} = 3m_d^{\prime(o)}, m_\mu^{\prime(o)} = 3m_s^{\prime(o)}$, and $m_\tau^{\prime(o)} = 3m_b^{(o)}$. The prime signifies contribution from $\langle\underset{\sim}{45}\rangle$. With both $\langle\underset{\sim}{5}\rangle$ and $\langle\underset{\sim}{45}\rangle$ contributing to fermion masses, one can not, in general, obtain any mass relation. But one may choose specific patterns of Yukawa couplings of $\underset{\sim}{5}$ and $\underset{\sim}{45}$ with suitable discrete symmetries to obtain desired patterns of fermion masses. In particular, one may arrange to have masses for the $e-\mu$ sector given by [33]

$$\begin{array}{c} \\ e \\ \mu \end{array}\!\!\begin{array}{c} e \quad \mu \\ \begin{pmatrix} 0 & a \\ a & 3b \end{pmatrix} \end{array} ; \quad \begin{array}{c} \\ d \\ s \end{array}\!\!\begin{array}{c} d \quad s \\ \begin{pmatrix} 0 & a \\ a & b \end{pmatrix} \end{array} \tag{27}$$

with $m_b^{(o)} \simeq m_\tau^{(o)} \propto a$. Here a and b denote contributions from $\langle\underset{\sim}{5}\rangle$ and $\langle\underset{\sim}{45}\rangle$ respectively. Assuming $b \gg a$, one obtains $m_e/m_\mu \approx a^2/(9b^2) \simeq (1/9)(m_d/m_s)$, which is preserved in the presence of gluonic corrections. This is consistent with observation. In summary, allowing only renormalizable interactions, the Higgs-content of the minimal $SU(5)$ model needs to be extended to make it consistent with just the fermion masses. This, of course, reduces the predictive power of the minimal model considerably not only as regards fermion masses but also as regards the prediction of the grand unification mass scale (see later).

The $\langle\underset{\sim}{5}\rangle$ and $\langle\underset{\sim}{45}\rangle$ give masses to u, c and t quarks, but they keep ν_L massless.

4.8 The Meeting of the Running Coupling Constants

The analysis given below follows Ref. 6. It applies to all grand unification symmetries G which break in one step to $SU(2)_L \times U(1)_Y \times SU(3)^C$. Consider the breaking pattern:

$$\begin{array}{c} G = SU(5),\ SO(10),\ SU(16) \text{ or } E_8 \\ \downarrow \ \langle\Sigma_i\rangle = M/g,\ M_{X,Y} \sim M \\ SU(2)_L \times U(1)_Y \times SU(3)^C \\ \downarrow \ \langle\phi_j\rangle \simeq 250\ GeV \\ U(1)_{em} \times SU(3)^C . \end{array} \tag{28}$$

It is well known that coupling constants (as well as masses) are renormalized by radiative corrections. The renormalization effects depend on the momenta of the external legs. As a result, the effective coupling constants "run" as a function of momentum Q. For a grand unified theory, like $SU(5)$, $SO(10)$, $SU(16)$ or E_8, there is just one effective gauge coupling constant g at and above the grand unification scale M. Thus, at $Q \geq M$, the coupling constants of $SU(3)^C$, $SU(2)_L$ and suitably normalized $U(1)_Y$, denoted by $\bar{g}_3, \bar{g}_2$ and $\bar{g}_1$ respectively are "equal" [In general, they would at least be related by Clebsch-Gordan coefficients].

$$\bar{g}_3(M) = \bar{g}_2(M) = \bar{g}_1(M). \tag{29}$$

Now those gauge particles which are in G but outside of $SU(2)_L \times U(1)_Y \times SU(3)^C$ (i.e. the X_i and Y_i gauge particles in $SU(5)$) acquire masses of order M in the first stage of symmetry breaking. As a result, at momenta Q below M, these gauge particles make negligible contributions to renormalization effects which are damped by Q^2/M^2 and may, therefore, be dropped [34]. In fact, following the discovery of asymptotic freedom [15], it was realized [6] that the $SU(3)$ coupling constant would grow faster than the $SU(2)$-coupling constant with decreasing $Q \lesssim M$, while the $U(1)$-coupling constant would decrease (if one includes fermion and scalar contributions). This would then account for the observed disparities between the coupling constants at low energies i.e. $\bar{g}_3(m_W) \gg (\bar{g}_2(m_W), \bar{g}_1(m_W))$ - provided $M >> m_W$. This is one of the most beautiful gifts of grand unification, that was first noticed in Ref. 6.

To be more specific, the one-loop renormalization group analysis yields:

$$\beta^i \equiv \frac{d\bar{g}_i}{dt} = -b_i\bar{g}_i^3\ ;\ t \equiv \frac{1}{2}\ell n\, Q^2/\mu^2 . \tag{30}$$

Here $i = 1,2,3$ and μ is the renormalization or subtraction point. The coefficient b_i for $SU(N_i)$ is given by

$$b_i = \frac{1}{16\pi^2}\left[\frac{11N_i}{3} - \frac{2}{3}n_f - \frac{1}{3}t_2(s)_i\right] . \tag{31}$$

The first term is absent for the $U(1)$ coupling constant. Here n_f is the number of fermion flavors (left plus right handed up-quark count as one flavor). Thus, with three families, $n_f = 6$. The last term $t_2(s)$ denote contributions from light Higgs scalars with masses typically of order $m_W \ll M$. With a minimal set of Higgs scalars (*e.g.* $\underset{\sim}{5}$ and $\underset{\sim}{24}$ of $SU(5)$), the contribution of scalars to b_i is small compared to those of the other terms and may be dropped to a good approximation.

Choosing the renormalisation point $\mu = M$, where M is the grand unification scale, one then obtains (using the boundary conditions (29)):

$$1/g_i^2(Q) = 1/g^2(M) - 2b_i \ell n\, M/Q \tag{32}$$

$$b_1 = -\frac{(2/3)n_f}{16\pi^2}\,;\; b_{2,3} = \left[\frac{11}{3}(2,3) - \frac{2}{3}n_f\right]\frac{1}{16\pi^2} \tag{33}$$

These equations hold for $Q < M$. Contributions of scalars to b_i are dropped in (33) for simplicity. Combining the equations for g_1 and g_2 and using $\sin^2\theta_W \equiv e^2/g_2^2$ and $(e^2/g_1^2) = (3/5)\left(1 - \sin^2\theta_W\right)$, one thus obtains:

$$\sin^2\theta_W = \frac{3}{8} - \frac{55e^2}{96\pi^2}\ell n(M/Q) \tag{34}$$

Note that $\sin^2\theta_W$ equals 3/8 at the grand unification scale where $Q = M$. This is a general group theoretic property which holds for many other models including $SO(10)$, $SU(16)$ and E_6. Combining the equations for g_2 and g_3 we get:

$$\alpha/\alpha_s = \frac{3}{8} - [11\alpha/(24\pi)]\,\ell n\, M/Q\,. \tag{35}$$

Here $\alpha = e^2/4\pi$ and $\alpha_s = g_3^2/4\pi$, which should all be regarded as functions of the running momentum Q. Setting $Q = m_Z = 91.176 \pm .023 GeV$ and using $\alpha(m_Z) = (128.8)^{-1}$ and $\alpha_s(m_Z) = .115 \pm .008$, which are based on renormalisation group based extrapolation of observed values of α on the one hand (with the standard assumptions that the number of families = 3 and that there is just one Higgs-doublet) and the recent more accurate measurements at LEP [35] involving DELPHI, ALEPH, OPAL and L3-Collaborations on the other hand, one can predict M (using (35)) and, thereby, also $\sin^2\theta_W$ (using (34)). [See Ref. 36 and 37 for details of this anlaysis]. Including refinements due to Higgs contributions of the minimal model, one thus obtains:

$$M = \left(3.2^{+1.8}_{-1.0}\right) \times 10^{14}\, GeV$$

$$\sin^2\theta_W\,(m_Z) = 0.212^{+.003}_{-.004}\ \text{(Minimal } SU(5)\text{)}. \tag{36}$$

These are the predictions of the minimal $SU(5)$ model with the minimal Higgs content $\left(\underset{\sim}{5} + \underset{\sim}{24}\right)$ and the assumption of three families. Identical analysis would apply for other grand unification models like $SO(10)$, $SU(16)$ or E_6 as long as they break in one step to $SU(2)_L \times U(1)_Y \times SU(3)^C$, although the quantitative predictions, especially of M, would alter owing to their rich Higgs content (see remarks later). The best experimental value of $\sin^2\theta_W$ comes from the recent measurements at LEP of either the coupling constants measured at the Z^o-resonance or the value of m_Z alone. Using, for concreteness the DELPHI-value of $m_Z = 91.176 \pm .023 GeV$ [35], one obtains [See Ref. 36].

$$[\sin^2\theta_W(m_Z)]_{\overline{MS}} = 0.2336 \pm .0018\ .\ \text{(expt.)} \tag{37}$$

This allows for a range in the Higgs-mass from 45 to 1000 GeV. We see that the theoretical prediction of $\sin^2\theta_W$ given by (36) is clearly inconsistent with the observed value given by (37). This makes the minimal $SU(5)$ model inconsistent with observation.

The inconsistency is exhibited more dramatically by plotting theoretically extrapolated curves of $\alpha_1^{-1}, \alpha_2^{-1}$, and α_3^{-1} (see fig. 2) as functions of the running momentum Q, starting with their observed low-energy values (which include values of $\alpha_s \equiv \alpha_3, \alpha_2$ and α_1 at m_z). The curves are taken from the paper of Amaldi *et al.* in Ref. 36. As may be seen from fig. 2, the three curves clearly fail to meet at a common point.

We will see that the minimal $SU(5)$ model is excluded on the basis of measurements of the proton life-time as well. First, I present a general argument which shows that non-conservations of B and L follow as a consequence of quark-lepton unification.

4.9 Non-Conservations of B and L: A Generic Feature of Quark-Lepton Unification

The price which one must pay for putting quarks and leptons in one multiplet of a gauge symmetry G, is non-absoluteness of baryon- and lepton-number conservations. They must be violated at least spontaneously, if not explicitly, in the lagrangian, leading in almost all simple cases to the decay of a proton into leptons or antileptons. The proton life-time and decay modes will depend, of course, on the specific nature of the unification-symmetry and its spontaneous-breaking, but an unstable proton — or, at least some form of nonconservation of B and/or L will be common to all such schemes which comprise quark-lepton unification [4].

We already encountered an example of one kind of (B, L) violations in $SU(5)$, which are explicit in the lagrangian. But the property of (B, L) violations is much more general. The argument runs as follows: consider first the case in which one chooses to gauge either the "maximal" symmetry of the quark-lepton family, like $SU(16)$, [see ref. 26] so that both B and L are generators and are thus conserved in the basic lagrangian, or a subgroup (like $[SU(4)]^4$, $SU(8) \times SU(8)$ or $SO(10)$) containing $SU(4)$-color so that at least the linear combination $(B - L)$ is a generator. There are massless gauge particles coupled to B *and* L in the first case, and to $(B - L)$ in the second case, which are thus conserved in the basic lagrangian. These gauge particles must, however, acquire masses spontaneously in order that there be no conflict with the results from Eötvos type-experiments. The associated charges — *i.e.* B and L in the first case and $(B - L)$ in the second case — must, therefore, be violated *spontaneously*, although, naively, they are exact symmetries of the basic lagrangian. Alternatively, one may, of course, choose to gauge such subgroups of the maximal symmetry $SU(16)$ like $SU(5)$ and $SO(10)$ which retain quark-lepton unification but mix the diquark and quark-lepton gauge-currents of $SU(16)$ (see *e.g.* (23)) and, thereby, violate B and L *explicitly*. For $SO(10)$, the combination $(B - L)$ is still violated only spontaneously. *The violations of B and/or L must be there, owing just to quark-lepton unification and the constraints from Eötvos type experiments.* As mentioned above, specific nature of these violations would, of course, depend upon the unification scheme and the pattern of spontaneous breaking of the unification symmetry.

On the theoretical side, one can, *in general*, expect to see a variety of processes violating B and/or L; which include alternative modes of proton-decay, $n - \bar{n}$ oscillation [38] (satisfying $\Delta B = \pm 2, \Delta L = 0$) and neutrinoless double β-decay (satisfying $\Delta L = \pm 2, \Delta B = 0$). The processes in question, the associated selection-rules and the characteristic mass-scales for which these processes would be observable in current experiments (determined on dimensional grounds [39]) are listed in Table I.

4.10 Proton Life Time in the Minimal Models

If symmetries like $SO(10)$ or $SU(16)$ descend in one step to $SU(2) \times U(1) \times SU(3)^C$, which is the only possibility in $SU(5)$, one can show that the processes of type IA (*i.e.* $p \to e^+\pi^o, e^+\omega$, *etc.*) satisfying $\Delta(B - L) = 0$, which can occur through gauge boson exchanges (*e.g.* exchanges of (X, Y)-bosons) dominate over other processes. For minimal $SU(5)$ (with only $\underline{24}$ plus $\underline{5}$ of Higgs and 3 families), only $\Delta(B - L) = 0$ processes occur

and one gets a rather well-defined prediction for M (eq. (36)) and corresponding'y for proton lifetime. [See Ref. 37 for an analogous determination involving older data].

$$\Gamma(p \to e^+\pi^o)^{-1} \simeq 4 \times 10^{29.8\pm.6\pm.7} \left(\frac{M_X}{3.2 \times 10^{14}\, GeV}\right)^4 yr \tag{38A}$$

$$\lesssim 5.5 \times 10^{31}\ yr \quad \textit{(Minimal SU(5))} \tag{38B}$$

In the first equation (38A), allowance has been made for (a) an uncertainty by a factor of 5 in the rate due to uncertainty in the hadronic matrix element and (b) the uncertainty

in M_X as shown in (36). This may be compared with the experimental lower limit set by continued IMB search: [40].

$$\Gamma\left(p \to e^+\pi^o\right)^{-1}_{expt.} \gtrsim 5.5 \times 10^{32}\ yr. \tag{39}$$

This clearly excludes minimal $SU(5)$. The same conclusion is drawn from measurements of $\sin^2\theta_W$.

It is worth noting for comparison that one-step breaking of $SO(10) \to SU(2) \times U(1) \times SU(3)^C$ with minimal Higgs structure yields [41], taking recently determined value of Λ_S [36] as before,

$$\begin{aligned} M = 1.0 \times 10^{14}\ GeV, \quad \sin^2\theta_W = .217 \\ \Gamma(p \to e^+\pi^o)^{-1} \lesssim 10^{31}\ yr\ . \end{aligned} \tag{40}$$

In other words, M in this case is even lower than the case of minimal $SU(5)$, owing to contributions from the Higgs multiplets to the renormalization group equations. This is, of course, excluded by the IMB-limit on proton lifetime. The situation alters if we consider a two-step breaking of $SO(10)$ or of any other grand unification symmetry which contains $SU(4)$-color. This is discussed below.

4.11 Demise of Minimal SU(5): A Case For New Physics In the Grand Desert:

The inconsistency of minimal $SU(5)$, with both the observed value of $\sin^2\theta_W$ and the lower limit on proton lifetime should not, of course, be interpreted as evidence against the *idea* of grand unification. The fact that the predicted value of $\sin^2\theta_W$ is rather close to the observed value may suggest as one possibility that the idea of grand unification is correct but that there is some *new physics* lying between m_W and the grand unification scale M, which could still lie well below the Planck scale. In other words, the so-called grand desert — a must for $SU(5)$ — not a grand desert but rather it has some oases. This might be of some consolation to experimentalists.

What could be the nature of this new physics? Three possibilities suggest themselves. Two of these were in fact anticipated well in advance of the recent experimental measurements of $\sin^2\theta_W$ and τ_p.

(i) Higher Symmetries and Intermediate Scale: Symmetries like $SO(10)$, $SU(16)$ or E_6, which are aesthetically more elegant than $SU(5)$ (because they allow members of one family to be unified within one multiplet), permit a two-step breaking of the grand unification symmetry to the standard model symmetry involving an intermediate scale $M_I \sim 10^{11}\, GeV$ for $\sin^2\theta_W \simeq 0.23$. This has been noted by several authors [42-45,41].

For example, consider a two-step breaking of $SO(10)$ via two alternative routes (A and B). The mass-scales relevant to the two stages of SSB are denoted by M_U and M_I respectively.

$$SO(10) \xrightarrow[(A)]{M_U} SU(2)_L \times SU(2)_R \times SU(4)^c \xrightarrow{M_I} SU(2)_L \times U(1) \times SU(3)^c$$

$$SO(10) \xrightarrow[(B)]{M_U} SU(2)_L \times SU(2)_R \times U(1)_{B-L} \times SU(3)^c \xrightarrow{M_I} SU(2)_L \times U(1) \times SU(3)^c$$

Now there are three unknowns: M_U, M_I and $\sin^2\theta_W$, two of which can be predicted by the renormalization group equations for the running coupling constants provided one of them is introduced as an input. Tables IIA and IIB show the predicted values of M_U and M_I (for no. of families = 3 and $\Lambda_{\overline{MS}} = 160 MeV$) for different input values of $\sin^2\theta_W$ corresponding to the symmetry breaking patterns A and B respectively. It is interesting that M_U, and therefore proton lifetime, and $\sin^2\theta_W$ increase together [42-46,41], as needed by experiments.

Given the present experimental value of $\sin^2\theta_W = .2336 \pm .0018$ [35], which ranges between .23 and .235 say, the unification mass M_U for $SO(10)$ increases by a factor of nearly 4 to 6 (for Case A) and 4 to 7 (for Case B) compared to the $SU(5)$-value of $2.1 \times 10^{14} GeV$ (see eq. (36)) Correspondingly, allowing for the uncertainties in $\alpha_S(M_Z)$ and the hadronic matrix ele ent, as before, one obtains:

$$\begin{aligned}\Gamma(p \to e^+\pi^\circ)^{-1} &= (256 \quad 1296) \times (4.5 \times 10^{29\pm1.7} yr)(SO(10) - \text{Case A}) \\ &= (256 \quad 2401) \times (4.5 \times 10^{29\pm1.7} yr)(SO(10) - \text{Case B})\end{aligned} \quad (41)$$

Thus,

$$\begin{aligned}\Gamma(p \to e^+\pi^\circ)^{-1} &\leq 2.8 \times 10^{34}\ yrs\ (\text{Case A}) \\ &\leq 5.2 \times 10^{34}\ yrs\ (\text{Case B})\end{aligned} \quad (42)$$

In other words, *for the worst situation* from the viewpoint of observing proton decay, the inverse rate $\Gamma(p \to e^+\pi^\circ)^{-1}$ could be as high as (2 to 5)$\times 10^{34}$ yr (assuming a two-step breaking of $SO(10)$), but the actual value could be well below these limits (by factors of 10to100). For example, if $\sin^2\theta_W = .23, \Gamma(p \to e^+\pi^\circ)^{-1}$ should be less than 3.5×10^{33} yrs, for case A, and 5×10^{33} yrs, for Case B.

Given that the value of the weak angle $\sin^2\theta_W$ has gone up above the canonical value of 0.214, I feel that the idea of a two-step breaking of a grand unification symmetry, such as $SO(10)$, – should be considered seriously. Such a breaking, together with the observed value of $\sin^2\theta_W$, implies that the inverse rate for proton decaying via $(B-L)$-conserving $(e^+\pi^\circ)$-mode should lie in the range of $10^{32} - 10^{34}$ yrs (see eq. (41)). Strictly speaking, this is true with the inclusion of only gauge-boson-exchange-contributions to proton-decay amplitudes. In the presence of intermediate scales, Higgs-boson exchanges also become important. It has been noted [47,48] that Higgs-exchanges can contribute not only to $(B-L)$-conserving modes like $p \to e^+\pi^\circ$ but also to $(B-L)$ non-conserving ones, e.g. Type IIA and IIB satisfying $\Delta(B-L) = -2$ (see Table I):

$$\begin{aligned}&p \to e^-\pi^+\pi^+\ ,\ \mu^-\pi^+\pi^+\ ,\ e^-K^+\pi^+ \\ &n \to e^-\pi^+\ ,\ \mu^-\pi^+\ ,\ e^-K^+\end{aligned} \quad (43)$$

$$\begin{aligned}&p \to e^- + (e^+ + \nu) + (\pi^+ or K^+); p \to \mu^- + [(e^+\nu) or (\mu^+\nu)] + \pi^+ \\ &p \to \nu + [(e^+\nu) or (\mu^+\nu)]; n \to e^- + [(e^+\nu) or (\mu^+\nu)], etc.\end{aligned} \quad (44)$$

The amusing fact is that owing to the underlying $SU(4)$-color sy metry, the decay modes (44) can compete favorably and under certain circumstances ven supercede not only the decay modes (43) but also the $(B-L)$ conserving decay m des like $p \to e^+\pi^\circ$, etc. [See Ref. 48 for details].

Advantages of the Intermediate Scale

As we saw, an intermediate scale $M_I \sim 10^{11}\ GeV$, associated with a two-step breaking of a symmetry like $SO(10)$, leads to increased $\sin^2\theta_W$ and τ_p, in accord with experiments. Such an intermediate scale offers a few other advantages as well [46, 49-52].

(a) The Solar Neutrino Puzzle and Neutrino Masses: I have already noted in Section 4.4 the relevance of an intermediate scale of order $10^{11} GeV$ for spontaneous breakdown of $SU(2)_R$ and the generation of light neutrino masses via the see-saw mechanism such that they may be relevant to an explanation of the solar neutrino puzzle.

(b) Peccei-Quinn Symmetry Breaking: Several authors have noted that the existence of an intermediate scale of order $10^{11\pm1} GeV$, which could be associated, on the one hand, with the breaking of a symmetry like $SO(10)$, may also be associated with the breaking of Peccei-Quinn symmetry giving rise to axions whose scale is constrained on astrophysical grounds to lie in the range.

$$10^9 GeV \leq M_{PQ} \leq 10^{12} GeV \quad (45)$$

The lower bound is based on considerations of energy loss in stars due to emission of axions while the upper bound is based upon requiring that axions should not overclose the universe.

(ii) Marriage of Supersymmetry with Grand Unification:

An additional or alternative possibility for new physics in the desert is supersymmetry [7]. Supersymmetry is an attractive idea by itself demanding a symmetry between fermions and bosons. As a local [8] rather than a global symmetry it has the virtue that it *requires* the existence of gravity. As mentioned before, one fundamental reason for supersymmetry in its local form is thus the unification of gravity with the other forces — a feature that is promised in its full form by the superstring theories.

As a practical matter, supersymmetry provides at least a technical resolution [53] of the gauge-hierarchy problem mentioned in Section 4.7, because it tames quantum corrections and allows one to maintain a large *input* hierarchy in the masses of scalars (*e.g.* the $\underset{\sim}{24}$ and the $\underset{\sim}{5}$ of Higgs in $SU(5)$) without fine tuning, inspite of quantum corrections. This feature has turned out to be one of the most commonly cited reasons for supersymmetry.

If one combines the idea of supersymmetry with grand unification [54] and assumes that the supersymmetric partners of the particles in the standard model — *i.e.* squarks, sleptons, gluinos, winos, sinos, photinos, and higgsinos — have masses of order 1/10 to 10 TeV — something which one needs anyhow to avoid the unnatural fine tuning problem [55] — the superpartners contribute significantly to the renormalization group equations. As a result, the grand unification scale M as well as $\sin^2\theta_W$ are pushed up for a supersymmetric $SU(5)$ compared to their corresponding values in non-supersymmetric $SU(5)$. One typically obtains [36, 37].

$$M \simeq 10^{16\pm.3}\ GeV$$

$$\sin^2\theta_W = 0.237^{+.003}_{-.004} - \frac{4}{15}\frac{\alpha}{\pi}\ell n\left(m_{SUSY}/m_W\right) \qquad (46)$$

The value of $\sin^2\theta_W$ is in good agreement with experiments. This is reflected by the fact that, in contrast to non-supersymmetric minimal $SU(5)$, $\alpha_1^{-1}, \alpha_2^{-1}$ and α_3^{-1} now meet at a point around $10^{16} GeV$ [see fig. 3 which is taken from Ref. 56]. The high value of M implies that gauge boson exchanges would lead to a very slow proton-decay. These would give $\tau_p \approx 4\times 10^{33}$ to 10^{35} yrs. But, there are other important contributions to proton decay in SUSY theories due to exchanges of spin-0 squark as well as higgsinos, winos and gluinos. For example, the dimension-4 interactions $qq\tilde{q}$ and $q\tilde{q}\ell$ could give rise to very rapid proton decay through $\tilde{q}$ exchange, where $\tilde{q}$ denotes squark. These can be forbidden, however, by imposing discrete symmetries such as R-Parity on the theory. New contributions arise involving dimension 5 operators [57] which allow effective $qq\tilde{q}\tilde{\ell}$ as well as $\tilde{q}\tilde{q}q\ell$ interactions arising from exchanges of color-triplet higgsinos $\tilde{H}_3$. In one loop, involving for examples wino-exchange, these can give rise to dimension 6 $qqq\ell$ interactions, which are damped by only one power of a superheavy mass M that may be comparable within a factor of ten to M_U. Allowing for generation-mixing Yukawa

couplings, these typically lead to proton decays primarily via strange particle modes — *i.e.* $p \to \bar{\nu}K^+$, $n \to \bar{\nu}K^o$ — with lifetimes $\lesssim 10^{31}$ yrs. This is, of course, excluded by experiments. But in a SUSY theory, there is considerable uncertainty, in general, in the prediction of τ_p, and, if one wishes, one can avoid the dimension 5 operators by imposing symmetries. Not worrying much about proton lifetime, it seems fair to say, therefore, that if proton is found to decay *primarily* via the strange particle modes, that may indirectly be a signal for supersymmetry.

As is well known, from the absence so far of the superparticles - i.e. the squarks, the sleptons, the gluinos, the winos and the zino etc. – in the hadronic as well as the e^-e^+ machines, one infers that these superparticles can not be degenerate in mass with their SUSY partners - i.e. the observed ordinary particles. In other words, supersymmetry must somehow be broken. Globally supersymmetric grand unification models [54] do not, however, possess any simple or natural mechanism to break supersymmetry spontaneously. As a result, one inevitably ends up introducing soft supersymmetry breaking terms explicitly into the lagrangian. While such soft symmetry-breaking terms do not disturb the good quantum behavior of a SUSY theory (ignoring gravity), their introduction is still adhoc. In this respect, supergravity - grand unification models [58] possessing local supersymmetry do better, because they provide a more natural mechanism to break supersymmetry through the so-called *"hidden" sector*, which communicates with the known sector only through gravity. [Needless to say, if supersymmetry exists, it seems to make sense only as a local symmetry anyhow–i.e. in the form of supergravity - because gravity exists, and supersymmetry in the presence of gravity is necessarily supergravity]. Apriori, the idea of hidden sector may appear adhoc, but it turns out to be a compelling feature of a class of superstring theories, see remarks made later. Such a supergravity-grand unification model with a hidden sector generally leads effectively to a globally supersymmetric grand unification model in the observable sector near the unification-scale of order $10^{16} GeV$, with some universal soft SUSY breaking mass-terms m_o for all the scalars - i.e. the Higgs-Scalars, the squarks and the sleptons etc. Even with a positive m_o^2 near the Planck-Scale, radiative effects can turn the Higgs $(mass)^2$ to negative values for a sufficiently heavy top mass ($\sim 100 GeV$), depending upon m_o^2, and thereby induce electroweak-symmetry breaking [59]. These are some of the attractive features of supergravity–grand unification.

It is still far from being a complete theory, however, because, it allows an enormous degree of arbitrariness in the choice of the multiplets *and* of the parameters. In, particular, it does not account (a) for the origin of families, (b) that of the Higgs-multiplets which are needed and (c) for the partameters associated with their masses and the coupling constrants, which are rather large in number.

The same degree of arbitrariness exists, of course, for the minimal supersymmetric grand unification model, with or without local SUSY. *For this reason, its success as regards the meeting of the coupling constants [36], exhibited in fig. 3, which is no doubt* *impressive, should be taken with some caution.* Any eventual theory such as a superstring theory, or whatever, which removes the arbitrariness mentioned above need not reproduce, even as an effective theory, the minimal supersymmetric grand unification model near the Planck Scale – i.e. the precise gauge structure and the spectrum of particles associated with this minimal model. In fact there is not even a vague indication so far of such a model emerging from a superstring theory. Bearing this in mind, it is conceivable that the meeting of the coupling constants in the context of the minimal SUSY grand unification model might well be fortutious. There are in fact other ways by which such a meeting may occur. An example involving extra Higgs multiplets, notwithstanding its proliferation and conflict with proton lifetime, is given in Ref. 56. Another novel possibility will be mentioned below. *Whatever is the "final" picture, it should not only exhibit the meeting of the coupling constants but also remove the arbitrariness mentioned above.*

At any rate, if SUSY is relevant in controlling the mass of the Higgs scalar occuring in the electroweak theory, regardless of the relevance of the minimal SUSY grand unification model, the new physics mentioned above would involve the existence of $SUSY$ partners with masses $\lesssim 10\,TeV$ (say). These could be seen in either the running high energy machines, in particular the Tevatron (for masses of $SUSY$ partners less than about 200 GeV,) or future high energy accelerators — *i.e.* LHC and SSC (for SSC energies, the SUSY partners need to be lighter than about 5 TeV).

(iii) Planck Scale Grand Unification:

There is a third possibility which preserves the concept of grand unification but calls for a unification scale at or near the Planck scale. This goes most naturally with the idea that all forces including gravity are unified at and above the Planck scale where supersymmetry is also manifest. This is precisely the viewpoint of superstring theories. Within this viewpoint grand unification is intact above the Planck Scale, but may be lost just below it – for example in the process of compactification of a higher dimensional superstring theory to a four-dimensional one [11, 12]. This approach, while preserving the spirit of grand unification, has the potential for explaining the choice of the quark-lepton and the Higgs-multiplets and thereby the origin of families on the one hand, and that of the parameters associated with the quark-lepton masses, mixings and the Higgs-sector on the other hand.

I will present briefly two alternative possibilities which may arise within this approach in the concluding section. One of these corresponds to the conventional case in which the relevant superstring theory is assumed to yield elementary quarks, leptons and Higgs bosons near the compactification Scale. The other corresponds to a *non-conventional one* in which the fundamental fields emerging from the superstring theory, may represent preonic substructures. These bind at a lower scale to make composite quarks, leptons and Higgs bosons. Either case could preserve fully the spirit of grand unification – in particular, the idea of the unity of matter and of its forces. The two cases differ, of course, as regards the level at which such a unity manifests. I will indicate that the second case, although non-conventional, possesses certain advantages. It is also the one which turns out to be most amenable to experimental tests.

Before entering into this discussion in the concluding chapter, let me present some ideas on the origin of baryon-asymmetry of the universe.

4.12 Baryogenesis

While proton-decay in the laboratory is yet to be discovered, there is fairly strong indication that some form of baryon-nonconservation must have taken place in the very early universe. This is based on (a) the assumption (motivated by aesthetics) that the universe started with no initial matter-anti-matter asymmetry and (b) the observation that today at least our galaxy is made of all matter, in short that we exist. As mentioned in the introduction, this line of work started with a not well-noticed work of Sakharov in 1967 [1], preceded by a preliminary remark by Weinberg [1]. A similar idea was suggested independently by V. Kuzmin in a paper in 1970 [60], which also remained unnoticed. The basic idea was rediscovered by a number of authors [61] in 1978-79, after the nonconservation of baryon number emerged as a compelling feature due to grand unification [4,5].

To repeat, three ingredients are essential to generate an excess of baryons over antibaryons, subject to the assumption mentioned above. (i) Non-conservation of B, the need for which is apparent, (ii) C and CP violations and (iii) thermodynamic inequilibrium for baryon non-conserving reactions during the relevant epoch in the early universe. The need for the last two ingredients will be seen below.

It is interesting that any realistic grand unification model ends up satisfying all three ingredients. As we saw, non-conservation of B and L is a generic feature of grand unification. For example, the X gauge bosons of $SU(5)$ or $SO(10)$, to be denoted by X_V, and the corresponding Higgs scalars X_S decay into both $qq(B=2/3)$ and $\bar{q}\bar{\ell}(B=-1/3)$ channels (see fig.1) and thus violate B and L.

The necessity for C and CP violations follows by noting that these symmetries interchange particles and anti-particles. Thus, as long as they are conserved, no asymmetry between baryons and antibaryons can arise. To be more specific, if the branching ratios for the qq and $\bar{q}\,\bar{\ell}$ decay modes of X are denoted by r and $1-r$ respectively, the net baryon number created by each X-particle decay is $B_X = r(2/3)+(1-r)(-1/3)$. The $\bar{X}$ particles would decay into the corresponding antiparticle channels $\bar{q}\bar{q}(B=-2/3)$ and $q\ell(B=1/3)$, with the same total decay rate as X (by CTP). Let us denote the branching ratios for $\bar{X}$ decaying into these two modes by $\bar{r}$ and $1-\bar{r}$ respectively. As a result, the net baryon number created by the decays of each $(X,\bar{X})$ - pair is $\Delta B_X = (r-1/3)-(\bar{r}-1/3) = r-\bar{r}$. This would vanish if either C or CP is absolutely conserved, because in that case $r = \bar{r}$ and whatever baryon number is gained or lost by the decays of the X-particles would be exactly compensated by the decays of the $\bar{X}$-particles. This is the reason why the non-conservations of C and CP are also needed to generate a baryon excess.

Any realistic grand unification model must, of course, introduce either explicit violations of C and CP (as in $SU(5)$ with complex Yukawa couplings) or spontaneous violations of these symmetries (which may arise, naturally, e.g., in a left-right symmetric model like $SO(10)$), so as to account for observed violations of these symmetries. It turns out that, even with C and CP violations, some additional features must be fulfilled. In particular an interference between a tree and a loop-amplitude possessing a non-vanishing absorptive (imaginary) part is needed to ensure $r \neq \bar{r}$. Since each loop typically involves a factor of α (gauge) $\sim 10^{-2}$ or α (Yukawa) $< 10^{-3}$, ΔB_X is damped at least by ("α")^n where the absorptive part arises in n loops.

Finally, even if there are processes violating B, C and CP, if such processes are in thermal equilibrium, no baryon asymmetry can emerge. This is because, in thermal equilibrium, all states of a given energy are equally populated, and by CPT, particle and antiparticle states have the same mass and thus must have the same density. The only exception to this would have arisen if there was an initial (or pre-existing) difference between particle and anti-particle–populations and if such a difference was maintained owing to a conservation law. Since, by assumption, we are dismissing such an initial difference, it is necessary that there must exist B, C and CP violating processes that are *out of thermal equilibrium* to generate a net baryon-excess.

The condition of thermal inequilibrium of B-violating processes is met if the rates of all relevant B-violating processes, involving either the reactions mediated by the X-particles of the type $a + b \rightarrow c + d$ or the decays of the X-particles, are slower than the expansion rate of the universe during a certain epoch. The $2 \rightarrow 2$ reaction rates ($n\sigma$) are given by $\Gamma_R \simeq \alpha_X^2 T^5/(T^2 + M_X^2)^2$, where α_X and M_X denote the relevant couplings and the mass of the X-particles and T denotes the temperature. The decay rates of the X-particles (for $T < M_X$) are given by $\Gamma_D \simeq \alpha_X M_X$ where as the expansion rate of the universe is given by $\Gamma_{exp} \simeq g_*^{1/2} T^2/M_{Pl}$, where g_* is the number of available helicity states (typically, $g_* \approx 10^2$). Comparing these rates and also noting that for $T > M_X$, the depletion of the X and the $\overline{X}$ particles through their decays is exactly compensated by inverse decay-processes (leading to $(X, \overline{X})$-production), it turns out that only a certain range of temperature <u>below</u> M_X is relevant for baryogenesis . [Details of this discussion may be found in Ref. 62.] At these temperatures, baryogenesis would occur under certain conditions (see below) primarily through the decays of X-particles. [Note that at temperatures $T < M_X$, the rates of both the reaction processes and the inverse decays (damped by the Boltzmann factor) become unimportant.]

Now the condition of thermal inequilibrium for the decay processes (i.e. $\Gamma_D < \Gamma_{exp}$)

is met if $\alpha_X M_X < g_*^{1/2} T^2/M_{Pl}$. Since $T < M_X$, this requires:

$$M_X > g_*^{-1/2} \alpha_X M_{Pl} \approx \left(\frac{\alpha_X}{10^{-2}}\right) 10^{16} GeV . \tag{47}$$

The gauge particles X_V do not satisfy this condition for the minimal models for which $\alpha_{X_V} \approx 1/50$ and $M_{X_V} \approx (2 to 4) \times 10^{14}$ Gev. But, it can be satisfied for the Higgs scalars X_S for which the relevant Yukawa coupling α_{X_s} can be of order 10^{-5} to 10^{-6}. Thus the decays of the Higgs scalars can be responsible for baryogenesis in the minimal grand unification models. As a qualitative remark, however, it is worth noting that the condition of thermal inequilibrium given by eq. (47) typically requires that the X-particles must be heavier than about 10^{12} GeV. This feature goes naturally with grand unification in that the unification of coupling constants also requires the X-particles to be very heavy.

While grand unification thus has all the right ingredients for generating baryon-excess in the early universe, it turns out that the minimal $SU(5)$-model yields a number for the baryon to photon ratio (n_B/n_γ) that is too small by about six orders of magnitude compared to the observed ratio of $\approx 10^{-10}$. This is because a nonvanishing absorptive part involving C and CP violations arises only in a very high order of perturbation theory for the minimal $SU(5)$ model. But this difficulty can be circumvented by extending the gauge and/or the Higgs sector. Before ending this section, I should mention that, meanwhile, two new developments have taken place through the interplay of particle physics and cosmology. These bring new angles which must be taken into account before the issue of baryogenesis can be settled.

1. <u>Inflation</u>:

First, is the idea of inflation [63] which simultaneously resolves two major puzzles in cosmology. These are:

(i) <u>The Flatness Problem</u>: Question: why is the mass-density of the early universe ρ, extrapolated from its observed value today to the past, so incredibly close to its critical value ρ_c that is just suficient to eventually stop the expansion of the universe? [Noting that the ratio $\Omega \equiv \rho/\rho_c$ is known to lie in the range [.1 to 2] today and that $\Omega - 1$ varies as $t(t^{2/3})$ during radiation (matter) dominated era, one obtains $|\Omega - 1| < 10^{-16}$ at $t =$ 1 sec, and $|\Omega - 1| < 10^{-49}$ at $t \approx 10^{-35}$ sec, which is the grand unification time-scale. (See e.g. [62].) Standard cosmology simply assumes this feature as part of the initial conditions, but does not provide an explanation for it.

(ii) <u>The Horizon Problem</u>: Question: why is the universe so uniform - i.e. homogeneous and isotropic – at large scales, as revealed especially by the cosmic background radiation? Isotropy of the radiation extends to better than one part in 10^4 across the sky subtending angles from 10″ to 180°. Assuming standard cosmology, however, this involves regions which are causally disconnected, at the time of recombination, when photons last scattered against matter. The question arises: why is the radiation still so isotropic involving regions which could not have had any interaction with each other? Standard cosmology must introduce this feature of homogeneity also as part of initial conditions, but can not explain it.

The basic assumption of inflation [63], which neatly resolves both the puzzles and many more, is that the universe, while undergoing a phase transition was somehow hung up, during a short period of time (Δt) in its early history, in a false vacuum, which dominates the energy density of the universe during this epoch. Following this short epoch, it, of course, did make the transition to the true vacuum - i.e. the state of lowest energy. During the short interval Δt since the energy-density of the false vacuum remained essentially constant, the universe, following Einstein's equations, expanded exponentially rather than by a power-law. The expansion factor given by $Z = e^{H\cdot\Delta t}$, where H is the relevant Hubble parameter, can thus be enormous $\gtrsim 10^{30}$, say, even for a relatively short time interval $\Delta t \gtrsim 70 H^{-1}$ (see below for relevant values of H^{-1}). Such a large expansion factor, as we will see below, resolves both puzzles.

First, it may be noted that since the energy density ρ_f of the false vacuum remains constant during the short interval Δt, while the volume of the universe increases by the enormous factor $Z^3 \gtrsim 10^{90}$ (say), the total matter-energy of the universe (i.e. all energy other than gravitational) increases by the same factor $\gtrsim 10^{90}$ during the short inflationary epoch. It can be argued that this does not, however, violate any conservation principle, if one adequately takes into account the exchange of energy between matter and gravitational fields [64]. The reason is that the pressure p that enters into the energy-momentum tensor $T_{\mu\nu}$ for the false vacuum is negative and is equal to $-\rho_f$. [From the Lorentz invariance of the false vacuum, its energy-momentum tensor must have the form $T_{\mu\nu} = \rho_f g_{\mu\nu} = \rho_f diag(1,-1,-1,-1)$ which implies that $p = -\rho_f$.] In short, the matter-energy of the system increases because work is done by the false vacuum pushing the system outward against the *negative* pressure [64].

The evoluation of the scale factor $R(t)$ that gives the size of the universe is governed by the two Einstein's equations:

$$\ddot{R}/R = -\frac{4\pi G}{3}(\rho + 3p) \tag{48}$$

$$H^2 \equiv (\dot{R}/R)^2 = \frac{8\pi G\rho}{3} - k/R^2 \tag{49}$$

We see that, for negative pressure $p = -\rho_f$, the effective source of gravity $(\rho + 3p) = -2\rho_f$ is negative. So, the RHS of Eq. (48) is positive. *Thus gravity, for a false vacuum, acts as a repulsive force and thereby accelerates the expansion rather than slowing it down.* It is this negative pressure of the false vacuum which drives inflation and, as we saw, leads to the enormous *increase* of matter-energy. If true, it would mean that essentially all of matter-energy that we see today has arisen during the inflationary epoch. This is what Guth rightly calls a "free lunch".

To be specific, let me present the so-called *new* inflationary picture [65] which retains all the successes of the original inflationary model proposed by Guth [63], but avoids some of its technical flows (see e.g. Ref. 62). Let us furthermore assume that inflation is associated with the phase transition corresponding to the spontaneous breaking of a grand unification symmetry, so that one would expect a critical temperature $T_c \sim 10^{14}$ GeV and thus the age of the universe just preceding the inflationary transition to be $t \sim H^{-1} \sim (g_*^{1/2} T_c^2/M_{pl})^{-1} \sim 10^{-34}$sec. One way to achieve a large expansion factor $Z \gtrsim 10^{30}$ by the exponential growth is to assume that the potential energy density of the universe V lying in the false vacuum at temperatures $T < T_c$, viewed as a function of the relevant Higgs scalar field ϕ, is approximately flat having a constant positive val-$V_o \sim (10^{14} GeV)^4$ over a *broad range* in ϕ spanning from $\phi \approx 0$ to $\phi = \phi_e$ (say) [65]. As ϕ exceeds ϕ_e, V drops rapidly towards its global minimum representing the true vacuum which is attained at $\phi = \sigma \approx 10^{14}$ GeV (say). The value of V at this global minimum is commonly assumed to be zero or nearly zero to account for the zero or vanishingly small cosmological constant today (such a vanishing value of the cosmological constant, of course, still remains, however, as an unexplained mystery [66]).

If the range $\Delta\phi$ in ϕ over which V is nearly flat is broad enough [67] compared to H, the time Δt for ϕ to roll down the flat region would be long enough compared to the expansion time scale $H^{-1} \sim 10^{-34}$ sec, and this would lead to sufficient inflation. For example, a roll-over time $\Delta t \gtrsim 70 H^{-1} \approx 10^{-32}$ sec would lead to an expansion factor $Z = e^{H\cdot\Delta t} \gtrsim 10^{30}$, as desired. In practice, Z may in fact far exceed 10^{30}.

With an expansion factor $Z \gtrsim 10^{30}$ (say), a very tiny domain, representing a single fluctuation region and having a pre-inflation size of order only 10^{-24} cm, *which is thus perfectly smooth and causally coherent*, would grow, during the short interval $\Delta t \approx 10^{-32}$ s, to a size $\gtrsim 3 \times 10^6$ cm. This inflated region would then grow adiabatically, following standard cosmology, to more than account for the present size of the observed universe. In this picture, the observed universe is thus only an infinitesimal fraction of the full cosmos, all of which evolved out of a *single domain or fluctuation region* in the process of phase transition in the early universe. By the same token, the size of the *observed universe* in the inflationary picture at times before the grand unified phase transition was clearly much smaller than what it would be in the standard scenario (by almost the factor Z). As a result, it lay well within a single fluctuation region maintaining thermal contact. This then naturally accounts for the large-scale isotropy and homogeneity of the observed universe today and solves the horizon problem.

In this picture, there would, of course, be an "infinity" of other domains which must have evolved into their own universes and which are causally disconnected from each other now. Our own universe is just one among these.

Using eq. (49) and the fact that the critical density is given by $\rho_c = 3H^2/8\pi G$, one obtains: $(\Omega - 1) = k/H^2R^2$, where $\Omega = \rho/\rho_c$. We see that if R increases exponentially by $e^{H\cdot\Delta t} \gtrsim 10^{30}$ during inflation, $(\Omega - 1)$, being proportional to $1/R^2$, approaches zero very rapidly during inflation regardless of its initial value. This explains naturally the flatness problem -i.e. why Ω is so incredibly close to one. This is indeed a prediction of the inflationary picture.

One truly fascinating aspect of the inflationary scenario is the creation of an enormous amount of matter and radiation out of almost nothing. As the domain reprsenting a single fluctuation region expands exponentially, it supercools rapidly to essentially zero temperature, while, (as mentioned before) its net matter-energy increases by $Z^3 \gtrsim 10^{90}$ (say). This is all potential energy. Following the inflationary epoch, the relevant scalar field ϕ approaches the position of true minimum $\phi = \sigma$ and the potential energy drops rapidly to its minimum value which is assumed to be zero. The scalar field oscillates around the position of this minimum and, by its decay, creates new particles. In short, the enormous potential energy stored in the false vacuum during inflation is converted into the creation of new particles (matter and radiation) possessing kinetic energy. As a result, the universe reheats (non-adiabatically) typically to temperatures $T \sim T_c \sim 10^{14}$ GeV or less (see below). At this point, it rejoins standard cosmology and expands adiabatically. In the proce s of reheating, the entropy of the universe increases by the enormous factor $Z^3 \gtrsim 10^{90}$. *Thus, in the inflationary picture, essentially all matter, energy and entropy which we see today is created by the expansion and subsequent decay of the false vacuum.*

Inflationary scenario has the additional advantage that it dilutes unwanted relics such as superheavy 't Hooft-P lyakov magnetic monopoles ($\sim 10^{16}$ GeV) [68], which would be produced during the sp taneous breaking of grand unified symmetries, provided that such a breaking takes pl ce either before or during inflation. Without such a dilution, grand unified theories w ld be in conflict with cosmology.

To summarise, inflat onary picture is an attractive idea in that it resolves a number of major puzzles. Thus it seems compelling that one way or another standard cosmology must be supplemented by a short period of non-adiabatic inflationary expansion in the early universe. It should be stated, however, that a compelling derivation of the full inflationary picture including its initial phase from an underlying theory of particle physics is still missing. To achieve a sufficiently flat potential near the origin $\phi \approx 0$ for the new inflationary model [65], the general tendency is to assume a radiatively generated Coleman-Weinberg effective potential with an assumed zero or vanishingly small effective mass and also a tiny quartic coupling for ϕ during the inflationary regime [67]. To achieve small enough density perturbations $\delta\rho/\rho \sim 10^{-4}$, which would inevitably arise through quantum effects, one must further assume that the quartic coupling parameter λ multiplying $\ell n\phi/\sigma$ in the Coleman-Weinberg potential should in fact be truly small

($\lambda \sim 10^{-14}$) (see Ref. 62 for a discussion of this and other phenomenological constraints which must be satisfied by an acceptable inflationary model).

It is thus imperative that an underlying theory should naturally yield such miniscule quartic coupling and mass parameter for the relevant Higgs scalar; otherwise one faces the problem of unnatural fine tuning. One idea that has been considered in this regard is that the inflaton field is a weakly self-coupled gauge-singlet so that its quartic coupling λ does not receive radiative contributions from gauge interactions, which would be too large. In this case, the inflaton field can not be directly responsible for breaking the grand unification symmetry, though indirectly, through its coupling to the relevant Higgs scalar (like the 24 of $SU(5)$), it may still induce such a breaking. A model of this type has been proposed in Ref. 69. It still appears to have, however, much arbitrariness in the choice of fields and their coupling parameters, some of which have to be chosen to be as small as $10^{-5} - 10^{-6}$. Thus the fine tuning problem is not fully avoided.

A more attractive framework which ultimately may succeed in yielding naturally a fully satisfactory model of inflation is that of supersymmetry in its local form - i.e. supergravity. This is because, first of all, supersymmetry has the virtue of preserving small parameters against radiative corrections. For motivations and some early suggestions in this respect see e.g. Ref. 70 and 71. The second virtue is that supe ravity models, may in some cases even *explain* the smallness of the relevant parameters see remarks below), provided they are forbidden in the limit of supersymmetry. Ulti ately, the question would be whether any such model fits naturally within a more com lete picture pertaining to an underlying theory of particle physics that is capable of dressing a variety of problems in a unified manner and not just inflation.

I believe that in the context of local supersymmetry, the inflaton field may well be a *gauge non-singlet* and can in fact be involved in the spontaneous breaking of a bigger symmetry, while also inducing inflation. Its effective mass and quartic coupling parameters (in the false vacuum) are small, because they are forbidden in the limit of global $SUSY$. In this case, $SUSY$ breaking may be induced by gravity and the mass and quartic coupling parameters in question may in turn be induced radiatively due to $SUSY$ breaking. These parameters would then be calculable, following Coleman and Weinberg, and would be naturally small. A model of this type has recently been shown to emerge [72] from within a locally supersymmetric preonic theory which possesses many attractive features (see comments in the last section). Within this picture, the inflaton field is associated with a composite Δ_R having the same transformation property as the "elementary" Δ_R that was introduced in section 4.3 to break spontaneously $SU(2)_R \times SU(4)^C$ to $U(1)_Y \times SU(3)^C$. The quartic coupling of this composite field, which is shown to vanish in the limit of SUSY, is derived in this case to be $\approx 10^{-14\pm1}$ and the mass-parameter is shown to be vanishingly small, just as desired. The problem of small numbers, needed to complement the new inflationary picture, is thus solved completely and naturally within this model. It is not clear in this picture, as in all $SUSY$ models, however, why the universe passes through the false vacuum at $\phi \approx 0$ instead of passing smoothly into the true vacuum. [See Ref. 62 for a discussion of possible remedies of this problem.]

Supergravity models, in general, also face the so-called "gravitino-problem". Gravitinos are weakly interacting and thus very long-lived. They could adversely affect the energy-density of the universe, depending upon their mass [73]. In the inflationary context, this problem is alleviated because their primordial abundance is greatly diluted by inflation. They are, however, reproduced during reheating following inflation, with a density which is proportional to the reheat temperature. It turns out [74] that their decays (assuming that their mass is of order 100 GeV to 1 TeV, as suggested by the simplest class of supergravity models) would noticeably disturb primordial nucleosynthesis, unless the reheat temperature is less than about 10^9 GeV. Such a reheat temperature is in fact quite likely with an inflaton field like Δ_R having a VEV and, therefore, a critical temperature for initiating inflation around an intermediate scale 10^{11} GeV, as in Ref. 72 rather than the grand unification scale of $10^{14} - 10^{16}$ GeV.

I now return to the implications of inflation on baryogenesis. It is clear that even if a baryon-excess existed just prior to inflation (for whatever reasons) it would be much diluted during inflation. Thus, baryogenesis must occur after the last inflation in order that one may account for the baryon-asymmetry of the universe observed today. There are a few alternative possibilities in this regard especially because of the recently proposed ideas on electroweak baryogenesis (see below). If one ignores the electroweak phenomena for a moment, however, and if baryon-excess is to be produced by the out-of-equilibrium decays of either the superheavy gauge basons of mass 10^{14} GeV or, more promisingly, the decays of the scalar bosons of mass 10^{12} GeV (say), which occur in grand unified theories, it is necessary that the universe must reheat to temperatures $\gtrsim 10^{11}$ GeV, so that a sufficient number of such heavy bosons may be produced during reheating.

Since the reheat temperature is expected to be $T_{RH} \simeq [\Gamma_\phi M_{Pl}]^{1/2}$, where Γ_ϕ is the width of the decaying inflaton one needs $\Gamma_\phi > 10^3$ GeV, so that $T_{RH} \gtrsim 10^{11}$ GeV. This in turn needs the effective coupling α_ϕ of ϕ to exceed about 10^{-8}. This is not easy to satisfy if ϕ is a weakly coupled gauge singlet. But, such an effective coupling arises naturally if ϕ is a gauge non-singlet.

One intriguing possibility is that baryon asymmetry is produced by the decay of the inflaton field itself. In this case, the reheat temperature can be considerably lower than that mentioned above. This general idea has been considered by several authors with alternative possibilities for the inflaton. As one possibility, if one way or another, there are sufficient number of Δ_R's and/or equivalently ν_R's following reheating, their out of equilibrium decays utilising $< \Delta_R > \neq 0$ and/or the Majorana mass of ν_R's could generate lepton-excess. It has been suggested that leptogenesis can in turn lead to baryogenesis utilising electroweak $B+L$ violation [75] (see discussions below). As regards ν_R-decays to light fermions, they would be governed by effective Yukawa couplings which

are ultimately related to the familiar Dirac masses of the known fermions either directly, as in Ref. 76, or via a see-saw mechanism, generating masses for all known fermions, as in Ref. 77. Such decays can be out-of-equilibrium in the first case, especially for the ν_R that goes with the lightest family. This could thus give rise to leptogenesis, which, in conjunction with electroweak physics, can yield baryogenesis.

One would of course, still need C and CP violations to be relevant in the decay processes so as to realise lepton asymmetry. If C and CP violations arise spontaneously through phases - e.g. in a set of condensates - arising during the proccess of inflation, such violations would be present in the decays of Δ_R and/or ν_R as well. At the same time, the so-called domain wall problem associated with spontaneous violation of discrete symmetries would disappear due to inflation (this is yet another advantage of inflation).

While each of the variants mentioned above possesses some merits, the details in terms of numbers combining the scenarios for inflation and baryogenesis still need to be worked out to check the viability in each case.

2. Electroweak $B+L$ Violation

Last but not least, there is yet another factor that is relevant to the baryon asymmetry of the universe which we see today. As alluded to before, even electroweak physics, governed by the $SU(2)_L \times U(1)$ gauge symmetry, leads to baryon violation through quantum non-perturbative effects. This is because the triangle graph involving the left chiral fermions in the loop, coupled to the currents of $B+L, W_L^a$ and W_L^b at the three vertices of the triangle, possesses Adler-Bell-Jackiw anomaly. As a result, the $B+L$-current is not conserved [78] and one obtains:

$$\partial_\mu J\mu^{B+L} = 2n_f[\frac{g^2}{32\pi^2}F_a^{\mu\nu}\tilde{F}_{a\mu\nu}] \tag{50}$$

Here n_f denotes the number of chiral families and $F^a_{\mu\nu}(\tilde{F}_{a\mu\nu})$ denote the field strengths (dual field strengths) of the $SU(2)_L$ gauge symmetry. There is no anomaly, however, in the $(B-L)$ current. Thus $(B-L)$ is conserved by electroweak physics.

$B+L$ is violated due to transitions of the $SU(2)_L$ gauge field configurations from one vacuum to another which differ in baryon number by $2n_f$ units. At zero temperature, the amplitude for quantum tunnelling between such neighboring vacua differing in $B+L$ by $2n_f$ is mediated by the $SU(2)_L$-instantons. This is exponentially suppressed by the familiar non-perturbative factor $e^{-2\pi/\alpha_W} \sim e^{-180}$, which is, of course, negligible today as well as in the early universe, as was first noted by 't Hooft [78].

At higher temperatures, the height of the barrier separating two vacua is given by $E_o(T) \sim m_W(T)/\alpha_W$ and the amplitude for crossing the barrier is proportional to $e^{-E_o/T}$. Kuzmin, Rubakov and Shaposhnikov [79] pointed out that at temperatures $T \gg m_W$, in particular with $T \gtrsim 1$ TeV, baryon violation proceeds without any exponential damping because, as T approaches the critical temperature T_c for $SU(2)\times U(1)$-breaking, $m_W(T)$ and thus the classical energy barrier E_s approach zero. As a result, $(B+L)$-violating but $(B-L)$-conserving processes, induced by electroweak physics, are found to be in thermal equilibrium ($\Gamma_{B+L} > \Gamma_{exp}$) for a wide range of temperature 200 GeV $\lesssim T \lesssim 10^{12}$ GeV [80]. This in turn implies that any net $B+L$ created in earlier epochs will be completely erased by electroweak physics during the temperature interval 200-10^{12} GeV. Noting that $\Delta B = \frac{1}{2}\Delta(B+L) + \frac{1}{2}\Delta(B-L)$, it follows that the net baryon number of the universe that is seen today, if it arose entirely through non-electroweak physics, must be given by the net $(B-L)/2$ that might have been produced either near the grand unification scale or some intermediate scale. This clearly requires non-conservation of $B-L$. Such a possibility can not, therefore, arise within the minimal $SU(5)$ because it conserves $B-L$. It can, however, arise within symmetries containing $SU(4)$-color which gauge $(B-L)$ such as $SO(10)$ with a two-scale breaking. These violate $B-L$ spontaneously (see e.g., Ref. 47 and 48 and discussions in Sec. 4.11).

As I mentioned earlier, the out-of-equilibrium lepton-violating decays of ν_R's into antileptons, could generate lepton asymmetry. Such lepton asymmetry, combined with strong $B+L$ violation due to electroweak physics, could yield the desired baryon asymmetry of the universe. A specific model of this kind involving ν_R-decays, but untied to inflation, has recently been shown to yield the desired magnitude for baryon asymmetry [76]. It is important that a satisfactory inflationary scenario should allow sufficient number of Δ_R's and/or ν_R's to be produced, following inflation and reheating, for such a mechanism to work. Note that models of this kind fruitfully utilize both electroweak physics as well as the physics of higher symmetries such as $SU(2)_L\times SU(2)_R\times SU(4)^C$, which makes the introduction of ν_R and Δ_R compelling. These models, coupled with supersymmetry (see e.g. Ref. 72), seem to me to have the potential for providing natural and realistic solutions to both inflation and baryogenesis.

One additional possibility is that electroweak B+L violation may allow baryogenesis to occur entirely through electroweak physics at temperatures below 200 GeV. Such a possibility was suggested in Ref. 81 and models of this kind have been proposed in Ref. 82. If they succeed in being realistic (this is certainly not clear at present), they would permit one to decouple scenarios for inflation and reheating from those of baryogenesis. This may, apriori, be an advantage.

One way or another, it is clear that at least inflation and very likely also baryogenesis are intimately related to the physics at short distances which goes well beyond that of the standard model. This then brings me to the concluding section.

V. GRAND UNIFICATION IN A BROADER PERSPECTIVE: A POINT OF VIEW AND A SUMMARY

Grand unification is an attractive concept because it is the only known consistent quantum field theoretical framework we have which explains:

- The quantisation of electric charge.
- The existence of quarks and leptons with $Q(e^-) = -Q(p)$.
- The existence of the strong, the electromagnetic and the weak forces with $g_3 \gg g_2 \gg g_1$ at low energies, but with $g_3 = g_2 = g_1$ at high energies.

The last two features provide a unification of basic matter and of its forces. *In addition, grand unification provides a compelling reason for the existence of fundamental new physics at a superheavy scale (or scales) lying between the Planck and the electroweak scales.* Such short-distance physics is almost certainly necessary to understand (a) the tiny masses of the neutrinos (if they are not zero) and very likely also (b) inflation, (c) baryogenesis and (d) the origin of dark matter.

For these reasons, I believe that grand unification *as a concept* is, very likely, a step in the right direction. It is thus likely to be contained either *explicitly* or *implicitly* within a more fundamental theory. The standard model emerges effectively from this theory at some scale, lying between the Planck and the electroweak scales, preserving all the *constraints* of grand unification. These constraints relate (a) the gauge coupling constants like g_3, g_2 and g_1 to each other and also (b) the quarks to the leptons. Furthermore, as mentioned above, they guarantee quantisation of electric charge and imply non-conservations of B and/or L. *In a broader sense, therefore, it is these inter-relationships between the fundamental particles and their forces and the associated consequences of quantization of electric charge and the non-conservations of B and/or L which grand unification brings to physics. I believe that it is precisely these aspects of grand unification which should last one way or another regardless of the manner in which they are realized.* Some important details (consequences) including the value of $\sin^2\theta_W$ and the proton life-time could, however, vary widely depending upon the nature of the fundamental theory and the manner in which grand unification or at least the constraints of grand unification, mentioned above, emerge effectively from it at low energies.

For this reason, non-observation of proton-decay (so far) and increased value of $\sin^2\theta_W$, both of which contradict minimal $SU(5)$, can hardly be regarded as evidences against the central concept of grand unification. As discussed in the text, either supersymmetry, or new physics at an intermediate scale $\sim 10^{11}$ GeV, or both, lead to increased τ_p and $\sin^2\theta_W$, in accord with observation. Furthermore, there are altogether new possibilities which go beyond conventional grand unification and are likely to remove its short comings (see below). These could also naturally differ from minimal $SU(5)$ as regards the predictions for τ_p and $\sin^2\theta_W$. It remains to be seen which of these (if any) might be relevant to nature.

Having said this in favor of grand unification as a concept, let me mention that I do not, however, believe that it will survive, even nearly, in the traditional sense, with the effective symmetries and field-contents of the types proposed in the 1970's, with or without supersymmetry [4,5,26,54,58]. This is because the traditional framework, while having the successes mentioned above, do not have the natural ingredients to account for (a) the origin of diverse mass scales from M_{Planck} to m_ν, (b) the origin of families with hierarchical masses and mixings and (c) the co-existence of gravity with the other three forces. Furthermore, as alluded to before in sec. 4.11, traditional grand unification possesses a host of *arbitrary parameters* in the form of the masses and the quartic and the Yukawa couplings of the Higgs scalars. It also permits much arbitrariness in the choice of the fermion and the scalar multiplets. A fundamental theory should, however, be devoid of such arbitrariness. Thus it seems clear that grand unification as a concept must have its origin as part of a bigger idea which is capable of removing this arbitrariness.

Origin From Superstrings

The one idea which has such a scope is that of superstrings [11]. As mentioned in sec. III, superstring theories appear to be most promising because they almost certainly yield a well-behaved theory of quantum gravity and seem capable of unifying all the forces of nature including gravity. In this sense, they comprise the central concept of grand unification and go well beyond.

The formulation of superstring theories, despite their promise, is not based, however, on a fundamental principle as yet, unlike the case of Einstein's general theory of relativity which is based on the principle of equivalence and the case of Yang-Mills theories based on the principle of local gauge invariance. In this sense, it is very likely that some fundamental ingredient is still missing from string theories which could provide guidance in the matter of choice at a fundamental level.

Talking of the matter of choice, superstring theories at present lack uniqueness in their formulations and especially in their solutions. There is first of all the choice between three types of superstring theories which arise in the critical dimension $d = 10$ [83]: (i) Type I with unoriented open and closed strings, based on the gauge group $SO(32)$; (ii) Types IIA and IIB with only oriented closed strings and (iii) the heterotic, with only oriented closed strings, based on either $SO(32)$ or $E_8 \times E_8'$. Of these, the heterotic $E_8 \times E_8'$ theory (see Gross et al. and Candelas et al. in Ref. 11) seems to be the only one which holds some promise in describing correctly the quark-lepton physics at low energies in $d = 4$, though even this is far from clear at the moment (see remarks below). If it is the right choice, one would like it to be uniquely selected out over the other

possible choices either by quantum effects (which may make all but one inconsistent) or some principle. This, however, is not the case at present.

The second choice which needs to be made arises due to the non-uniqueness of the solutions. Even if one accepts the heterotic $E_8 \times E_8'$ theory to be the right choice in $d = 10$, and thus the fundamental theory is free from adjustable dimensionless parameters as desired, there is a *plethora of classically allowed solutions* for the four-dimensional ground state or the vacuum, corresponding to the compactification of the extra six spatial dimensions onto, for example a Calabi-Yau manifold [Candelas et al. Ref. 11] or orbifold [84]. The different Calabi-Yau spaces are determined by different parameters (moduli). In as much as there is an enormous vacuum degeneracy corresponding to a lack of preference for one solution over the other, one is faced with making an *arbitrary choice* between different sets of parameters.

To add to the problem of non-uniqueness in formulation and/or solutions, a host of superstring theories have been constructed directly in four space-time dimensions [85], some of which are of interest from a phenomenological standpoint. These are apriori consistent just like the compactified versions of the ten-dimensional models in that they also satisfy local reparametrization invariance, world sheet supersymmetry, modular invariance and conformal invariance [83]. It may be possible to regard some of the four-dimensional constructions as compactifications of the ten-dimensional string models. But the construction directly in four-dimensions turns out to provide a much simpler method to obtain these solutions. Leaving aside the question of the parentage and the interconnections between these different solutions, the four-dimensional constructions offer certain new features and have greatly extended the space of consistent string models in four dimensions.

The primary task of string theory at present is, therefore, not necessarily to find new solutions but to understand its dynamics — especially its non-perturbative aspects — well enough, hopefully through the development of a suitable quantum field theory of strings, so that one may understand how the right vacuum is picked (perhaps) uniquely and, equally important, how supersymmetry breaks at a scale much below the Planck scale. One difficulty in this regard is that one really does not have the lagrangian for the string theory. What one has are the rules for constructing the S-matrix elements in perturbation theory. On the other hand, string perturbation theory, when summed to all orders, in powers of the coupling constant, is found to diverge badly [86]. This is proved for bosonic strings but believed to be true for all string theories. Optimistically, this is a desirable feature, because if string perturbation theory, summed to all orders, did make sense, one would not understand how to (a) break supersymmetry which is respected in each order, (b) generate a mass for the dilaton and (c) choose the unique vacuum. *Thus, it is clear that non-perturbative effects are essential to accomplish these tasks and make sense out of string theories.* It remains to be seen how and how soon one may gain insight into these effects.

In the meantime, given the vastness of the space of solutions without a reliable criterion to choose between them, the current status of "superstring phenomenology" has been reduced to the question: *is there any solution among the Calabi-Yau spaces, orbifolds, four-dimensional constructions or any other which resembles or nearly resembles the known world?* If such a solution (or solutions) could be found, the justification for its selection on dynamical grounds may be sought later. It seems fair to say, however, that an answer in the affirmative even to this question, at least within the conventional approach comonly pursued so far (see below), is still missing.

This brings me to remark on the phenomenological prospects of two alternative routes by which superstring theories may make connection with the low energy world. The ultimate justification for either one will, of course, have to wait until we gain insight on the question of the choice of the right vacuum.

1. The Conventional Route of Elementary Quarks and Leptons:

It is commonly believed that the appropriate superstring theory, *e.g.* the compactified ten-dimensional $E_8 \times E_8'$ heterotic theory (see Gross *et al.* and Candelas *et al.* in Ref. 11) or one of the so-called four-dimensional superstring theories [85], yields, below the Planck scale, an effective symmetry containing the $SU(2)_L \times U(1)_Y \times SU(3)^C$ gauge symmetry and $N = 1$ local supersymmetry. Furthermore, it is customary to assume that the quarks, leptons and the Higgs bosons are elementary. Thus almost all phenomenological studies of superstring theories have focused on identifying at least a subset of the massless fields emerging from the superstring theories with the known quarks, leptons and the Higgs bosons.

Enthusiasm in this regard was triggered by the pioneering work of Candelas *et al.* [11] who showed that the compactification of the ten-dimensional heterotic $E_8 \times E_8'$ theory onto a Calabi-Yau manifold K naturally leads to 27's, $\overline{27}$'s and singlets of a broken E_8 belonging to one of the E_8's and that the number of *chiral* families (i.e. $n_{27} - n_{\overline{27}}$) is given by the *topology of the compact manifold* K — *i.e.* by one-half of the magnitude of the Euler characteristic of K. It may be recalled that E_6 is a desirable grand unification group containing $SO(10)$ and $SU(2)_L \times SU(2)_R \times SU(4)^C$ and that 27 is the desired representation containing $16 + 10 + 1$ of $SO(10)$.

An additional desirable feature is that at least a number of Yukawa couplings between three 27's which govern the masses and the mixings of the quarks and the leptons can be calculated at the string tree-level on the basis of the topology *and* the complex structure of the compact manifold K [87]. This is true even after one allows for all the independent unknown field normalisations. Although these calculations are apriori subject to drastic changes due to higher order string-loop corrections, some non-renormalisation theorems have been proven which show that at least perturbatively, to all orders in the corresponding σ-model perturbation theory [88] and also in the string-loop expansion [89], the tree-level results on the Yukawa couplings are not renormalised. Furthermore, it has

been shown [90] that inclusion of non-perturbative effects as manifested by world-sheet instantons [91] do not renormalise the 27^3 couplings although they do renormalise the $\overline{27}^3$ couplings [91]. It remains to be seen whether other non-perturbative effects, *e.g.* those due to space-time instantons, which are known to invalidate the non-renormalisation theorem [92], would alter the 27^3 couplings.

Finally, Calabi-Yau manifold of a rather special nature has been constructed by Tian and Yau [93] possessing three generations of massless chiral quarks and leptons in 27's of E_6, and a four-dimensional gauge symmetry, obtained by the "flux breaking" of E_6 [94], which is either $[SU(3)]^3$ or $SU(6) \times U(1) \subset E_6$. The former is phenomenologically favored. The phenomenology of this model has been studied by a number of authors [95]. For a comprehensive review of some of these works, see Ref. [96].

Some generic problems characterising all these studies are the following. The theory after compactification near the Planck scale must pass through several stages of symmetry breaking at some *intermediate scales* which must lie between the Planck and the electroweak scales. These involve the breaking of supersymmetry and also the Planck-scale gauge symmetry (such as $[SU(3)]^3$) to $SU(2)_L \times U(1)_Y \times SU(3)^C$. The breaking of $[SU(3)]^3$ is assumed to take place through the VEVs of the scalar partners of ν_R and the $SU(2)_L \times SU(2)_R$-singlet neutrino N in the 27 of E_6. Such VEVs would be generated if the corresponding scalar $(mass)^2$-terms, induced by SUSY-breaking and evolved through renormalization group, become negative. Thus, which fields acquire VEV's and at what scales, depend on the nature of SUSY-breaking and also on the normalized Yukawa couplings. Since one does not have a handle from the string theory to derive either the nature of supersymmetry breaking or all the normalized Yukawa couplings, one is obliged in practice to proceed as best as possible by making certain *adhoc assumptions* about supersymmetry and intermediate scale gauge symmetry-breakings to see whether any superstring-generated solution may conceivably make connection with the real world in a desirable way.

Supersymmetry breaking was initially attributed to the formation of a single gaugino-condensate [97,98] in the hidden E_8'-sector, which, together with the VEV of an antisymmetric tensor field was assumed to maintain zero vacuum energy [98]. This idea, pursued in detail, seems, however, to be inconsistent in many respects [99], *e.g.* the relevant VEV is found to be quantised in Planck units and SUSY-breaking occurs at the Planck scale. Furthermore, with a single gaugino condensate in the hidden sector, the theory is found to be strongly coupled (*i.e.* the gauge couplings are large) at the compactification scale.

It has subsequently been suggested that the hidden sector may contain a product of simple groups (rather than a single E_8'), which in general is allowed, and that correspondingly several gaugino condensates may form in the hidden sector at differing scales [100]. It has also been noted that three ingredients are in fact necessary [101] to implement realistic supersymmetry breaking through the hidden sector: (i) the formation of at least two gaugino condensates, (ii) a stage of spontaneous breaking of the gauge symmetry (*e.g.* $[SU(3)]^3$) at an intermediate scale below the compactification scale and (iii) moduli dependent superpotential due to non-perturbative effects. All of these ingredients may well arise within the string theory. But the consistency of any such scheme in detail is not at all clear. In particular, the answers to the questions of whether the gaugino-condensation scales and the intermediate gauge symmetry-breaking scale(s) have just the right values and whether supersymmetry breaking in the hidden sector transmits efficiently through gravitational interaction into the observable sector so that ultimately the Higgsinos, the squarks, the sleptons, the gluinos and the winos have masses in the desirable range of order 100 GeV-1 TeV are far from clear.

Apart from the problem of supersymmetry breaking, which seems rather generic, there are some practical problems which also arise in general in superstring theories and must be circumvented before any particular solution may be regarded as viable. These include obtaining:

(i) sufficiently stable proton;

(ii) sufficiently light neutrinos;

(iii) consistency with renormalization group analysis leading to the right value for $\sin^2\theta_W$ and

(iv) the right pattern and the values for the masses and the mixings of the quarks and the leptons.

For a review of the status of three-generation models, with regard to these issues, see Ref. [96,102] and references therein. Briefly speaking, the neutrinos can be kept light by a see-saw mechanism involving the $SU(2)_L$-doublet neutrinos ν_L's and the singlet N with a heavy Majorana mass $\sim M_I^2/M_{Pl}$, where $M_I \sim 10^{16}$ GeV; $\sin^2\theta_W$ is on the borderline of being reasonable ($\sim$ .23) if $M_I \gtrsim 10^{15} \sim 10^{16}$ GeV. As regards proton-stability, while dimension 4 baryon-violating terms are naturally absent in the model due to discrete symmetries, dimension 5 and dimension 6 operators would be suppressed sufficiently provided the exotic D-quarks have masses $\gtrsim 10^{16}$ GeV. This in turn requires a delicate adjustment of VEVs of the intermediate scale fields [96], which, without a reason, is not really a solution. Last but not least, the model hardly leads to any prediction for the quark-lepton masses because the relevant Yukawa couplings involve unknown field normalisations. If one chooses alternative symmetry breaking patterns (Higgs fields) at the intermediate scale [103], one usually encounters explicit conflict with phenomenology such as unrealistic lepton masses and blowing up of $SU(3)^C$ coupling at around 10^{14} GeV. Thus, it seems fair to say that three-generation Calabi-Yau models are either inconsistent or need adhoc assumptions and delicate adjustment of parameters to be consistent with low-energy phenomenology. At any rate, they do not explain the

hierarchical quark-lepton masses and mixings and the origin of the diverse mass scales. Also, they do not seem to have some crucial testable predictions at low energies.

Leaving aside the three-generation Calabi-Yau models, phenomenology of several other superstring-derived models have also been pursued with varying degree of success. Notable among these are two four-dimensional superstring models [85], leading to the so-called flipped $SU(5) \times U(1)'$-model [104] and the $SU(2)_L \times SU(2)_R \times SU(4)^C$ model [4,105]. The former has been studied at some length [106]. It flips the role of u^c and d^c compared to the original $SU(5)$-model (see sec. 4.7) and puts u^c in $\bar{5}$ and d^c in 10 of $SU(5)$. The right-handed neutrino ν^c is assigned to 10, while the positron e^c is introduced as a singlet of $SU(5)$. Electric charge gets contributions from both $SU(5)$ and $U(1)'$. Unlike the original $SU(5)$, which needs an adjoint Higgs to break to the standard model, the flipped $SU(5) \times U(1)'$ is broken to the standard model by a $10_H + \overline{10}_H$ pair of Higgs utilizing the VEV of scalar ν_H^C $\left(\bar{\nu}_{\bar{H}}^C\right)$.

The model [104] is derived in the free fermionic formulation [85] of four-dimensional strings and is based on the gauge group $[SU(5)\times U(1)'\times U(1)^4]\times[SO(10)\times SU(4)]_{hidden}$. The massless spectrum of the model consists of four families plus one anti-family (or equivalently three ordinary families), one pair of Higgs $10_H + \overline{10}_H$, 4 pairs of electroweak Higgses $5_h + \bar{5}_h$, a bunch of $SU(5) \times U(1)'$-singlets, several $SU(5)$-singlets carrying fractional electric charges $\pm\frac{1}{2}$ and matter-fields with non-trivial representation in the hidden $SO(10) \times SU(4)$-sector. The fractionally charged particles get confined by the hidden sector forces.

Recent studies seem to reveal a number of inconsistencies of this model. First $\sin^2\theta_W$ and/or α_3 turn out to be unacceptly large unless one modifies the model by introducing extra matter multiplets with masses $\lesssim 10^{13}$ GeV and with couplings which are vector-like [107]. Second, allowing for non-renormalizable interactions even up to the eighth order and for the breaking of the hidden sector $SO(10) \times SU(4)$ to $SO(7) \times SO(5)$ (which is shown to be advantageous), the model turns out to possess two massless Higgs isodoublets and one pair of massless colored triplets [108]. Unless the colored triplets become superheavy, they would induce extra rapid proton decay through effective dimension five operators. Leaving aside these pending problems, the model does not fare too well with the fermion masses either. It yields one good relation $m_b = m_\tau$, but one bad relation $m_s = m_\mu$ at the unification scale [108]. Furthermore, it has no real prediction for the other fermion masses and mixing angles. In addition to these difficulties, the model needs to introduce by hand a number of scales spanning between 10^{18} to 10^{14} GeV and supersymmetry breaking. Without going into details, it seems that at least a number of these adverse remarks would apply to the four-dimensional string-construction [105] of the $SU(2)_L \times SU(2)_R \times SU(4)^C$ model as well.

In summary, it seems fair to say that the route where superstrings lead to elementary quarks and leptons has not yet yielded a single model that is manifestly realistic. Furthermore, it has not yet shed much light on the origin of diverse mass scales — from

M_{Planck} to m_ν — and on the parameters of the standard model which include the fermion masses and the mixings. This brings me to present a possible alternative which seems to have certain advantages.

2. The Unconventional Route of Elementary "Preons":

Given that (a) there is such a big gap in energy between the Planck scale and the electroweak scale, where quarks and leptons appear elementary and that (b) the case of elementary quarks and leptons emerging from the superstring theories has not led as yet to any promising picture, it seems in order to consider the possibility that quarks and leptons are not elementary. In this case, the elementary objects emerging from superstrings may represent preonic substructures [13,10] which ultimately bind, utilising a "metacolor" gauge force, to give composite quarks, leptons and Higgs bosons. Apriori, preonic ideas possess some attractive features:

(i) They provide a dynamical origin of the Higgs mechanism like technicolor [109], but without the proliferation and the problems of technicolor. Quarks, leptons and all Higgs-like particles can be built as composites of the same set of elementary objects — no extended technicolor is needed — thereby providing utmost economy. Furthermore, a class of preon models neatly avoids the problem of excessive flavor-changing neutral current processes [10,77], which poses a major obstacle for standard technicolor.

(ii) Because of dynamical symmetry breaking, preonic ideas also avoid the gauge-hierarchy problem [30] altogether, even without the introduction of supersymmetry. Local supersymmetry still seems to be necessary, however, in preonic theories, in order that one may account for certain peculiarities of preon dynamics such as why quarks and leptons are so much lighter than their compositeness scale [110].

Finally, contrary to common impression, there does exist, in my opinion, a viable preon model [10] with manifest economy in field content and parameters, which has evolved over the last several years. It needs just six positive and six negative *massless* chiral superfields $(\Phi_\pm^{a,\sigma})_{a=1,2...6;\sigma=1,2...N}$, which define two basic flavors (u, d) and four basic colors (r, y, b, ℓ) including lepton color [4]. They couple to an asymptotically free "metacolor" gauge force generated by a symmetry $SU(N)_M$ and the flavor-color gauge forces generated, for example, by $SU(2)_L \times SU(2)_R \times SU(4)^C$ near the Planck scale. Thus the model has at most three parameters represented by the gauge couplings g_N, g_2 and g_4; these would reduce to just one, if there is an underlying unity of forces, as we envisage (see below).

Such a model has not yet been derived from a superstring theory, although there does not appear to be any bar, in principle, in this regard [111], especially in the context of four-dimensional constructions [85]. Nevertheless, if one introduces the economical preon-picture mentioned above through an effective lagrangian just below the Planck scale and assumes $N = 1$ local supersymmetry, one seems to be able to derive a number of advantages [112]. These include:

(i) *A Unification of Scales*: It is shown [10] that the model is capable of generating all the diverse scales — from M_{Planck} to m_ν — and thereby the small numbers such as $(m_W/M_{Pl}) \sim (m_t/M_{Pl}) \sim 10^{-17}$, $(m_C/M_{Pl}) \sim 10^{-19}$, $(m_e/M_{Pl}) \sim 10^{-22}$, and $(m_\nu/M_{Pl}) \lesssim 10^{-27}$ — *in terms of just one fundamental input parameter*: the coupling constant α_M associated with the metacolor force. Choosing $\alpha_M \approx .07$ to $.05$ at $M_{Pl}/10$, the metacolor force becomes strong at a scale $\Lambda_M \approx 10^{11}\ GeV$ for $N = 5$ to 6. Thus, one small number $(\Lambda_M/M_{Pl}) \sim 10^{-8}$ arises naturally through calculable renormalisation effects, which lead to a slow logarithmic growth of α_M. The remaining small scales arise [110,10] primarily due to the constraints of the index theorem [113] which forbids SUSY breaking (in the class of models under consideration) in the limit of global SUSY. Thus SUSY breaks only through the *collaboration of gravity* with the metacolor force. As a result, the SUSY-breaking metagaugino and matter-fermion condensates [114], some of which also break $SU(2)_L \times U(1)$, are damped by (Λ_M/M_{Pl}) [110]. This ends up giving $m_W \sim m_t \sim (\frac{1}{10})\Lambda_M(\Lambda_M/M_{Pl}) \sim (\frac{1}{10})M_{Pl}(\Lambda_M/M_{Pl})^2 \sim M_{Pl}(10^{-17}) \sim 100\ GeV$. A double see-saw mechanism yields $m_{\nu_L} \lesssim (K_i^2)\Lambda_M(\Lambda_M/M_{Pl})^2 \lesssim (10^{-3})M_{Pl}(\Lambda_M/M_{Pl})^3 \sim M_{Pl}(10^{-27})$ [10].

In this way, one can conceive of a *common origin* of all the diverse scales — from M_{Planck} to m_ν. Local supersymmetry and compositeness of quarks and leptons (which implies the presence of a new metacolor force) play crucial roles in achieving this remarkable result. One novel feature arising from the constraints of the index theorem is that electroweak symmetry breaking is dynamical, but contrary to common belief, it has its origin in a new gauge force in the observable sector [115] that becomes strong at a superheavy scale $\Lambda_M \sim 10^{11}\ GeV$, rather than at 1 TeV.

(ii) *Family Replication:* Owing to fermion-boson pairing in SUSY, the model provides a compelling reason for family replication and a good reason (subject to the assumption of saturation at the level of minimum dimensional composite operators) for having just three chiral families [116]. In the process, it predicts that there must also exist *two complete vector-like families* (Q and Q') [10] with masses of order 200 GeV-2 TeV, where Q couples vectorially to W_L's and Q' to W_R's.

(iii) *Inter-family Hierarchy:* The three chiral families acquire masses by a see-saw mechanism almost entirely through their mixings with the two vector-like families [10,77]. One finds that although the three families (e, μ and τ) are made of the stuff and are on par with each other as regards binding [116], the symmetry of the theory forces a hierarchical mass-pattern with one of them being massless, barring corrections of order 1 MeV [77]. In this way, one understands naturally why $\hat{m}_e << \hat{m}_\mu << \hat{m}_\tau$, with $m_e \sim O(1\ MeV)$ and $m_t \sim 100\ GeV$. The model also exhibits a hierarchical CKM mixings.

In other words, the preon model provides at least a qualitative understanding of the bulk of the parameters associated with the standard model. In particular, it explains why $m_e/m_t \sim 10^{-5}$.

(iv) *CP Violation:* The model provides an elegant mechanism for spontaneous CP violation, which is shown to vanish (for the observed processes, if the masses of the electron family were set to zero [77]. The model predicts an electric dipole moment for the neutron $\approx (1 \text{ to } \frac{1}{30}) \times 10^{-25}$ *ecm*, which is observable.

(v) *Crucial Tests:* One distinguishing feature of the preon model is that it leads to several crucial predictions [77] by which it can really be falsified if it is wrong. First, owing to the mixing of the chiral with the vector-like families, there are new contributions to $K^o - \overline{K}^o$, $K_L \to \mu^+\mu^-$ and $K_L \to \bar{\mu} e$, which are, however, smaller typically by one to two orders of magnitude than that of the standard model, while those for $B^o - \overline{B}^o$ are comparable to that of the standard model. In other words, the model is safe so far, unlike technicolor. It predicts intriguing new processes such as: (a) $Z \to t\bar{c}$ with an amplitude of order (2 to $\frac{1}{2}$)% of the normal Z^o-coupling, (b) $Z \to c\bar{u}$ which gives $\Delta m(D - \overline{D}) \approx (10 \text{ to } 3) \times 10^{-14}$ *GeV*, which is at least 20 times bigger than that of the standard model, (c) $Z \to \bar{\mu} e$ leading to $B(\mu \to 3e) \approx (1 \text{ to } 5) \times 10^{-13}$, (d) enhanced and observable $\nu_\mu - \nu_\tau$ oscillations [117] and (e) a (3 to 10)% *increase* in top and tau life-times compared to those expected in the standard model together with a predictable *decrease* in LEP neutrino counting from 3 [77,118].

One dramatic prediction and hallmark of the model is, of course, the existence of two vector-like families Q and Q' whose leptons (E, E', N, N') have masses in the range of (200-600) *GeV* and quarks (U, D, U', D') have masses $\approx$ (600-2000) *GeV* [10]. The existence of such heavy *vector-like* (rather than chiral) fermions is found to be perfectly compatible [118] with precision electroweak tests including measurements of the ρ and the S or ϵ_3 paramters [119]. These vector-like families should provide rich new physics to be probed at SSC, LHC and TeV-range e^-e^+ colliders.

(vi) *Supersymmetry Breaking:* The preonic theory requires the presence of a new metacolor force in the *observable sector* which becomes strong at a superheavy scale $\Lambda_M >> 1$ *TeV*. Such a force would not, however, be permissible if quarks and leptons were elementary. The existence of such a force, in the presence of gravity, allows the possibility of a dynamical breaking of supersymmetry directly in the observable sector (albeit with a damping by the Planck mass [110]), which thus transmits efficiently into the masses of the squarks, the gluinos and the winos. This may well be an advantage over a dynamical breaking of supersymmetry occurring entirely in the hidden sector of a superstring theory [97-101], which seems to be the only possibility in such a theory if quarks and leptons are elementary.

(vii) *Grand Unification at the Preon Level:* The preon model [10] based on the effective gauge symmetry $G_P = SU(N)_M \times SU(2)_L \times SU(2)_R \times SU(4)^C$, which operates just below the Planck scale, possesses at least three of the main ingredients of grand unification — *i.e.* (a) quark-lepton unification because of $SU(4)^{color}$, (b) quantisation of electric charge because of the non-abelian nature of the gauge symmetry G_P and (c) spontaneous violation of $B - L$ because of the gauging of $B - L$ (see sec. 4.9) [120]. Note also that the preon-content and the gauge symmetry G_P respect left-right-

symmetry. At the composite level, therefore, ν_L is accompanied by ν_R. Furthermore, subject to the assumption that the preonic condensate $< \Delta_R > \sim \Lambda_M \sim 10^{11}$ *GeV* forms [120], one has essentially the same physics in the preonic theory at the intermediate scale Λ_M as in a two-step breaking of a left-right symmetric grand unified theory like $SO(10)$ with elementary quarks and leptons and elementary Higgs Δ_R (see secs. 4.11 and 4.3). This includes the see-saw generated light ν_L's and heavy Majorana ν_R's.

What about the remaining aspect of grand unification — *i.e.* the unity of forces? The gauge symmetry G_P, subject to $L - R$ symmetry, would seem to allow three gauge coupling constants g_M, g_2 and g_4. However, if G_P and the associated preon-content (specified above) arise from an underlying superstring theory, which could happen, in particular, through a four-dimensional construction [85], these coupling constants would in fact be equal at the string unification scale $M_{SU} \approx 10^{18}$ *GeV*. In this case, the spirit of grand unification would be fully retained at the level of preons.

Although the derivation of a preon model resembling that proposed in Ref. 10 from a superstring theory is still awaited, it is intriguing to ask whether the coupling constants g_1, g_2 and g_3 extrapolated in the context of the preon model from their measured values at low energies do indeed meet with each other as well as with g_M near the Planck scale, for any reasonable choice of the metacolor gauge symmetry G_M. Note that the extrapolation involves three regions of particle-spectrum: (I) from 1 *GeV* to about 1 *TeV* where one has the standard three families of quarks and leptons and the standard gauge bosons, (II) from 1 *TeV* to the metacolor scale $\Lambda_M \sim 10^{11}$ *GeV* where the spectrum consists of the three chiral and two vector-like quark-lepton families, the standard model gauge bosons and the superpartners of all these and (III) from $\Lambda_M \sim 10^{11}$ *GeV* to the unification-scale $M_{SU} \approx 10^{18}$ *GeV*. In region (III), the spectrum is supersymmetric as in region II and is made of preons (specified above) and the gauge particles of the metacolor gauge symmetry G_M as well as of the flavor-color gauge symmetry G_{fc}, which may be chosen to be either $\mathcal{G}_o = SU(2)_L \times SU(2)_R \times SU(4)^C$ or in general even a subgroup of $\mathcal{G}_o$ containing $SU(2)_L \times U(1) \times SU(3)^C$. [The precise nature of G_M and G_{fc} may hopefully get determined ultimately by an underlying superstring theory if preons have their origin from such a theory.]

Note that the extrapolation based on renormalization group equations is fully determined in regions I and II because the spectrum and the gauge symmetry are fixed, while in region III, there is only a few discrete choices which can be made as regards the metacolor gauge symmetry G_M and the flavor-color gauge symmetry G_{fc}. The scale Λ_M is fixed at about 10^{11} *GeV* by requiring consistency with the hierarchy of scales. It is found [121] that with the measured low-energy values of g_1, g_2 and g_3 and thus with $\sin^2\theta_W(m_Z) \simeq .23$ as an input, the coupling constants including g_M, remarkably enough, do indeed converge to a common value within a few percent of each other [122] near the Planck scale, for the choice $G_M = SU(5)_M$ and $G_{fc} = SU(2)_L \times U(1)_{I3R} \times SU(4)^C$ [123]. A sketch of the running of the coupling constants α_1, α_2, α_3 and α_M for this particular choice of G_M and G_{fc} is shown in Fig. 4. Note that α_2 and α_3 (just barely) loose asymptotic freedom in region II owing to contributions from composite SUSY partners. Yet, all the coupling constants including α_1, α_2 and α_3 remain perturbative [125] in the entire range of extrapolation except of course α_M which becomes strong just around Λ_M. I find that such a meeting of the coupling constants at the preonic level near the Planck scale, despite the few discrete choices available for G_M and G_{fc}, is intriguing and non-trivial. At the very least, it shows that the unity of all gauge forces may well occur at the preon level rather than the quark-lepton level.

Thus, in my opinion, the preonic picture, though unconventional, should be regarded as a *viable contender* of the elementary quark-lepton picture. Even without being derived at present from an underlying superstring theory, it exhibits utmost economy in field-content and parameters [10], and is capable of addressing some major issues such as the puzzles of hierarchy of mass scales [10], family replication [116] and inter-family mass-splittings [77]. Furthermore, as discussed above, grand unification of forces can well materialise through preons. *In this case, all the constraints of grand unification would be retained at low energies at the level of composite quarks and leptons, including a relationship between g_1, g_2 and g_3 and a prediction for $\sin^2\theta_W$.* Finally, since $B-L$ is gauged through $SU(4)^{color}$, *it must be violated spontaneously in the preonic theory as well just as in $SO(10)$.* One would expect a few condensates to form at the metacolor scale $\Lambda_M \sim 10^{11}\ GeV$, which happens to match precisely the intermediate scale arising in a two-step breaking of $SO(10)$. (See sec. 4.11). One would, in general, expect some of these condensates to break B, L and $B-L$. These could thus induce $(B-L)$ violating proton-decays of the type discussed in sec. 4.11 [see Ref. 47,48], and lepton-number violating Majorana masses for the neutrinos (through $< \Delta_R > \sim O(\Lambda_M)$). In these respects, the preon theory, subject to the assumption of the formation of a few condensates like $< \Delta_R >$, could thus resemble the intermediate-scale physics of $SO(10)$ with elementary quarks, leptons and Higgs bosons.

On the negative side, the preon picture possesses a certain arbitrariness with regard to the choice in the pattern of symmetry breaking and correspondingly in the precise set of condensates which form near Λ_M [114]. While this type of arbitrariness is present in one form or another in alternative approaches as well (compare with the needed VEVs of the scalar partners of ν_R and N in the three-generation Calabi-Yau models), it is clearly desirable to understand the dynamics of locally supersymmetric gauge theories to shed some light on these issues. It also remains to be seen whether the preon-content and gauge symmetry of the type suggested in Ref. 10 can be derived from an underlying superstring theory. These are some of the challenges confronting the preon picture.

On the positive side, the preon picture distinguishes itself by offering several crucial tests as listed above which can help falsify it if it is wrong or "establish" it if it is right. Some of these tests can be performed in the very near future by ongoing and forthcoming facilities and some have to wait until SSC, LHC and/or TeV-range e^-e^+-

forthcoming facilities and some have to wait until SSC, LHC and/or TeV-range e^-e^+-
colliders are available. These tests can also clearly help distinguish the preon picture from the picture of elementary quarks and leptons.

Regardless of which of these two pictures survive, assuming that the survivor has its origin in a superstring theory, grand unification in the traditional sense will be lost. This is because, in this case, (a) the gauge symmetry below the Planck scale, (b) the field-content and (c) the unification-scale are likely to depart significantly from those in traditional grand unification [4,5], even including its supersymmetric extensions [54,58]. Yet, grand unification, *viewed from a broader perspective*, in particular all the constraints of grand unification as outlined in the beginning of this chapter, would be fully retained, for either of the two alternative pictures. Assuming that a fundamental theory should be devoid of arbitrariness, I thus believe that grand unification will survive, but only in this broader sense.

At this juncture, clues from improved experimental studies, including (i) searches for proton-decay at the awaited Superkamiokande facilities, (ii) studies of solar neutrinos at SAGE and GALLEX and of neutrino-oscillations in the laboratory, (iii) searches for small departures from standard model-predictions as regards universality and flavor-changing neutral current processes of the type listed above, and (iv) searches for SUSY-partners as well as for vector-like families at the proposed SSC, LHC and TeV-range e^-e^+-colliders, are of crucial importance. These clues are necessary and they may well be sufficient to ascertain whether the belief in grand unification, in the sense outlined above, is well-founded.

The research is supported in part by the National Science Foundation. It is a pleasure to thank K.S. Babu, M. Cvetič, S.J. Gates, T. Hübsch, R.N. Mohapatra, H. Nishino, Abdus Salam and H. Stremnitzer for many stimulating discussions and J. Sucher for helpful comments on the manuscript. I am grateful to Academician Professor L.V. Keldysh, to Professors V. Fainberg, V. Vanyashin and V. Zaikin, and to Dr. S. Afanasiev for the kind hospitality extended to me in Moscow.

References and Footnotes

1. A.D. Sakharov, *Pisma Zh. Eksp. Teor. Fiz.* **5**, 32 (1967). A preliminary remark in this regard was made by S. Weinberg in *Lectures on Particles and Fields*, ed. by S. Deser and K. Ford (Prentice Hall), 1964; p. 482.

2. S. Weinberg, *Phys. Rev. Lett.* **19**, 1269 (1967); Abdus Salam, in *Elementary Particle Theory*, Nobel Symposium, ed. by N.Svartholm (Almqvist, Stockholm, 1968), p. 367; S.L. Glashow, *Nucl. Phys.* **22**, 57a (1961);

3. G. 't Hooft, *Nucl. Phys.* *B* **33**, 173 (1971).

4. J.C. Pati and Abdus Salam; Proc. 15^{th} High Energy Conference, Batavia, reported by J.D. Bjorken, Vol. 2, p.. 301 (1972); *Phys. Rev. D* **8**, 1240 (1973); *Phys. Rev. Lett.* **31**, 661 (1973); *Phys. Rev. D* **10**, 275 (1974).

5. H. Georgi and S.L. Glashow, *Phys. Rev. Lett.* **32**, 438 (1974).

6. H. Georgi, H. Quinn and S. Weinberg, *Phys. Rev. Lett.* **33**, 451 (1974).

7. J. Wess and B. Zumino, *Nucl. Phys. B* **70**, 39 (1974); *Phys. Lett.* **49B**, 52 (1974); D. Volkov and V.P. Akulov, *JETP Lett.* **16**, 438 (1972).

8. D. Freedman, S. Ferrara and P. Van Nieuwenhuizen, *Phys. Rep. D* **13**, 3214 (1976); S. Deser and B. Zumino, *Phys. Lett.* **62B** 335 (1976).

9. An early attempt in this direction was based on the idea that quarks and leptons are made of a fermion and a boson constituent, the former carrying a flavor and the latter carrying a color quantum number including lepton color. This was initiated in the footnote 7 of the paper by J.C. Pati and Abdus Salam, *Phys. Rev. D* **10**, 275 (1974) and has subsequently been developed by them and by other authors. For a review of the composite models and a list of other references see M. Peskin, Proc. 1981 Int'l Symp. on Lepton and Photon Int. at High Energies, Bonn 1981, ed. W. Pfeil, p. 880.

10. J.C. Pati, "A Model For a Unification of Scales — From M_{Planck} to m_ν," *Physics Letters* **228**, 228 (1989). Other relevant references may be found here.

11. M.B. Green and J.H. Schwarz, *Phys. Lett.* **149B**, 117 (1984); ibid **151B**, 21 (1985); D.J. Gross, J.A. Harvey, E. Martinec and R. Rohm, *Phys. Rev. Lett.* **54**, 502 (1985); P. Candelas, G.T. Horowitz, A. Strominger and E. Witten, *Nucl. Phys. B* **258**, 46 (1985). For a review of certain important issues see *e.g.* D. Gross, Rapporteur Talk, High Energy Conf. held at Munich (August, 1988), page 310.

12. For example, the ten-dimensional $SO(32)$ or the heterotic $E_8 \times E_8'$ theories involve one gauge coupling constant at the basic level. The more recently discovered four-dimensional constructions (see later) also possess this feature.

13. T. Hübsch, H. Nishino and J.C. Pati, *Phys. Letters* **163B**, 111 (1985).

14. The idea of gauging $SU(3)$-color symmetry was first introduced by M. Han and Y. Nambu in *Phys. Rev.* **139**, B 1006 (1965) where, however, $SU(3)$-color symmetry was broken explicitly by electromagnetism rendering the theory unrenormalisable; furthermore, an additional medium strong force was presumed to accompany the $SU(3)$-color gauge force. This idea was reintroduced in the modern context in Ref. 4 (where quark charges were allowed to be either integral or fractional) and independently by H. Fritzsch and M. Gell-Mann, *Proc.* 15^{th} High Energy Conf. held at Batavia, Vol. 2, page 135 (1972). The idea of $SU(3)$-color as a global symmetry was introduced implicitly through the idea of paraquarks by O.W. Greenberg, *Phys. Rev. Lett.* **13**, 598 (1964) and explicitly, in a manner commonly used, by Han and Nambu (Reference above).

15. D. Gross and F. Wilczek, *Phys. Rev. Lett.* **30**, 1343 (1973); H.D. Politzer, *Phys. Rev. Lett.*, 1346 (1973).

16. R.N. Mohapatra and J.C. Pati, *Phys. Rev. D* **11**, 566 (1975); *Phys. Rev. D* **11**, 2558 (1975).

17. G. Senjanović and R.N. Mohapatra, *Phys. Rev. D* **12**, 1502 (1975).

18. A. Davidson, *Phys. Rev. D* **20**, 776 (1979); R.N. Mohapatra and R. Marshak, *Phys. Lett.* **91B**, 222 (1980).

19. M. Gell-Mann, P. Ramond and R. Slansky, in *Supergravity*, ed. by D. Freedman and P. Van Nieuwenhuizen (North Holland, 1979), p. 315; T. Yanagida, Proc. Workshop on "Unified Theory and Baryon Number in the Universe," eds. O. Sawata and A. Sugamoto, KEK, Japan (1979); R.N. Mohapatra and G. Senjanovic, *Phys. Rev. Lett.* **44**, 912 (1980).

20. R.N. Mohapatra and J.C. Pati, *Phys. Rev. D* **11**, 566 (1975). For an extension to the case of three generations see D. Chang, *Nucl. Phys. B* **214**, 435 (1983).

21. Owing to gluonic renormalisation, the up-quark member of a given family is expected to be heavier than the Dirac mass of the neutrino in the same family (at least in the absence of $\langle\xi\rangle = \langle(2,2,\underline{15})\rangle$).

22. S.P. Mikheyev and Y. Yu Smirnov, *Nuovo Cim.* **9C**, 17 (1986).

23. L. Wolfenstein, *Phys. Rev. D* **17**, 2369 (1978).

24. See *e.g.* H.A. Bethe and J.N. Bahcall, *Phys. Rev.* **D44**, 2962 (1991) and references therein.

25. To obtain a tiny majorana mass ν_L in the $SU(2)_L \times U(1)_Y$ model, one would have to introduce a new Higgs multiplet transforming as a triplet of $SU(2)_L$ with $Y_W = 0$ just for this purpose. Its VEV would have to be unnaturally small compared to that of the Higgs doublet, or else the ρ-parameter would be disturbed in conflict with observation. In this sense, a massive neutrino (even with a majorana mass) is rather unnatural in the standard model and its left-right asymmetric extensions such as $SU(5)$.

26. For $[SU(4)]^4$, see J.C. Pati and Abdus Salam, *Phys. Lett.* **58B**, 333 (1975). For $SU(8) \times SU(8)$, see J.C. Pati, Proc. Einstein Centennial Symposium, (April 1979), page 221, ed. by Y. Ne'eman (Addison-Wesley) and N. Deshpande and P. Mannheim, Univ. of Oregon preprint (1979). For $SO(10)$, see H. Fritzsch and P. Minkowski, *Ann. Phys. (NY)* **93**, 193 (1975), and H. Georgi, Proc. AIP Conf. Williamsburg (1974). For $SU(16)$, see J.C. Pati, Abdus Salam and J. Strathdee, II, *Nuov. Cim.* **26A**, 77 (1975); *Nucl. Phys.* **B185**, 445 (1981). For E_6, see F. Gursey, P. Ramond and P. Sikivie, *Phys. Lett.* **60B**, 177 (1976); F. Gürsey and P. Sikivie, *Phys. Rev. Lett.* **36**, 775 (1976); P. Ramond, *Nucl. Phys.* **B110**, 214 (1976).

27. For a comprehensive discussion of $SO(10)$ and E_6 and their representations, see R. Slansky, *Phys. Reports* **79**, 1 (1981).

28. See *e.g.* M.S. Chanowitz, J. Ellis and M. Gaillard, *Nucl. Phys.* **B128**, 506 (1977).

29. M. Gell-Mann, P. Ramond and R. Slansky, (Ref. 19).

30. E. Gildener, *Phys. Rev.* **D14**, 1667 (1976).

31. The $\underline{15}$ of $SU(5)$ contains $SU(3)^C$-singlet electrically neutral member, which, however, transforms as a member of a $SU(2)_L$-triplet. VEV of such a member can give Majorana mass to ν_L. The VEV of $\langle H_{15}\rangle$ must be much smaller than that of $\langle H_5\rangle$ to preserve the $\Delta I_W = \frac{1}{2}$ rule (*i.e.* $\rho \equiv m_W/(m_Z \cos\theta_W) = 1$). Furthermore, $\langle H_{15}\rangle$ and/or the corresponding Yukawa coupling would have to be artificially small compared to the corresponding parameters of $\langle H_5\rangle$ to account for $(m_{\nu_e}/m_{e^-}) \lesssim 10^{-5}$.

32. A.J. Buras, J. Ellis, M.K. Gaillard and D.V. Nanopoulos, *Nucl. Phys.* **B135**, 66 (1978).

33. H. Georgi and C. Jarlskog, *Phys. lett.* **86B**, 297 (1979); P. Frampton, S. Nandi and J. Scanio, *Phys. lett.* **85B**, 255 (1979).

34. T. Appelquist and J. Carrazone, *Phys. Rev.* **D11**, 2856 (1975).

35. For results of ALEPH, DELPHI, L3 and OPAL Collaborations, see Plenary talks by T. Hebbeker and J. Carter, *Proc. Lepton-Photon Symposium*, Geneva, July '91; and a recent preprint CERN-PPE/91-232 (Dec. 20, 1991).

36. U. Amaldi, W. de Boer, H. Fürstenau, *Phys. Lett.* **B260**, 447 (1991); J. Ellis, S. Kelley and D.V. Nanopoulos, *Phys. Lett.* **B260**, 131 (1991); P. Langacker and M. Luo, *Phys. Rev.* **D44**, 817 (1991).

37. For an earlier analysis, which is still fairly relevant today, see U. Amaldi *et al.*, *Phys. Rev.* **D36**, 1385 (1987), and W. Marciano, *Proc. Eighth Grand Unification Workshop* held at Syracuse, N.Y. in April 1987, ed. by K.C. Wali, published by the World Scientific, pages 185-199.

38. V. Kuzmin, *JETP Lett.* **12**, 228 (1980); S.L. Glashow in *Recent Developments in Gauge Theories*, Cargese, 1979; R.N. Mohapatra and R.E. Marshak, *Phys. Rev. Lett.* **44**, 1316 (1980).

39. S. Weinberg, *Phys. Rev. Lett.* **43**, 1566 (1979); F. Wilczek and A. Zee, *Phys. Rev. Lett.* **43**, 1571 (1979).

40. IMB-3 Collab., R. Becker, S. Zendy, *et al.*, *Phys. Rev.* **D42**, 2974 (1990). See report by F. Reines at this conference for the latest IMB results and also the results of Kamiokande and other experimental groups.

41. See *e.g.* D. Chang, R.N. Mohapatra, J.M. Gipson, R.E. Marshak and M.K. Parida, *Phys. Rev.* **D31**, 1718 (1985); J. Gipson and R.E. Marshak, *Phys. Rev.* **D31**, 1705 (1985). For a recent work, see R.N. Mohapatra and M.K. Parida, UMD-PP-92-179.

42. J.C. Pati and Abdus Salam, Proc. First Workshop on Grand Unification, New Hampshire (April, 1980); p. 115; J.C. Pati, Abdus Salam and J. Strathdee, *Nucl. Phys.* **B185**, 445 (1981); J.C. Pati, Abdus Salam and U. Sarkar, *Phys. Lett.* **113B**, 330 (1983).

43. T. Rizzo and G. Senjanović, *Phys. Rev.* **D25**, 235 (1982); A. Raychaudhri and U. Sarkar, preprint (1982).

44. F. del Aguilla and L. Ibanes, *Nucl. Phys.* **B177**, 60 (1981).

45. M. Fukugita, T. Yanagida and M. Yoshimura, *Phys. Lett.* 106B, 183 (1981).

46. R.N. Mohapatra and G. Senjanović, *Z. Phys.* **C17**, 1292 (1983); D. Reiss, *Phys. Lett.* **B109**, 365 (1982); R. Holman, G. Lazarides and Q. Shafi, *Phys. Rev.* **D27**, 995 (1983) and Q. Shafi and F.W. Stecker, *Phys. Rev. Lett.* **53**, 1292 (1984).

47. J.C. Pati, A. Salam and U. Sarkar, *Phys. Lett.* **133B**, 330 (1983).

48. J.C. Pati, *Phys. Rev. D*, Rap. Comm. **29**, 1549 (1984).

49. See, *e.g.* D. Chang and G. Senjanović, *Phys. Lett.* **188**, 231 (1987). See also Ref. [57-59].

50. See also K. Kang and M. Shin, *Mod. Phys. Lett.* *A* **1**, 585 (1986).

51. See also P. Langacker, R.D. Peccei and T. Yanagida, DESY Preprint (1986). In this reference as well as in Ref. 50, lepton number is used as a global symmetry. Thus, unlike in the case of $SU(4)$-color, $B-L$ is not a local symmetry in these models.

52. J.C. Pati, Proc. Eighth Grand Unification Workshop held at Syracuse (April, 1987), page 361, ed. by K.C. Wali.

53. E. Witten, *Nucl. Phys.* *B* **186**, 513 (1981); R. Kaul (1981).

54. S. Dimopoulos and H. Georgi, *Nucl. Phys.* *B* **193**, 150 (1981); N. Sakai, *Z. Physik* *C* **11**, 153 (1982).

55. Within a supersymmetric theory, the Higgs scalar would get a contribution to its $(mass)^2$ given by $\delta m_H^2 \sim \frac{\alpha}{\pi}\left|m_B^2 - m_F^2\right|$, where m_B and m_F are the masses of the boson and the fermion ir a SUSY multiplet. For $m_H \simeq \delta m_H \lesssim 1\,TeV$, one thus needs the SUSY partner of ordinary particles (like quarks, leptons and gluons which are much lighter t an 100 GeV) to have masses $\lesssim (\pi/\alpha)^{\frac{1}{2}}$ $(1\ TeV)$.

56. V. Amaldi, W. de Boer ·d H. Fürstenau (Ref. 36).

57. S. Weinberg, *Phys. Rev.* *D* **26**, 287 91982); N. Sakai and T. Yanagida, *Nucl. Phys.* *B* **197**, 533 (1982).

58. A.H. Chemseddine, R. Arnowitt and P. Nath, *Phys. Rev. Lett.* **49**, 970 (1982); R. Barbieri, S. Ferrara and C.A. Savoy, *Phys. Lett.* **119B**, 343 (1982); L.E. Ibanez, *Phys. Lett.* 118 B, 73 (1982); J. Ellis, D.V. Nanopoulos and K. Tamvakis, *Phys. Lett.* **121B**, 123 (1983).

59. L. Alvares-Gaume, J. Polchinski and M. Wise, *Nucl. Phys.* *B* **221**, 495 (1983); J. Ellis, J. Hagelin, D. Nanopoulos and K. Tamvakis, *Phys. Lett.* **125B**, 275 (1983).

60. V. Kuzmin, *Pisma Zh. Eksp. Teor. Fiz.* **12**, 335 (1970).

61. A.V. Ignatiev, N.V. Krasnikov, V.A. Kuzmin and A.N. Tavkhelidze, *Phys. Lett.* **76B**, 436 (1978); M. Yoshimura, *Phys. Rev. Lett.* **41**, 281 (1978); **42**, 740(E) (1979); S. Dimopoulos and L. Susskind, *Phys. Rev.* *D* **18**, 4500 (1978); D. Touissant, S.B. Treiman, F. Wilczek and A. Zee, *Phys. Rev.* *D* **19**, 1036 (1979); S. Weinberg, *Phys. Rev. Lett.* **42**, 850 (1979); J. Ellis, M.K. Gaillard and D.V. Nanopoulos, *Phys. Lett.* **80B**, 360 (1979).

62. E.W. Kolb and M.S. Turner, "The Early Universe", Addison-Wesley Publication.

63. A. Guth, *Phys. Rev.* *D* **23**, 347 (1981).

64. For a discussion of this point, See A. Guth, Proc. of the Nuffield Workshop on "The Very Early Universe", 1982, p. 172-205, ed. by G. Gibbons, S. Hawking and S. Siklos, Cambridge Univ. Press.

65. A.D. Linde, *Phys. Lett.* **108B**, 389 (1982); A. Albrecht and P.J. Steinhardt, *Phys. Rev. Lett.* **48**, 1220 (1983).

66. For some novel suggestions in this regard, see S. Coleman, *Nucl. Phys.* *B* **307**, 867 (1988); ibid *B* **310**, 643 (1988) and references therein.

67. Such a flat potential is attributed to a Coleman-Weinberg ective potential with a vanishing mass and a very small quartic coupling λ for tl e inflaton lying in the false vacuum [65].

68. G. 't Hooft, *Nucl. Phys.* *B* **79**, 276 (1974); A. Polyakov, *JETP Lett.* **20**, 194 (1974).

69. Q. Shafi and A. Vilenkin, *Phys. Rev. Lett.* **52**, 691 (1984); S.Y. Pi, *Phys. Rev. Lett.* **52**, 1725 (1984).

70. J. Ells, D. Nanopoulos, K. Olive and K. Tamvakis, *Phys. Lett.* **118B**, 335 (1982).

71. R. Holman, P. Ramond and G.G. Ross, *Phys. Lett.* **137B**, 343 (1984); B. Ovrut and P.J. Steinhardt, *Phys. Rev. Lett.* **53**, 732 (1984).

72. M. Cvetič, T. Hübsch, J.C. Pati and H. Stremnitzer, *Phys. Rev.* *D* **40**, 1311 (1989).

73. S. Weinberg, *Phys. Rev. Lett.* **48**, 1303 (1982).

74. J. Ellis, J.E. Kim and D. Nanopoulos, *Phys. Lett.* **145B**, 181 (1984); M. Yu Khlopov and A.D. Linde, *Phys. Lett.* **138B**, 265 (1984).

75. V. Kuzmin, V. Rubakov and M. Shaposhnikov, *Phys. Lett.* **155B**, 36 (1985); M. Fukugita and T. Yanagida, *Phys. Lett.* *B* **174**, 45 (1986); G. Lazarides and Q. Shafi, *Phys. Lett.* *B* **258**, 305 (1991).

76. M.A. Luty, *Phys. Rev.* *D* **45**, 455 (1992).

77. K. Babu, J.C. Pati and H. Stremnitser, "Fermion Masses and CP Violation in a Model with Scale Unification", *Phys. Rev. Lett.* **67**, 1688 (1991).

78. G. 't Hooft, *Phys. Rev. Lett.* **37**, 8 (1976).

79. V. Kuzmin, V. Rubakov and M. Shaposhnikov, Ref. 75.

80. P. Arnold and L. McLerran, *Phys. Rev. D* **36**, 581 (1987); **37**, 1020 (1988).

81. M. Shaposhnikov, *Nucl. Phys.* **B287**, 757 (1987); *Nucl. Phys.* **B299**, 797 (1988); L. Maclerran, *Phys. Rev. Lett.***62**, 1075 (1989).

82. See *e.g.* L. McLerran, M. Shaposhnikov, N. Turok and M. Voloshin, Princeton U. PUPT-90-121224; N. Turok and J. Zadrozny, PUPT-90-1216.

83. M. Green, J.H. Schwarz and E. Witten, "Superstring Theory", Vols. 1 and 2, Cambridge Univ. Monographs (1987).

84. L. Dixon, J. Harvey, C. Vafa and E. Witten, *Nucl. Phys.* **B261**, 678 (1985); *Nucl. Phys.* **B274**, 285 (1986).

85. H. Kawai, D. Lewellen and S.H. Tye, *Phys. Rev. Lett.***57**, 1832 (1986) and *Nucl. Phys.* **B288**, 1 (1987); I. Antoniadis, C. Bachas and C. Kounnas, *Nucl. Phys.* **B289**, 87 (1987); I. Antoniadis and C. Bachas, *Nucl. Phys.* **B298**, 586 (1988); W. Lerche, D. Lust and A. Schellekens, *Nucl. Phys.* **B287**, 477 (1987).

86. D. Gross and V. Periwal, *Phys. Rev. Lett.***60**, 2105 (1988).

87. A. Strominger and E. Witten, *Comm. Math. Phys.***101**, 341 (1985).

88. E. Witten, *Nucl. Phys.* **B268**, 79 (1986).

89. M. Dine and N. Seiberg, *Phys. Rev. Lett.***57**, 2625 (1986).

90. J. Distler and B.R. Greene, *Nucl. Phys.* **B309**, 295 (1988).

91. M. Dine, N. Seiberg, X.G. Wen and E. Witten, *Nucl. Phys.***278**, 769 (1987); **B289**, 319 (1987).

92. S.J. Rey, *Phys. Rev.* **D43**, 526 (1991).

93. Two three-generation manifolds, which are diffeomorphic to each other, are proposed respectively by G. Tian and S.T. Yau, Proc. Argonne Symposium on Anomalies, Geometry and Topology (1985) and R. Schimmrigk, *Phys. Lett.* **B193**, 175 (1987).

94. Y. Hosotani, *Phys. Lett.* **B126**, 309 (1983); E. Witten, *Nucl. Phys.* **B258**, 75 (1985); J. Breit, B. Ovrut and G. Segre, *Phys. Lett.* **B158**, 33 (1985); A. Sen, *Phys. Rev. Lett.***55**, 33 (1985).

95. B. Greene, K. Kirklin, P. Miron and G. Ross, *Nucl. Phys.* **B278**, 667 (1986) and *Phys. Lett.***192B**, 111 (1987) and references therein; N. Ganoulis, G. Lazarides and Q. Shafi, *Nucl. Phys.* **B323**, 374 (1989); M. Eastwood and T. Hübsch, UTTG-25-89; P. Candelas and S. Kalara, *Nucl. Phys.* **B298**, 357 (1988). This is an incomplete sampling of references.

96. For a review of superstring-phenomenology especially of 3-generation Calabi-Yau models and other relevant issues, see B.R. Greene, Proc. 1990 Trieste Summer School in HEP, Publ. by World Scientific, pages 334-420.

97. J.P. Derendinger, L.E. Ibanez and H .P. Nilles, *Phys. Lett.***155B**, 65 (1985); H.P. Nilles, *Phys. Lett.***115B**, 193 (1982).

98. M. Dine, R. Rohm, N. Seiberg and E. Witten, *Phys. Lett.***156B**, 55 (1985)

99. R. Rohm and E. Witten, *Ann. Phys.* **170**, 454 (1986).

100. L. Dixon, V. Kaplunovsky, J. Louis and M. Peskin, SLAC/Texas Preprint (unpublished) and references therein.

101. J.A. Casas, C. Munoz and G.G. Ross, Proc. Workshop on "Current Problems in Particle Physics and Cosmology," Ioannina, Greece (June 1990), Publ. by North-Holland, page 60-72, and references therein.

102. R. Arnowitt and P. Nath, Proc. 1990 Trieste Summer School in HEP, Publ. by World Scientific, p. 664-676.

103. See *e.g.* S. Kalara and R. Mohapatra, *Phys. Rev.* **D36**, 3474 (1987); J. Ellis, K. Enqvist, D.V. Nanoupoulos and K.A. Olive, CERN Preprint 4713 (1987).

104. I. Antoniadis, J. Ellis, J.S. Hagelin and D.V. Nanopoulos, *Phys. Lett.* **B194**, 231 (1987); **B231**, 65 (1989).

105. I. .Antoniadis and G.K. Leontaris, *Phys. Lett.* **B216**, 333 (1989); I. Antoniadis, G.K. Leontaris and J. Rizzos, *Phys. Lett.* **B245**, 161 (1990); G.K. Leontaris and N.D. Tracas, Ioanniana Preprint 260/91.

106. See *e.g.* S. Kalara, J. Lopez and D.V. Nanopoulos, *Nucl. Phys.* **B353**, 421 (1991); **B251**, 369 (1990); J. Lopez and D.V. Nanopoulos, CERN-TH 6097/91 and references therein.

107. I.. Antoniadis, J. Ellis, S. Kelley and D.V. Nanopoulos, CERN-TH 6169/91.

108. I. Antoniadis, J. Rizos and K. Tamvakis, CPTH-A092.1191; Ioannina IOA-265 (Nov. 1991). For earlier works on the question of B-violation, see J. Ellis, J. Lopez and D.V. Nanopoulos, *Phys. Lett.* **B252**, 53 (1990) and G. Leontaris and K. Tamvakis, *Phys. Lett.* **B260**, 333 (1991).

109. For a review see E. Farhi and L. Susskind, *Phys. Rep.* **74**, 277 (1981) and references therein.

110. J.C. Pati, M. Cvetič and H. Sharatchandra, *Phys. Rev. Lett.* **58**, 851 (1987).

111. This is being explored in collaboration with I. Antoniadis and G.K. Leontaris.

112. For a comprehensive review of the model, see J.C. Pati, "Natural Origins of Small Numbers and Family Replication Through Supersymmetry and Compositeness," Proc. Workshop on "Current Problems in Particle Physics and Cosmology," Ioannina, Greece (June 1990), Publ. by North Holland, page 115-127.

113. E. Witten, *Nucl. Phys.* **B185**, 513 (1981); **B202**, 253 (1983); S. Cecotti and L. Girardello, *Phys. Lett.* **B110**, 39 (1983); E. Cohen and L. Gomez, *Phys. Rev. Lett.* **52**, 237 (1984).

114. While the preonic theory [10] is most economical in parameters at the fundamental level, in particular it has no fundamental Higgs bosons and thus no mass and quartic and Yukawa coupling parameters for these bosons, it inevitably needs to *assume* that a certain set of condensates form, albeit in accord with the index theorem [113] and the idea of attractive channels. The advantage of the preonic theory, nevertheless, is that the scales of these condensates, including those which are damped due to the index theorem (see text), do not have to be assumed. They are all *determined* in terms of Λ_M and (Λ_M/M_{Pl}) up to numerical coefficients of order unity. As a result, there are even very few *effective parameters — all of order unity* — which enter into the low energy theory of the composites. These not only generate large hierarchies in scales — *e.g.* between Λ_M and m_W through the damping factor (Λ_M/M_{Pl}) — but also (owing to symmetries of the fermion mass — matrix) between m_t and m_e (see Ref. 10 and 77).

115. The reader may note one similarity here between the preon model [10] and the hidden-sector mechanism of supergravity and superstring models [58, 97, 98], which are applied to the case of elementary quarks and leptons. In both cases gravity plays an essential role, in suppressing m_W strongly relative to an intermediate scale Λ_M or Λ_H. In some specific cases, one obtains, for example, $m_W \sim \Lambda_M(\Lambda_M/M_{Pl})$ for the former and $m_W \sim \Lambda_H(\Lambda_H/m_{Pl})^2$ for the latter. *The physics is, however, very different in the two cases.* In the case of the former, gravity is needed to overcome the constraints of the index theorem [113] which forbids a dynamical breaking of global SUSY, while for the latter, gravity is needed to transmit SUSY

breaking from the hidden to the observable sector. A related major difference is that the driving force causing SUSY-breaking acts in the observable sector for the former and entirely in the hidden sector for the latter. [For the preon theory, there may, of course, well be a hidden-sector force arising from superstrings, in addition to the metacolor force that acts in the observable sector].

116. K. Babu, J.C. Pati and H. Stremnitzer, "A Simple Reason Based on Supersymmetry for Replication of Chiral Families," *Phys. Lett.* **B256**, 206 (1991).

117. K. Babu, J.C. Pati and H. Stremnitzer, "Confronting the Scale-Unifying Model Through $\nu_\mu - \nu_\tau$ Oscillations," *Phys. Lett.* **B264**, 347 (1991).

118. K. Babu, J.C. Pati and Xinmin Zhang, "Tau-Anomaly and Vector-Like Families," Preprint UMDHEP 92-152.

119. D.C. Kennedy and B.W. Lynn, *Nucl. Phys.* **B322**, 1 (1989); M. Peskin and T. Takeuchi, *Phys. Rev. Lett.* **65**, 964 (1990); G. Altarelli and R. Barbieri, *Phys. Lett.* **B253**, 161 (1991); G. Altarelli, R. Barbieri and S. Jadach, CERN Preprint-TH 6124/91.

120. $B - L$ is violated spontaneously through a preonic condensate $\langle\Delta_R\rangle$ transforming as $(1, 3, 10^*)$ under $SU(2)_L \times SU(2)_R \times SU(4)^C$. It is assumed that such a condensate forms. It preserves SUSY and is thus of order $\Lambda_M \sim 10^{11}\ GeV$. It breaks $SU(2)_R \times SU(4)^C$ to $U(1)_Y \times SU(3)^C$ at Λ_M and gives a heavy Majorana mass to ν_R, just like the vev of the elementary Δ_R (see sec. 4.3).

121. K. Babu and J.C. Pati, "Unity of Forces at the Preon Level," To appear.

122. Considering that at $10^{16}\ GeV$, effects of non-renormalizable interactions (which inevitably arise within superstring theories) and quantum gravity can not be completely ignored, it seems that one should not expect convergence of the coupling constants to better than say 5% without the inclusion of such effects.

123. With the inclusion of threshold effects involving composite Higgs scalars such as Δ_R's near the metacolor scale Λ_M, only a few other choices of the metacolor and flavor-color symmetries including $G_M = SU(4)_M$ and $G_{fc} = SU(2)_L \times SU(2)_R \times SU(4)^C$ seem to be consistent with the union of the coupling constants. This is being explored in collaboration with M. Parida.

124. In this regard, it is intriguing that α_1, α_2 and α_3 remain in the perturbative range $(\lesssim \cdot 15)$ all the way from 1 to $10^{16}\ GeV$ only for the minimal choice of the preon-content, which introduces just 2 basic flavors and 4 basic colors [10]. Any increase in the preon-content, for example in the number of flavors from 2 to 4, would make α_3 blow up at about 10^7 GeV (because of SUSY). In other words, the minimal choice is also the maximal allowed one [121].

Table I
Complexion of (B,L) Violations

Type	Process	Selection Rules	Char. Mass Scale
IA	$p \to e^+\pi^0, e^+\omega, \bar{\nu}\pi^+$ $n \to e^+\pi^-$	$\Delta B = \Delta L = -1$ $\Delta(B-L) = 0$	$10^{14\pm1}$ GeV
IB	$p \to \mu^+ K^0, \bar{\nu}_\mu K^+$	same	same
IC	$p \to \mu^+\pi, \mu^+\eta$	same	same
IIA	$p \to e^-\pi^+\pi^+, \mu^-\pi^+\pi^+$ $n \to e^-\pi^+, \mu^-\pi^+$	$\Delta B = -\Delta L = -1$ $\Delta(B-L) = -2$ $\Delta(B+L) = 0$	$10^{11\pm1}$ GeV†
IIB	$p \to e^-e^+\nu\pi^+$ $p \to \mu^-e^+\nu\pi^+$ $p \to \nu e^+\nu, \nu\mu^+\nu$ $n \to e^-e^+\nu, \mu^-e^+\nu$	same	same†
III	$p \to e^-\nu\nu\pi^+\pi^+$ $n \to e^-\nu\nu\pi^+$	$\Delta B = -\frac{\Delta L}{3} = -1$ $\Delta\left(B + \frac{L}{3}\right) = 0$	$10^{8\pm1}$ GeV
IV	$p \to e^+\bar{\nu}\bar{\nu}$, *etc.*	$\Delta B = \frac{\Delta L}{3} = -1$	$10^{4.5\pm.5}$ GeV
V	$n \leftrightarrow \bar{n}$ oscillation deuteron $\to$ pions	$\lvert\Delta B\rvert = 2, \Delta L = 0$ $\Delta(B-L) = \pm 2$	$10^{4.5\pm.5}$ GeV††
VI	$nn \to ppe^-e^-$ neutrino-less Double β-decay	$\Delta B = 0, \Delta L = 2$ $\Delta(B-L) = -2$	$10^{4.5\pm.5}$ GeV

† See Ref. 47 and 48.

†† See Ref. 38.

Table IIA

$SO(10) \to SU(2) \times SU(2) \times SU(4) \to SU(2) \times U(1)_Y \times SU(3)$

$\sin^2\theta_W$	M_I (GeV)	M_U (GeV)	$\Gamma(p \to e^+\pi^0)^{-1}$ in yrs.
.217	1.0×10^{14}	1.0×10^{14}	$0.3 \times 10^{30\pm1.7} \le 1.5 \times 10^{30}$
.225	8×10^{12}	4.3×10^{14}	$7.2 \times 10^{30\pm1.7} \le 3.5 \times 10^{32}$
.23	1.2×10^{12}	8×10^{14}	$7.2 \times 10^{31\pm1.7} \le 3.5 \times 10^{33}$
.235	1.2×10^{11}	1.2×10^{15}	$3.6 \times 10^{32\pm1.7} \le 1.7 \times 10^{34}$

Table IIB

$SO(10) \to SU(2) \times SU(2) \times U(1) \times SU(3) \to SU(2) \times U(1)_Y \times SU(3)$

$\sin^2\theta_W$	M_I (GeV)	M_U (GeV)	$\Gamma(p \to e^+\pi^0)^{-1}$ in yrs.
.217	1.2×10^{14}	1.2×10^{14}	$0.6 \times 10^{29\pm1.7} \le 3 \times 10^{30}$
.225	2.5×10^{12}	4.3×10^{14}	$7.2 \times 10^{29\pm1.7} \le 3.5 \times 10^{32}$
.23	1.3×10^{11}	9×10^{14}	$1.0 \times 10^{32\pm1.7} \le 5 \times 10^{33}$
.235	8×10^{9}	1.6×10^{15}	$1.0 \times 10^{33\pm1.7} \le 5 \times 10^{34}$

***Entries in the second and third columns are taken from the first two papers in Ref. 41, based on a two-loop renormalisation group analysis including Higgs contributions and the older value of $\alpha_s(m_W) \simeq 0.10$. The entries would alter somewhat if one uses the recent value of α_s, although the basic trend and pattern of M_I and M_U remain intact. (See *e.g.* the last paper in Ref. 41 for a recent analysis).**

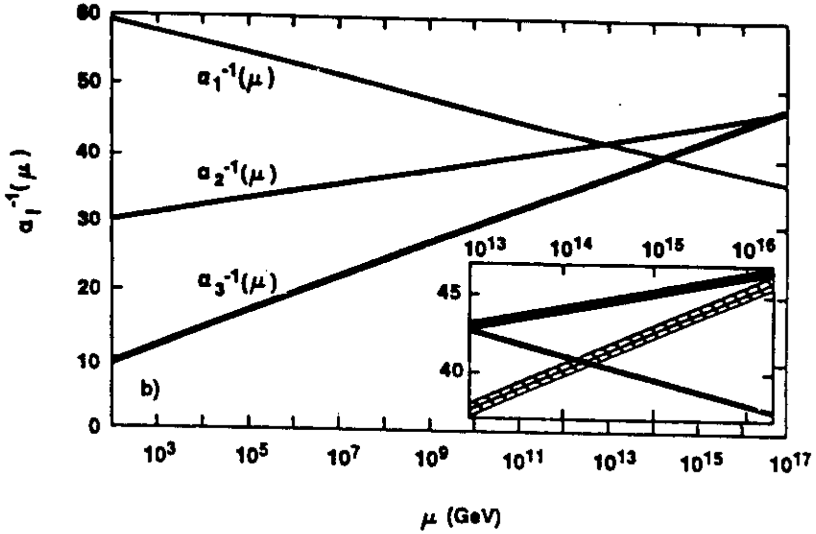


Fig. 2 Running of coupling constants in minimal $SU(5)$ (Ref. 56).

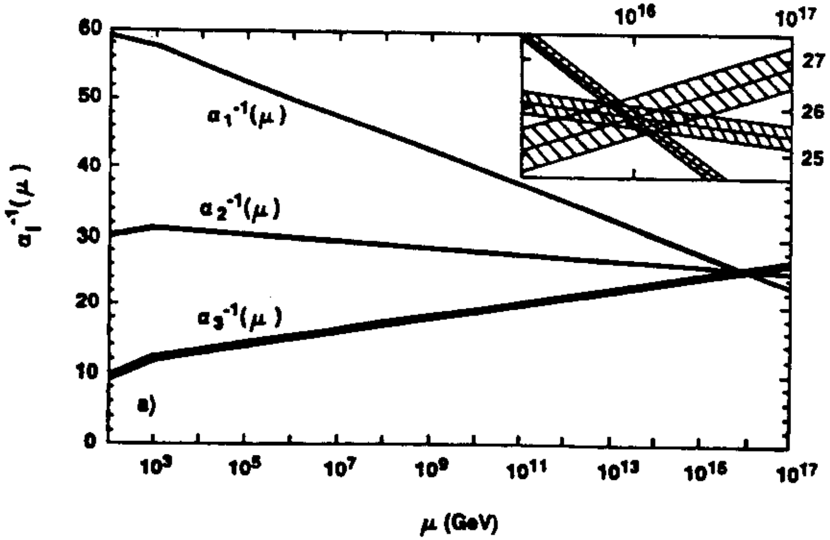


Fig. 3 Running of couplin constants in SUSY minimal $SU(5)$ (Ref. 56).

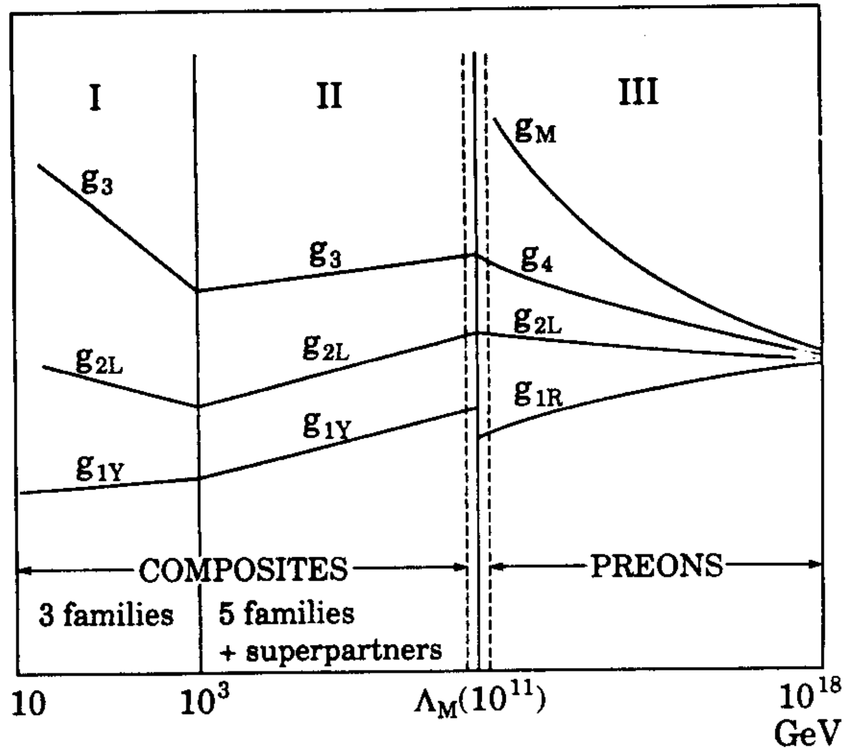


Fig. 4 Unity of Forces at the Preon Level. The figure sketches the running of the coupling constants for the preon model. While it is not drawn to scale, it represents the results of an actual calculation [121], corresponding to the choice $G_M = SU(5)_M$ and $G_{fc} = SU(2)_L \times I_{3R} \times SU(4)^C$ (see text). The five families in region II represent three chiral and two vector-like families which are predicted by the preon model [10,116]. Noting that g_M, g_4, g_{2L} and g_{1R} clearly exhibit the tendency to converge to a point near the Planck scale, it would seem that grand unification, viewed in a broader sense, can well materialize at the preonic level.