DISCOVERY OF PROTON DECAY: A MUST FOR THEORY, A CHALLENGE FOR EXPERIMENT*

Jogesh C. Pati

Department of Physics, University of Maryland, College Park
MD 20742, USA.

April 3, 2000

Abstract

It is noted that, but for one missing piece – proton decay – the evidence in support of grand unification is now strong. It includes: (i) the observed family-structure, (ii) the meeting of the gauge couplings, (iii) neutrino-oscillations, (iv) the intricate pattern of the masses and mixings of all fermions, including the neutrinos, and (v) the need for $B - L$ as a generator, to implement baryogenesis. Taken together, these not only favor grand unification but in fact select out a particular route to such unification, based on the ideas of supersymmetry, SU(4)-color and left-right symmetry. Thus they point to the relevance of an effective string-unified G(224) or SO(10)-symmetry.

A concrete proposal is presented, within a predictive SO(10)/G(224)-framework, that successfully describes the masses and mixings of all fermions, including the neutrinos - with eight predictions, all in agreement with observation. Within this framework, a systematic study of proton decay is carried out, which pays special attention to its dependence on the fermion masses, including the superheavy Majorana masses of the right-handed neutrinos. The study shows that a conservative upper limit on the proton lifetime is about $(1/2 - 1) \times 10^{34}$ yrs, with $\pi K^+$ being the dominant decay mode, and as a distinctive feature, $\mu^+ K^0$ being prominent. This in turn strongly suggests that an improvement in the current sensitivity by a factor of five to ten (compared to SuperK) ought to reveal proton decay. Otherwise some promising and remarkably successful ideas on unification would suffer a major setback.

1 Introduction

It has been recognized since the early 1970’s that the price one must pay to achieve a unification of quarks and leptons and simultaneously a unity of the three gauge forces, commonly called “grand unification”, is proton decay [1, 2, 3, 4]. This important process,

which would provide the window for viewing physics at truly short distances \( (< 10^{-30} \text{ cm}) \), is yet to be seen. Nevertheless, as I will stress in this talk, there have appeared over the years an impressive set of facts, including the meeting of the gauge couplings and neutrino-oscillations, which not only favor the hypothesis of grand unification, but in fact select out a particular route to such unification, based on the ideas of supersymmetry \([5]\) and SU(4)-color \([2]\). These facts together provide a clear signal that the discovery of proton decay cannot be far behind.

To be specific, working within the framework of a unified theory \([6]\), that incorporates the ideas mentioned above, I would argue that an improvement in the current sensitivity for detecting proton decay by a modest factor of five to ten should either produce real events, or else the framework would be excluded. In this sense, and as I will elaborate further, the discovery of proton decay is now crucial to the survival of some elegant ideas on unification, which are otherwise so successful. By the same token, proving or disproving their prediction on proton decay poses a fresh challenge to experiment.

The pioneering efforts by several physicists \([7]\) in the mid 1950’s through the early 70’s had provided a lower limit on the proton lifetime of about \(10^{26} \text{ yrs}\), independent of decay modes, and \(10^{29}-10^{30} \text{ yrs}\) for the \(e^+\pi^0\)-mode. Subsequent searches at the Kolar Goldfield and the NUSEX detectors in the early 80’s \([8]\) pushed this limit to about \(10^{31} \text{ yrs}\) in the \(e^+\pi^0\)-mode. Following the suggestion of proton decay in the context of grand unification, and thanks to the initiative of several experimenters, two relatively large-size detectors - IMB and Kamiokande - were built in the 80’s, where dedicated searches for proton decay were carried out with higher sensitivity. These detectors helped to push the lower limit in the \(e^+\pi^0\) channel to about \(10^{32} \text{ yrs}\). This in turn clearly disfavored the minimal non-supersymmetric SU(5)-model of grand unification \([3]\) - a conclusion that was strengthened subsequently by the measurements of the gauge couplings at LEP as well (see discussion later).

The searches for proton decay now continues with still greater sensitivity at the largest detector so far - at SuperKamiokande, completed in 1996. It is worth noting at this point that, although these detectors have not revealed proton decay yet, they did bring some major bonuses of monumental importance. These include : (a) the detection of the neutrinos from the supernova 1987a, (b) confirmation of the solar neutrino-deficit, and last but not least, (c) the discovery of atmospheric neutrino-oscillation. The SuperK water-Cerenkov detector with a fiducial volume of 22.5 kilotons currently provides (with three years of running) a lower limit on the inverse rate of proton decay of about \(1.6 \times 10^{33} \text{ yrs}\) for the theoretically favored \((\pi K^+)\)-channel \([9]\) and of about \(3.8 \times 10^{33} \text{ yrs}\) for the \((e^+\pi^0)\)-channel \([10]\). It has the capability of improving these limits by a factor of two to three in each case within the next decade, unless of course it discovers real events for proton decay or strong candidate events in the meantime. I will return to this point and its relevance to theoretical expectations for proton decay in just a bit.

While proton decay is yet to be observed, it is worth stressing at this point, that the hypothesis of grand unification, especially that based on the ideas of SU(4)-color, left-right gauge symmetry, and supersymmetry, is now supported by several observations. As I will explain in sections 2-5, these include :

(a) **The observed family structure** : The five scattered multiplets of the standard model, belonging to a family, neatly become parts of a whole (a single multiplet), with their
weak hypercharges precisely predicted by grand unification. Realization of this feature calls for an extension of the standard model symmetry $G(213) = SU(2)_L \times U(1)_Y \times SU(3)_C$ minimally to the symmetry group $G(224) = SU(2)_L \times SU(2)_R \times SU(4)_C$ [2], which can be extended further into the simple group SO(10) [11], but not SU(5) [3]. The G(224) symmetry in turn introduces some additional attractive features (see Sec. 2), including especially the right-handed (RH) neutrinos ($\nu_R$’s) accompanying the left-handed ones ($\nu_L$’s), and $B-L$ as a local symmetry. As we will see, both of these features, which are special to G(224), now seem to be needed on empirical grounds.

(b) Meeting of the gauge couplings : Such a meeting is found to occur at a scale $M_X \approx 2 \times 10^{16}$ GeV, when the three gauge couplings are extrapolated from their values measured at LEP to higher energies, in the context of supersymmetry [12]. This dramatic phenomenon supports the ideas of both grand unification and supersymmetry. These in turn may well emerge from a string theory [13] or M-theory [14] (see discussion in Sec. 3).

(c) Mass of $\nu_r \sim 1/20$ eV : Subject to the well-motivated assumption of hierarchical neutrino masses, the recent discovery of atmospheric neutrino-oscillation at SuperKamiokande [15] suggests a value for $m(\nu_r) \sim 1/20eV$. It has been argued (see e.g. Ref. [16]) that a mass for $\nu_r$ of this magnitude points to the need for RH neutrinos, and that it goes extremely well with the hypothesis of a supersymmetric unification, based on either a string-unified G(224) symmetry or SO(10). The SUSY unification-scale as well as SU(4)-color play crucial roles in making this argument.

(d) Some intriguing features of fermion masses and mixings : These include : (i) the ”observed” near equality of the masses of the b-quark and the $\tau$-lepton at the unification-scale ($m_0^b \approx m_0^\tau$); (ii) the empirical Georgi-Jarlskog relations: $m_0^u \sim m_0^d/3$ and $m_0^e \sim 3m_0^\mu$, and (iii) the observed largeness of the $\nu_\mu$-$\nu_\tau$ oscillation angle ($\sin^2 2\theta_{osc} \geq 0.83$) [15], together with the smallness of the corresponding quark mixing parameter $V_{bc}(\approx 0.04)$ [17]. As shown in recent work by Babu, Wilczek and me [6], it turns out that these features and more can be understood remarkably well (see discussion in Sec 5) within an economical and predictive SO(10)-framework based on a minimal Higgs system. The success of this framework is in large part due simply to the group-structure of SO(10). For most purposes, that of G(224) suffices.

(e) Baryogenesis : To implement baryogenesis [18] successfully, in the presence of electroweak sphaleron effects [19], which wipe out any baryon excess generated at high temperatures in the $(B-L)$-conserving mode, it has become apparent that one would need $B-L$ as a generator of the underlying symmetry, whose spontaneous violation at high temperatures would yield, for example, lepton asymmetry (leptogenesis). The latter in turn is converted to baryon-excess at lower temperatures by electroweak sphalerons. This mechanism, it turns out, yields even quantitatively the right magnitude for baryon excess [20]. The need for $B-L$, which is a generator of SU(4)-color, again points to the need for G(224) or SO(10) as an effective symmetry near the unification-scale $M_X$.

The success of each of these five features (a)-(e) seems to be non-trivial. Together they make a strong case for both supersymmetric grand unification and simultaneously for the G(224)/SO(10)-route to such unification, as being relevant to nature. However, despite these successes, as long as proton decay remains undiscovered, the hallmark of grand unification - that is quark-lepton transformability - would remain unrevealed.
The relevant questions in this regard then are: What is the predicted range for the lifetime of the proton - in particular an upper limit - within the empirically favored route to unification mentioned above? What are the expected dominant decay modes within this route? Are these predictions compatible with current lower limits on proton lifetime mentioned above, and if so, can they still be tested at the existing or possible near-future detectors for proton decay?

Fortunately, we are in a much better position to answer these questions now, compared to a few years ago, because meanwhile we have learnt more about the nature of grand unification. As noted above (see also Secs. 2 and 4), the neutrino masses and the meeting of the gauge couplings together seem to select the supersymmetric G(224)/SO(10)-route to higher unification. The main purpose of my talk here will therefore be to address the questions raised above, in the context of this route. For the sake of comparison, however, I will state the corresponding results for the case of supersymmetric SU(5) as well.

My discussion will be based on a recent study of proton decay by Babu, Wilczek and me [6], which, relative to previous ones, has three distinctive features:

(a) It systematically takes into account the link that exists between proton decay and the masses and mixings of all fermions, including the neutrinos.

(b) In particular, in addition to the contributions from the so-called “standard” $d = 5$ operators [22] (see Sec. 6), it includes those from a new set of $d = 5$ operators, related to the Majorana masses of the RH neutrinos [21]. These latter are found to be as important as the standard ones.

(c) The work also incorporates GUT-scale threshold effects, which arise because of mass-splittings between the components of the SO(10)-multiplets, and lead to differences between the three gauge couplings.

Each of these features turn out to be crucial to gaining a reliable insight into the nature of proton decay. Our study shows that the inverse decay rate for the $\pi K^+$-mode, which is dominant, is less than about $7 \times 10^{33}$ yrs. This upper bound is obtained by making generous allowance for uncertainties in the matrix elements and the SUSY-spectrum. Typically, the lifetime should of course be less than this bound. Furthermore, due to contributions from the new operators, the $\mu^+ K^0$-mode is found to be prominent, with a branching ratio typically in the range of 10-50%. By contrast, minimal SUSY SU(5), for which the new operators are absent, would lead to branching ratios $\leq 10^{-3}$ for this mode. Thus our study of proton decay, correlated with fermion masses, strongly suggests that discovery of proton decay should be around the corner. In fact, one expects that at least candidate events should be observed in the near future already at SuperK. However, allowing for the possibility that the proton lifetime may well be closer to the upper bound stated above, a next-generation detector providing a net gain in sensitivity in proton decay-searches by a factor of 5-10, compared to SuperK, would certainly be needed not just to produce proton-decay events, but also to clearly distinguish them from the background. It would of course also be essential to study the branching ratios of certain sub-dominant but crucial decay modes, such as the $\mu^+ K^0$. The importance of such improved sensitivity, in the light of the successes of supersymmetric grand unification, is emphasized at the end.
2 Advantages of the Symmetry G(224) as a Step to Higher Unification

The standard model (SM) based on the gauge symmetry $G(213) = \text{SU}(2)_L \times \text{U}(1)_Y \times \text{SU}(3)_C$ has turned out to be extremely successful empirically. It has however been recognized since the early 1970’s that, judged on aesthetic merits, it has some major shortcomings. For example, it puts members of a family into five scattered multiplets, without providing a compelling reason for doing so. It also does not provide a fundamental reason for the quantization of electric charge. Nor does it explain the co-existence of quarks and leptons, and that of the three gauge forces, with their differing strengths. The idea of grand unification was postulated precisely to remove these shortcomings. That in turn calls for the existence of fundamentally new physics, far beyond that of the standard model. As mentioned before, recent experimental findings, including the meetings of the gauge couplings and neutrino-oscillations, seem to go extremely well with this line of thinking.

To illustrate the advantage of an early suggestion in this regard, consider the five standard model multiplets belonging to the electron-family as shown:

$$\begin{pmatrix} u_r & u_y & u_b \\ d_r & d_y & d_b \end{pmatrix}^\frac{2}{3}_L; \begin{pmatrix} u_r & u_y & u_b \\ d_r & d_y & d_b \end{pmatrix}^\frac{2}{3}_L; \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}^{\frac{1}{2}}_L; \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}^{\frac{1}{2}}_L.$$

Here the superscripts denote the respective weak hypercharges $Y_W$ (where $Q_{em} = I_{3L} + Y_W/2$) and the subscripts L and R denote the chiralities of the respective fields. If one asks: how one can put these five multiplets into just one multiplet, the answer turns out to be simple and unique. As mentioned in the introduction, the minimal extension of the SM symmetry $G(213)$ needed, to achieve this goal, is given by the gauge symmetry $G(224)$:

$$G(224) = \text{SU}(2)_L \times \text{SU}(2)_R \times \text{SU}(4)^C.$$

Subject to left-right discrete symmetry $(L \leftrightarrow R)$, which is natural to $G(224)$, all members of the electron family fall into the neat pattern:

$$F_{e,L,R} = \begin{bmatrix} u_r & u_y & u_b & \nu_e \\ d_r & d_y & d_b & e^- \end{bmatrix}_{L,R}.$$

The multiplets $F_{e,L}^e$ and $F_{e,R}^e$ are left-right conjugates of each other and transform respectively as $(2,1,4)$ and $(1,2,4)$ of $G(224)$; likewise for the muon and the tau families. Note that the symmetries $\text{SU}(2)_L$ and $\text{SU}(2)_R$ are just like the familiar isospin symmetry, except that they operate on quarks and well as leptons, and distinguish between left and right chiralities. The left weak-isospin $\text{SU}(2)_L$ treats each column of $F_{e,L}^e$ as a doublet; likewise $\text{SU}(2)_R$ for $F_{e,R}^e$; the symmetry $\text{SU}(4)$-color treats each row of $F_{e,L}^e$ and $F_{e,R}^e$ as a quartet, interpreting lepton number as the fourth color. Note also that postulating either $\text{SU}(4)$-color or $\text{SU}(2)_R$ forces one to introduce a right-handed neutrino $(\nu_R)$ for each family as a singlet of the SM symmetry. This requires that there be sixteen two-component fermions in each family, as opposed to fifteen for the SM. The symmetry $G(224)$ introduces an elegant charge formula:

$$Q_{em} = I_{3L} + I_{3R} + \frac{B - L}{2}$$
expressed in terms of familiar quantum numbers \(I_{3L}, I_{3R}\) and \(B-L\), which applies to all forms of matter (including quarks and leptons of all six flavors, gauge and Higgs bosons). Note that the weak hypercharge given by \(Y_W = I_{3R} + \frac{B-L}{2}\) is now completely determined for all members of the family. The values of \(Y_W\) thus obtained precisely match the assignments shown in Eq. (1). Quite clearly, the charges \(I_{3L}, I_{3R}\) and \(B-L\), being generators respectively of \(SU(2)_L\), \(SU(2)_R\) and \(SU(4)^c\), are quantized; so also then is the electric charge \(Q_{em}\).

In brief, the symmetry \(G(224)\) brings some attractive features to particle physics. These include:

(i) Organization of all 16 members of a family into one left-right self-conjugate multiplet;
(ii) Quantization of electric charge;
(iii) Quark-lepton unification (through \(SU(4)\) color);
(iv) Conservation of parity at a fundamental level [2, 23];
(v) Right-handed neutrinos (\(\nu^R\)) as a compelling feature; and
(vi) \(B-L\) as a local symmetry.

As mentioned in the introduction, the two distinguishing features of \(G(224)\) - i.e. the existence of the RH neutrinos and \(B-L\) as a local symmetry - now seem to be needed on empirical grounds.

Believing in a complete unification, one is led to view the \(G(224)\) symmetry as part of a bigger symmetry, which itself may have its origin in an underlying theory, such as string theory. In this context, one might ask: Could the effective symmetry below the string scale in four dimensions (see sec.3) be as small as just the SM symmetry \(G(213)\), even though the latter may have its origin in a bigger symmetry, which lives however only in higher dimensions? I will argue in Sec. 4 that the data on neutrino masses and the need for baryogenesis provide an answer to the contrary, suggesting clearly that it is the effective symmetry in four dimensions, below the string scale, which must minimally contain either \(G(224)\) or a close relative \(G(214) = SU(2) \times I_{3R} \times SU(4)^C\).

One may also ask: does the effective four dimensional symmetry have to be any bigger than \(G(224)\) near the string scale? In preparation for an answer to this question, let us recall that the smallest simple group that contains the SM symmetry \(G(213)\) is \(SU(5)\) [3]. It has the virtue of demonstrating how the main ideas of grand unification, including unification of the gauge couplings, can be realized. However, \(SU(5)\) does not contain \(G(224)\) as a subgroup. As such, it does not possess some of the advantages listed above. In particular, it does not contain the RH neutrinos as a compelling feature, and \(B-L\) as a local symmetry. Furthermore, it splits members of a family into two multiplets: \(\bar{5} + 10\).

By contrast, the symmetry \(SO(10)\) has the merit, relative to \(SU(5)\), that it contains \(G(224)\) as a subgroup, and thereby retains all the advantages of \(G(224)\) listed above. (As a historical note, it is worth mentioning that these advantages had been motivated on aesthetic grounds through the symmetry \(G(224)\) [2], and all the ideas of higher unification were in place [1, 2, 3], before it was noted that \(G(224)\) (isomorphic to \(SO(4) \times SO(6)\)) embeds nicely into \(SO(10)\) [11]). Now, \(SO(10)\) even preserves the 16-plet family-structure of \(G(224)\) without a need for any extension. By contrast, if one extends \(G(224)\) to the still higher symmetry \(E_6\) [24], the advantages (i)-(vi) are retained, but in this case, one must extend the family-structure from a 16 to a 27-plet, by postulating additional fermions. In this sense, there seems to be some advantage in having the effective symmetry below the string scale to be
minimally G(224) (or G(214)) and maximally no more than SO(10). I will compare the relative advantage of having either a string-derived G(224) or a string-SO(10), in the next section. First, I discuss the implications of the data on coupling unification.

3 The Need for Supersymmetry: MSSM versus String Unifications

It has been known for some time that the precision measurements of the standard model coupling constants (in particular $\sin^2 \theta_W$) at LEP put severe constraints on the idea of grand unification. Owing to these constraints, the non-supersymmetric minimal SU(5), and for similar reasons, the one-step breaking minimal non-supersymmetric SO(10)-model as well, are now excluded [25]. But the situation changes radically if one assumes that the standard model is replaced by the minimal supersymmetric standard model (MSSM), above a threshold of about 1 TeV. In this case, the three gauge couplings are found to meet [12], at least approximately, provided $\alpha_3(m_Z)$ is not too low (see Figs. in e.g. Refs. [23, 13]). Their scale of meeting is given by

$$M_X \approx 2 \times 10^{16} \text{GeV} \ (\text{MSSM or SUSY SU(5)})$$

This dramatic meeting of the three gauge couplings, or equivalently the agreement of the MSSM-based prediction of $\sin^2 \theta_W(m_Z)_{\text{Th}} = 0.2315 \pm 0.003$ [26] with the observed value of $\sin^2 \theta_W(m_Z) = 0.23124 \pm 0.00017$ [17], provides a strong support for the ideas of both grand unification and supersymmetry, as being relevant to physics at short distances.

The most straightforward interpretation of the observed meeting of the three couplings and of the scale $M_X$, is that a supersymmetric grand unification symmetry (often called GUT symmetry), like SU(5) or SO(10), breaks spontaneously at $M_X$ into the standard model symmetry G(213). In the context of string or M theory, which seems to be needed to unify all the forces of nature including gravity and also to obtain a good quantum theory of gravity, an alternative interpretation is however possible. This is because, even if the effective symmetry in four dimensions emerging from a higher dimensional string theory is non-simple, like G(224) or G(213), string theory can still ensure familiar unification of the gauge couplings at the string scale. In this case, however, one needs to account for the small mismatch between the MSSM unification scale $M_X$ (given above), and the string unification scale, given by $M_{st} \approx g_{st} \times 5.2 \times 10^{17} \text{GeV} \approx 3.6 \times 10^{17} \text{GeV}$ (Here we have put $\alpha_{st} = \alpha_{GUT}(\text{MSSM}) \approx 0.04$) [27]. Possible resolutions of this mismatch have been proposed. These include: (i) utilizing the idea of string-duality [28] which allows a lowering of $M_{st}$ compared to the value shown above, or alternatively (ii) the idea of a semi-perturbative unification that assumes the existence of two vector-like families, transforming as $(16 + \overline{16})$, at the TeV-scale. The latter raises $\alpha_{GUT}$ to about 0.25-0.3 and simultaneously $M_X$, in two loop, to about $(1/2 - 2) \times 10^{17} \text{GeV}$ [29] (Other mechanisms resolving the mismatch are reviewed in Refs. [30] and [31]). In practice, a combination of the two mechanisms mentioned above may well be relevant.\footnote{I have in mind the possibility of string-duality [28] lowering $M_{st}$ for the case of semi-perturbative uni-}
While the mismatch can thus quite plausibly be removed for a non-GUT string-derived symmetry like G(224) or G(213), a GUT symmetry like SU(5) or SO(10) would have an advantage in this regard because it would keep the gauge couplings together between $M_{sl}$ and $M_X$ (even if $M_X \sim M_{sl}/20$), and thus not even encounter the problem of a mismatch between the two scales. A supersymmetric GUT-solution (like SU(5) or SO(10)), however, has a possible disadvantage as well, because it needs certain color triplets to become superheavy by the so-called double-triplet splitting mechanism (see Sec. 6 and Appendix), in order to avoid the problem of rapid proton decay. However, no such mechanism has emerged yet, in string theory, for the GUT-like solutions [32].

Non-GUT string solutions, based on symmetries like G(224) or G(2113) for example, have a distinct advantage in this regard, in that the dangerous color triplets, which would induce rapid proton decay, are often naturally projected out for such solutions [33, 34]. Furthermore, the non-GUT solutions invariably possess new “flavor” gauge symmetries, which distinguish between families. These symmetries are immensely helpful in explaining qualitatively the observed fermion mass-hierarchy (see e.g., Ref. [34]) and resolving the so-called naturalness problems of supersymmetry such as those pertaining to the issues of squark-degeneracy [35], CP violation [36] and quantum gravity-induced rapid proton decay [37].

Weighing the advantages and possible disadvantages of both, it seems hard at present to make a priori a clear choice between a GUT versus a non-GUT string-solution. As expressed elsewhere [31], it therefore seems prudent to keep both options open and pursue their phenomenological consequences. Given the advantages of G(224) or SO(10) in the light of the neutrino masses (see Secs. 2 and 4), I will thus proceed by assuming that either a suitable G(224)-solution with a mechanism of the sort mentioned above, or a realistic SO(10)-solution with the needed doublet-triplet mechanism, will emerge from string theory. We will see that with this broad assumption an economical and predictive framework emerges, which successfully accounts for a host of observed phenomena, and makes some crucial testable predictions. Fortunately, it will turn out that there are many similarities between the predictions of a string-unified G(224) and SO(10), not only for the neutrino and the charged fermion masses, but also for proton decay. I next discuss the implications of the mass of $\nu_\tau$ suggested by the SuperK data.

## 4 Mass of $\nu_\tau$: Evidence In Favor of the G(224) Route

One can obtain an estimate for the mass of $\nu_\tau$ in the context of G(224) or SO(10) by using the following three steps (see e.g., Ref. [16]):

(i) Assume that B−L and $I_{3R}$, contained in a string-derived G(224) or SO(10), break

ification (for which $\alpha_{sl} \approx 0.25$, and thus, without the use of string-duality, $M_{sl}$ would be about $10^{18}$ GeV) to a value of about $(1-2)\times10^{17}$ GeV (say), and semi-perturbative unification [29] raising the MSSM value of $M_X$ to about $5\times10^{16}$ GeV $\approx M_{sl}/(1/2$ to $1/4)$ (say). In this case, an intermediate symmetry like G(224) emerging at $M_{sl}$ would be effective only within the short gap between $M_{sl}$ and $M_X$, where it would break into G(213). Despite this short gap, one would still have the benefits of SU(4)-color that are needed to understand neutrino masses (see sec.4). At the same time, Since the gap is so small, the couplings of G(224), unified at $M_{sl}$, would remain essentially so at $M_X$, so as to match with the “observed” coupling unification, of the type suggested in Ref. [29].
near the unification-scale:

$$M_X \sim 2 \times 10^{16} \text{ GeV},$$  \hfill (6)

through VEVs of Higgs multiplets of the type suggested by string-solutions - i.e. $\langle (1, 2, 4)_H \rangle$ for G(224) or $\langle \overline{10}_H \rangle$ for SO(10), as opposed to $126_H$ \cite{38}. In the process, the RH neutrinos ($\nu_R^i$), which are singlets of the standard model, can and generically will acquire superheavy Majorana masses of the type $M_{ij}^2 \nu_R^i \nu_R^j$ by utilizing the VEV of $\langle \overline{10}_H \rangle$ and effective couplings of the form:

$$L_M (SO(10)) = f_{ij} \ 16_i \cdot 16_j \ \overline{16}_H \cdot \overline{16}_H / M + h.c.$$

A similar expression holds for G(224). Here $i, j = 1, 2, 3$, correspond respectively to $e$, $\mu$ and $\tau$ families. Such gauge-invariant non-renormalizable couplings might be expected to be induced by Planck-scale physics, involving quantum gravity or stringy effects and/or tree-level exchange of superheavy states, such as those in the string tower. With $f_{ij}$ (at least the largest among them) being of order unity, we would thus expect $M$ to lie between $M_{\text{Planck}} \approx 2 \times 10^{18}$ GeV and $M_{\text{string}} \approx 4 \times 10^{17}$ GeV. Ignoring for the present off-diagonal mixings (for simplicity), one thus obtains $^2$:

$$M_{3R} \approx \frac{f_{33}(\overline{16}_H)^2}{M} \approx f_{33} (2 \times 10^{14} \text{ GeV}) \eta^2 (M_{\text{Planck}}/M)$$  \hfill (7)

This is the Majorana mass of the RH tau neutrino. Guided by the value of $M_X$, we have substituted $\langle \overline{10}_H \rangle = (2 \times 10^{16} \text{ GeV}) \eta$, with $\eta \approx 1/2$ to 2(say).

(iii) Now using SU(4)-color and the Higgs multiplet $(2, 2, 1)_H$ of G(224) or equivalently $10_H$ of SO(10), one obtains the relation $m_{\tau}(M_X) = m_\nu(M_X)$, which is known to be successful. Thus, there is a good reason to believe that the third family gets its masses primarily from the $10_H$ or equivalently $(2, 2, 1)_H$ (see sec.5). In turn, this implies:

$$m(\nu_{\text{Dirac}}^\tau) \approx m_{\text{top}}(M_X) \approx (100 - 120) \text{ GeV}$$  \hfill (9)

Note that this relationship between the Dirac mass of the tau-neutrino and the top-mass is special to SU(4)-color. It does not emerge in SU(5).

(ii) Given the superheavy Majorana masses of the RH neutrinos as well as the Dirac masses as above, the see-saw mechanism \cite{39} yields naturally light masses for the LH neutrinos. For $\nu_L^\tau$ (ignoring mixing), one thus obtains, using Eqs. (8) and (9),

$$m(\nu_L^\tau) \approx \frac{m(\nu_{\text{Dirac}}^\tau)^2}{M_{3R}} \approx [(1/20) \text{ eV} (1 - 1.44) / f_{33} \eta^2] (M/M_{\text{Planck}})$$  \hfill (10)

Now, assuming the hierarchical pattern $m(\nu_L^\tau) \ll m(\nu_L^\mu) \ll m(\nu_L^e)$, which is suggested by the see-saw mechanism, and further that the SuperK observation represents $\nu_L^\mu - \nu_L^\tau$ (rather than $\nu_L^\mu - \nu_L^e$) oscillation, the observed $\delta m^2 \approx 1/2 (10^{-2} - 10^{-3}) \text{ eV}^2$ corresponds to $m(\nu_L^\tau) \approx (1/15 - 1/40) \text{ eV}$. It seems truly remarkable that the expected magnitude of $m(\nu_L^\tau)$, given by

$^2$The effects of neutrino-mixing and of possible choice of $M = M_{\text{string}} \approx 4 \times 10^{17}$ GeV (instead of $M = M_{\text{Planck}}$) on $M_{3R}$ are considered in Ref. \cite{6}.
Eq. (10), is just about what is suggested by the SuperK data, if $f_{33} \eta^2 (M_{\text{Plank}}/M) \approx 1.3$ to 1/2. Such a range for $f_{33} \eta^2 (M_{\text{Plank}}/M)$ seems most plausible and natural (see discussion in Ref. [16]). Note that the estimate (10) crucially depends upon the supersymmetric unification scale, which provides a value for $M_{3R}$, as well as on SU(4)-color that yields $m(\nu^\text{Dirac})$. The agreement between the expected and the SuperK result thus clearly suggests that the effective symmetry below the string-scale should contain SU(4)-color. Thus, minimally it should be either G(214) or G(224), and maximally as big as SO(10), if not Eq.

By contrast, if SU(5) is regarded as either a fundamental symmetry or as the effective symmetry below the string scale, there would be no compelling reason based on symmetry alone, to introduce a $\nu_R$, because it is a singlet of SU(5). Second, even if one did introduce $\nu_R$ by hand, their Dirac masses, arising from the coupling $h_i^5 \langle 5_H \rangle \nu^i_R$, would be unrelated to the up-flavor masses and thus rather arbitrary (contrast with Eq. (9)). So also would be the Majorana masses of the $\nu^i_R$'s, which are SU(5)-invariant, and thus can be even of order string scale. This would give $m(\nu^i_L)$ in gross conflict with the observed value.

Before passing to the next section, it is worth noting that the mass of $\nu_\tau$ suggested by SuperK, as well as the observed value of $\sin^2 \theta_W$ (see Sec.3), provide valuable insight into the nature of GUT symmetry breaking. They both favor the case of a single-step breaking (SSB) of SO(10) or a string-unified G(224) symmetry at a scale of order $M_X$, into the standard model symmetry G(213), as opposed to that of a multi-step breaking (MSB). The latter would correspond, for example, to SO(10) (or G(224)) breaking at a scale $M_1$ into G(2213), which in turn breaks at a scale $M_2 << M_1$ into G(213). One reason why the case of single-step breaking is favored over that of multi-step breaking is that the latter can accommodate but not really predict $\sin^2 \theta_W$, whereas the former predicts the same successfully. Furthermore, since the Majorana mass of $\nu^i_R$ arises arises only after $B-L$ and $I_{3R}$ break, it would be given, for the case of MSB, by $M_{3R} \sim f_{33} (M_2^2/M)$, where $M \sim M_{\text{st}}$ (say). If $M_2 \ll M_X \sim 2 \times 10^{16}$ GeV, and $M > M_X$, one would obtain too low a value ($<< 10^{14}$ GeV) for $M_{3R}$ (compare with Eq.(8)), and thereby too large a value for $m(\nu^i_L)$, compared to that suggested by SuperK. By contrast, the case of SSB yields the right magnitude for $m(\nu_\tau)$ (see Eq. (10)).

Thus the success of the result on $m(\nu_\tau)$ discussed above not only favors the symmetry G(224) or SO(10), but also clearly suggests that $B-L$ and $I_{3R}$ break near the conventional GUT scale $M_X \sim 2 \times 10^{16}$ GeV, rather than at an intermediate scale $<< M_X$. In other words, the observed values of both $\sin^2 \theta_W$ and $m(\nu_\tau)$ favor only the simplest pattern of symmetry-breaking, for which SO(10) or a string-derived G(224) symmetry breaks in one step to the standard model symmetry, rather than in multiple steps. It is of course only this simple pattern of symmetry breaking that would be rather restrictive as regards its predictions for proton decay (to be discussed in Sec.6). I next discuss the problem of understanding the masses and mixings of all fermions.

### 5 Understanding Fermion Masses and Neutrino Oscillations in SO(10)

Understanding the masses and mixings of all quarks and charged leptons, in conjunction with those of the neutrinos, is a goal worth achieving by itself. It also turns out to be essential
for the study of proton decay. I therefore present first a recent attempt in this direction, which seems most promising [6]. A few guidelines would prove to be helpful in this regard. The first of these is motivated by the desire for economy and the rest by data.

1) Hierarchy Through Off-diagonal Mixings: Recall earlier attempts [40] that attribute hierarchical masses of the first two families to matrices of the form:

\[
M = \begin{pmatrix}
0 & \epsilon \\
\epsilon & 1
\end{pmatrix} m_s^{(0)}, \tag{11}
\]

for the \((d, s)\) quarks, and likewise for the \((u, c)\) quarks. Here \(\epsilon \sim 1/10\). The hierarchical patterns in Eq. (11) can be ensured by imposing a suitable flavor symmetry which distinguishes between the two families (that in turn may have its origin in string theory (see e.g. Ref [34]). Such a pattern has the virtues that (a) it yields a hierarchy that is much larger than the input parameter \(\epsilon: (m_d/m_s) \approx \epsilon^2 \ll \epsilon\), and (b) it leads to an expression for the cabibbo angle:

\[
\theta_c \approx \left| \sqrt{\frac{m_d}{m_s}} - e^{i\phi} \sqrt{\frac{m_u}{m_c}} \right|, \tag{12}
\]

which is rather successful. Using \(\sqrt{m_d/m_s} \approx 0.22\) and \(\sqrt{m_u/m_c} \approx 0.06\), we see that Eq. (12) works to within about 25% for any value of the phase \(\phi\). Note that the square root formula (like \(\sqrt{m_d/m_s}\)) for the relevant mixing angle arises because of the symmetric form of \(M\) in Eq. (11), which in turn is ensured if the contributing Higgs is a 10 of SO(10). A generalization of the pattern in Eq. (11) would suggest that the first two families (i.e. the \(e\) and the \(\mu\)) receive masses primarily through their mixing with the third family \((\tau)\), with \((1, 3)\) and \((1, 2)\) elements being smaller than the \((2, 3)\); while \((2, 3)\) is smaller than the \((3, 3)\). We will follow this guideline, except for the modification noted below.

2) The Need for an Antisymmetric Component: Although the symmetric hierarchical matrix in Eq. (11) works well for the first two families, a matrix of the same form fails altogether to reproduce \(V_{cb}\), for which it yields:

\[
V_{cb} \approx \left| \sqrt{\frac{m_s}{m_b}} - e^{i\chi} \sqrt{\frac{m_c}{m_t}} \right|. \tag{13}
\]

Given that \(\sqrt{m_s/m_b} \approx 0.17\) and \(\sqrt{m_c/m_t} \approx 0.06\), we see that Eq. (13) would yield \(V_{cb}\) varying between 0.11 and 0.23, depending upon the phase \(\chi\). This is too big, compared to the observed value of \(V_{cb} \approx 0.04 \pm 0.003\), by at least a factor of 3. We interpret this failure as a clue to the presence of an antisymmetric component in \(M\), together with symmetrical ones (thus \(m_{ij} \neq m_{ji}\)), which would modify the relevant mixing angle to \(\sqrt{\frac{m_i}{m_j}} \sqrt{\frac{m_{ij}}{m_{ji}}}\), where \(m_i\) and \(m_j\) denote the respective eigenvalues.

3) The Need for a Contribution Proportional to \(B-L\): The success of the relations \(m_b^0 \approx m_{\tau}^0\), and \(m_i^0 \approx m(\nu_i)^0_{\text{Dirac}}\) (see Sec. 4), suggests that the members of the third family get their masses primarily from the VEV of a SU(4)-color singlet Higgs field that is independent of \(B-L\). This is in fact ensured if the Higgs is a 10 of SO(10). However, the empirical observations of \(m_s^0 \sim m_{\mu}^0/3\) and \(m_d^0 \sim 3m_c^0\) [41] clearly call for a contribution
proportional to $B-L$ as well. Further, one can in fact argue that the suppression of $V_{bc}$ (in the quark-sector) together with an enhancement of $\theta_{\nu_e}^{osc}$ (in the lepton sector) calls for a contribution that is not only proportional to $B-L$ but is also antisymmetric in the family space (as suggested above in item (2)). We note below how both of these requirements can be met, rather easily, in SO(10), even for a minimal Higgs system.

4) Up-Down Asymmetry : Finally, the up and the down-sector mass matrices must not be proportional to each other, as otherwise the CKM angles would all vanish.

Following Ref. [6], I now present a simple and predictive mass-matrix, based on SO(10), that satisfies all three requirements, (2), (3) and (4). The interesting point is that one can obtain such a mass-matrix for the fermions by utilizing only the minimal Higgs system, that is needed anyway to break the gauge symmetry SO(10). It consists of the set:

$$H_{\text{minimal}} = \{45_H, 16_H, \overline{16}_H, 10_H\}. \quad (14)$$

Of these, the VEV of $\langle 45_H \rangle \sim M_X$ breaks SO(10) into G(2213), and those of $\langle 16_H \rangle = \langle \overline{16}_H \rangle \sim M_X$ break G(2213) to G(213), at the unification-scale $M_X$. Now G(213) breaks at the electroweak scale by the VEV of $\langle 10_H \rangle$ to $U(1)_{em} \times SU(3)^c$.

One might have introduced large-dimensional tensorial multiplets of SO(10) like 126 and 120, both of which possess cubic level Yukawa couplings with the fermions. In particular, the coupling $16_i 16_j (120_H)$ would give the desired family-antisymmetric as well as $(B-L)$-dependent contribution. We do not however introduce these multiplets in part because they do not seem to arise in string solutions [38], and in part also because mass-splittings within such large-dimensional multiplets tend to give excessive threshold corrections to $\alpha_3(m_z)$ (typically exceeding 20%), rendering observed coupling unification fortuitous. By contrast, the multiplets in the minimal set (shown above) do arise in string solutions leading to SO(10). Furthermore, the threshold corrections for the minimal set are found to be naturally small, and even to have the right sign, to go with the observed coupling unification [6].

The question is : does the minimal set meet all the requirements listed above? Now $10_H$ (even several 10's) can not meet the requirements of antisymmetry and $(B-L)$-dependence. Furthermore, a single $10_H$ cannot generate CKM-mixings. This impasse disappears, however, as soon as one allows for not only cubic, but also effective non-renormalizable quartic couplings of the minimal set of Higgs fields with the fermions. These latter couplings could of course well arise through exchanges of superheavy states (e.g. those in the string tower) involving renormalizable couplings, and/or through quantum gravity.

Allowing for such cubic and quartic couplings and adopting the guideline (1) of hierarchical Yukawa couplings, as well as that of economy, we are led to suggest the following effective lagrangian for generating Dirac masses and mixings of the three families [6] (for a related but different pattern, involving a non-minimal Higgs system, see Ref [42]).

$$L_{\text{Yuk}} = h_{33} 16_3 16_3 10_H + [h_{23} 16_2 16_3 10_H + a_{23} 16_2 16_3 10_H 45_H/M + g_{23} 16_2 16_3 16_H 16_H/M] + \{a_{12} 16_1 16_2 10_H 45_H/M + g_{12} 16_1 16_2 16_H 16_H/M\}. \quad (15)$$

Here, $M$ could plausibly be of order string scale. Note that a mass matrix having essentially the form of Eq. (11) results if the first term $h_{33}(10_H)$ is dominant. This ensures $m_b^0 \approx$
$m_t^0$ and $m_t^0 \approx m(\nu_{Dirac})^0$. Following the assumption of progressive hierarchy (equivalently appropriate flavor symmetries\(^7\)), we presume that $h_{23} \sim h_{33}/10$, while $h_{22}$ and $h_{11}$, which are set to be zeros, are progressively much smaller than $h_{23}$ (see discussion in Ref. [31]). Since $\langle 45_H \rangle \sim \langle 16_H \rangle \sim M_X$, while $M \sim M_{st} \sim 10 M_X$, the terms $a_{23}\langle 45_H \rangle/M$ and $g_{23}\langle 16_H \rangle/M$ can quite plausibly be of order $h_{33}/10$, if $a_{23} \sim g_{23} \sim h_{33}$. By the assumption of hierarchy, we presume that $a_{12} \ll a_{23}$, and $g_{12} \ll g_{23}$.

It is interesting to observe the symmetry properties of the $a_{23}$ and $g_{23}$-terms. Although $10_H \times 45_H = 10 + 120 + 320$, given that $\langle 45_H \rangle$ is along $B-L$, which is needed to implement doublet-triplet splitting (see Appendix), only 120 in the decomposition contributes to the mass-matrices. This contribution is, however, antisymmetric in the family-index and, at the same time, proportional to $B-L$. Thus the $a_{23}$ term fulfills the requirements of both antisymmetry and $(B-L)$-dependence, simultaneously\(^4\). With only $h_{ij}$ and $a_{ij}$-terms, however, the up and down quark mass-matrices will be proportional to each other, which would yield $V_{CKM} = 1$. This is remedied by the $g_{ij}$ coupling. Because, the $16_H$ can have a VEV not only along its SM singlet component (transforming as $\bar{\mathbf{7}}_R$) which is of GUT-scale, but also along its electroweak doublet component — call it $16_d$ — of the electroweak scale. The latter can arise by the the mixing of $16_d$ with the corresponding doublet (call it $10_d$) in the $10_H$. The MSSM doublet $H_d$, which is light, is then a mixture of $10_d$ and $16_d$, while the orthogonal combination is superheavy (see Appendix). Since $\langle 16_d \rangle$ contributes only to the down-flavor mass matrices, but not to the up-flavor, the $g_{23}$ and $g_{12}$ couplings generate non-trivial CKM-mixings. We thus see that the minimal Higgs system satisfies apriori all the qualitative requirements (2)-(4), including the condition of $V_{CKM} \neq 1$. I now discuss that this system works well even quantitatively.

With these six effective Yukawa couplings, the Dirac mass matrices of quarks and leptons of the three families at the unification scale take the form:

$$U = \begin{pmatrix} 0 & \epsilon' & 0 \\ -\epsilon' & 0 & \epsilon + \sigma \\ 0 & -\epsilon + \sigma & 1 \end{pmatrix} m_U, \quad D = \begin{pmatrix} 0 & \epsilon' + \eta' & 0 \\ -\epsilon' + \eta' & 0 & \epsilon + \eta \\ 0 & -\epsilon + \eta & 1 \end{pmatrix} m_D,$$

$$N = \begin{pmatrix} 0 & -3\epsilon' & 0 \\ 3\epsilon' & 0 & -3\epsilon + \sigma \\ 0 & 3\epsilon + \sigma & 1 \end{pmatrix} m_U, \quad L = \begin{pmatrix} 0 & -3\epsilon' + \eta' & 0 \\ 3\epsilon' + \eta' & 0 & -3\epsilon + \eta \\ 0 & 3\epsilon + \eta & 1 \end{pmatrix} m_D. \quad (16)$$

Here the matrices are multiplied by left-handed fermion fields from the left and by anti-fermion fields from the right. $(U, D)$ stand for the mass matrices of up and down quarks, while $(N, L)$ are the Dirac mass matrices of the neutrinos and the charged leptons. The

\(^7\)Although no explicit string solution with the hierarchy in $h_{ij}$ mentioned above, together with the $a_{ij}$ and $g_{ij}$ couplings of Eq. (15), exists as yet, flavor symmetries of the type alluded to, as well as SM singlets carrying flavor-charges and acquiring VEVs of order $M_X$ that can lead to effective hierarchical couplings, do emerge in string solutions. And, there exist solutions with top Yukawa coupling being leading (see e.g. Refs. [34] and [33]).

\(^4\)The analog of $10_H \cdot 45_H$ for the case of $G(224)$ would be $\chi_H \equiv (2, 2, 1)_H \cdot (1, 1, 15)_H$. Although in general, the coupling of $\chi_H$ to the fermions need not be antisymmetric, for a string-derived $G(224)$, the multiplet $(1, 1, 15)_H$ is most likely to arise from an underlying 45 of SO(10) (rather than 210); in this case, the couplings of $\chi_H$ must be antisymmetric like that of $10_H \cdot 45_H$. 

13
entries 1, $\epsilon$, and $\sigma$ arise respectively from the $h_{33}, a_{23}$ and $h_{23}$ terms in Eq. (15), while $\eta$ entering into $D$ and $L$ receives contributions from both $g_{23}$ and $h_{23}$; thus $\eta \neq \sigma$. Similarly $\eta'$ and $\epsilon'$ arise from $g_{12}$ and $a_{12}$ terms respectively. Note the quark-lepton correlations between $U$ and $N$ as well as $D$ and $L$, and the up-down correlations between $U$ and $D$ as well as $N$ and $L$. These correlations arise because of the symmetry property of $G(224)$. The relative factor of $-3$ between quarks and leptons involving the $\epsilon$ entry reflects the fact that $\langle 45_H \rangle \propto (B - L)$, while the antisymmetry in this entry arises from the group structure of $SO(10)$, as explained above$^4$. As we will see, this $\epsilon$-entry helps to account for (a) the differences between $m_s$ and $m_\mu$, (b) that between $m_d$ and $m_e$, and also, (c) the suppression of $V_{cb}$ together with the enhancement of the $\nu_\mu$-$\nu_\tau$ oscillation angle.

The mass matrices in Eq. (16) contain 7 parameters$^5$: $\epsilon, \sigma, \eta, m_D = h_{33} \langle 10_d \rangle, m_U = h_{33} \langle 10_U \rangle, \eta'$ and $\epsilon'$. These may be determined by using, for example, the following input values: $m_t^{\text{phys}} = 174$ GeV, $m_c(m_c) = 1.37$ GeV, $m_s(1$ GeV) = 110-116 MeV [43], $m_u(1$ GeV) $\approx 6$ MeV and the observed masses of $e, \mu$ and $\tau$, which lead to (see Ref. [6], for details):

$$\sigma \simeq 0.110, \quad \eta \simeq 0.151, \quad \epsilon \simeq -0.095, \quad |\eta'| \approx 4.4 \times 10^{-3} \quad \text{and} \quad \epsilon' \approx 2 \times 10^{-4}$$

$$m_U \simeq m(t)(M_U) \simeq (100-120) \text{GeV}, \quad m_D \simeq m(b)(M_U) \simeq 1.5 \text{GeV}. \quad (17)$$

We have assumed for simplicity that the parameters are real, because a good fitting suggests that the relative phases of at least $\sigma$, $\eta$ and $\epsilon$ are small ($< 10^6$ say). Such fitting also fixes their relative signs. Note that in accord with our general expectations discussed above, each of the parameters $\sigma$, $\eta$ and $\epsilon$ are found to be of order 1/10, as opposed to being $^6 \mathcal{O}(1)$ or $O(10^{-2})$, compared to the leading (3,3)-element in Eq. (16). Having determined these parameters, we are led to a total of five predictions involving only the quarks (those for the leptons are listed separately):

$$m_b^0 \approx m_b^0(1 - 8\epsilon^2); \quad \text{thus} \quad m_b(m_b) \simeq (4.6-4.9) \text{GeV} \quad (18)$$

$$|V_{cb}| \simeq |\sigma - \eta| \approx \left| \sqrt{m_s/m_b} \left| \frac{\eta + \epsilon}{\eta - \epsilon} \right|^{1/2} - \sqrt{m_c/m_t} \left| \frac{\sigma + \epsilon}{\sigma - \epsilon} \right|^{1/2} \right| \approx 0.045 \quad (19)$$

$$m_d(1\text{GeV}) \simeq 8 \text{MeV} \quad (20)$$

$$\theta_C \simeq \left| \sqrt{m_d/m_s} - e^{i\phi} \sqrt{m_u/m_c} \right| \quad (21)$$

$$|V_{ub}/V_{cb}| \simeq \sqrt{m_u/m_c} \simeq 0.07. \quad (22)$$

In making these predictions, we have extrapolated the GUT-scale values down to low energies using $\alpha_3(m_Z) = 0.118$, a SUSY threshold of 500 GeV and $\tan \beta = 5$. The results depend

---

$^5$Of these, $m_U^0 \approx m_l^0$ can in fact be estimated to within 20% accuracy by either using the argument of radiative electroweak symmetry breaking, or some promising string solutions (see e.g. Ref. [34]).

$^6$This is one characteristic difference between our work and that of Ref. [38], where the (2,3)-element is even bigger than the (3,3).
weakly on these choices, assuming $\tan \beta \approx 2-30$. Further, the Dirac masses and mixings of the neutrinos and the mixings of the charged leptons also get determined. We obtain:

$$m_{\nu_{\tau}}^D(M_U) \approx 100-120 \text{ GeV}; \quad m_{\nu_{\mu}}^D(M_U) \approx 8 \text{ GeV},$$  \hspace{1cm} \text{(23)}

$$\theta_{\mu \tau}^e \approx -3 \epsilon + \eta \approx \sqrt{m_{\mu}/m_{\tau}} \left| \frac{-3 \epsilon + \eta}{3 \epsilon + \eta} \right|^{1/2} \approx 0.437$$  \hspace{1cm} \text{(24)}

$$m_{\nu_e}^D \approx \left[ 9 \epsilon'^2/(9\epsilon^2 - \sigma^2) \right] m_U \approx 0.4 \text{ MeV}$$  \hspace{1cm} \text{(25)}

$$\theta_{\mu \tau}^e \approx \left| \frac{\eta' - 3 \epsilon'}{\eta' + 3 \epsilon'} \right|^{1/2} \sqrt{m_e/m_{\mu}} \approx 0.85 \sqrt{m_e/m_{\mu}} \approx 0.06$$  \hspace{1cm} \text{(26)}

$$\theta_{e \tau}^e \approx \frac{1}{0.85} \sqrt{m_e/m_{\tau}} (m_{\mu}/m_{\tau}) \approx 0.0012.$$  \hspace{1cm} \text{(27)}

In evaluating $\theta_{e \mu}^e$, we have assumed $\epsilon'$ and $\eta'$ to be relatively positive.

Given the bizarre pattern of quark and lepton masses and mixings, it seems remarkable that the simple pattern of fermion mass-matrices, motivated by the group theory of G(224)/SO(10), gives an overall fit to all of them which is good to within 10%. This includes the two successful predictions on $m_b$ and $V_{cb}$ (Eqs.(18 and (19)). Note that in supersymmetric unified theories, the “observed” value of $m_b(m_b)$ and renormalization-group studies suggest that, for a wide range of the parameter $\tan \beta$, $m_b^0$ should in fact be about 10-20% lower than $m_{\tau}^0$ [44]. This is neatly explained by the relation: $m_b^0 \approx m_{\tau}^0 (1 - 8\epsilon^2)$ (Eq. (18)), where exact equality holds in the limit $\epsilon \rightarrow 0$ (due to SU(4)-color), while the decrease of $m_b^0$ compared to $m_{\tau}^0$ by $8\epsilon^2 \approx 10\%$ is precisely because the off-diagonal $\epsilon$-entry is proportional to $B$-$L$ (see Eq. (16)).

Specially intriguing is the result on $V_{cb} \approx 0.045$ which compares well with the observed value of $\approx 0.04$. The suppression of $V_{cb}$, compared to the value of $0.17 \pm 0.06$ obtained from Eq. (13), is now possible because the mass matrices (Eq. (16)) contain an antisymmetric component $\propto \epsilon$. That corrects the square-root formula $\theta_{cb} = \sqrt{m_s/m_b}$ (appropriate for symmetric matrices, see Eq. (11)) by the asymmetry factor $[(\eta + \epsilon)/(\eta - \epsilon)]^{1/2}$ (see Eq. (19)), and similarly for the angle $\theta_{ct}$. This factor suppresses $V_{cb}$ if $\eta$ and $\epsilon$ have opposite signs. The interesting point is that, the same feature necessarily enhances the corresponding mixing angle $\theta_{\mu \tau}^e$ in the leptonic sector, since the asymmetry factor in this case is given by $[(3 \epsilon + \eta)/(3 \epsilon + \eta)]^{1/2}$ (see Eq. (24)). This enhancement of $\theta_{\mu \tau}^e$ helps to account for the nearly maximal oscillation angle observed at SuperK (as discussed below). This intriguing correlation between the mixing angles in the quark versus leptonic sectors – that is suppression of one implying enhancement of the other – has become possible only because of the $\epsilon$-contribution, which is simultaneously antisymmetric and is proportional to $B$-$L$. That in turn becomes possible because of the group-property of SO(10) or a string-derived G(224).4

Taking stock, we see an overwhelming set of evidences in favor of $B$-$L$ and in fact for the full SU(4)-color-symmetry. These include: (i) the suppression of $V_{cb}$, together with the enhancement of $\theta_{\mu \tau}^e$, just mentioned above, (ii) the successful relation $m_{\nu}^0 \approx m_{\tau}^0 (1 - 8\epsilon^2)$, (iii) the usefulness again of the SU(4)-color-relation $m_{\nu_{\text{Dirac}}}^0 \approx m_{\nu_{\text{Dirac}}}^0$ in accounting for $m_{\nu_{\text{Dirac}}}^0$ (see Sec. 4), and (iv) the agreement of the relation $|m_{\nu_{s}}^0/m_{\nu_{\mu}}^0| \approx |(\epsilon^2 - \eta^2)/(9\epsilon^2 - \eta^2)|$ with
the data, in that the ratio is naturally less than 1, if $\eta \sim \epsilon$. The presence of $9\epsilon^2$ in the denominator is because the off-diagonal entry is proportional to B-L. Finally, the need for (B-L)- as a local symmetry, to implement baryogenesis, has been noted in Sec.1.

Turning to neutrino masses, while all the entries in the Dirac mass matrix $N$ are now fixed, to obtain the parameters for the light neutrinos, one needs to specify those of the Majorana mass matrix of the RH neutrinos ($\nu^c_{R,\mu,\tau}$). Guided by economy and the assumption of hierarchy, we consider the following pattern:

$$M^R_\nu = \begin{pmatrix} x & 0 & z \\ 0 & 0 & y \\ z & 0 & 1 \end{pmatrix} M_R. \quad (28)$$

As discussed in Sec. 4, the magnitude of $M_R \approx (5-15) \times 10^{14}$ GeV can quite plausibly be justified in the context of supersymmetric unification\(^7\) (e.g. by using $M \approx M_{st} \approx 4 \times 10^{17}$ GeV in Eq. (8)). To the same extent, the magnitude of $m(\nu_\tau) \approx (1/10-1/30)$ eV, which is consistent with the SuperK value, can also be anticipated. Thus there are effectively three new parameters: $x$, $y$, and $z$. Since there are six observables for the three light neutrinos, one can expect three predictions. These may be taken to be $\theta^{\text{osc}}_{\nu_\mu,\nu_\tau}$, $m_{\nu_\tau}$ (see Eq. (10)), and for example $\theta^{\text{osc}}_{\nu_\mu,\nu_\mu}$.

Assuming successively hierarchical entries as for the Dirac mass matrices, we presume that $|y| \sim 1/10$, $|z| \leq |y|/10$ and $|x| \leq z^2$. Now given that $m(\nu_\tau) \sim 1/20$ eV (as estimated in Eq. (10)), the MSW solution for the solar neutrino puzzle [45] suggests that $m(\nu_\mu)/m(\nu_\tau) \approx 1/10-1/30$. The latter in turn yields: $|y| \approx (1/18$ to $1/23.6)$, with $y$ having the same sign as $\epsilon$ (see Eq. (17)). This solution for $y$ obtains only by assuming that $y$ is $O(1/10)$ rather than $O(1)$. Combining now with the mixing in the $\mu-\tau$ sector determined above (see Eq. (24)), one can then determine the $\nu_\mu-\nu_\tau$ oscillation angle. The two predictions of the model for the neutrino-system are then:

$$m(\nu_\tau) \approx (1/10-1/30) \text{ eV} \quad (29)$$

$$\theta^{\text{osc}}_{\nu_\mu,\nu_\tau} \approx \theta^{\ell}_{\mu\tau} - \theta^{\nu}_{\mu\tau} \approx \left(0.437 + \sqrt{\frac{m_{\nu_2}}{m_{\nu_1}}}\right). \quad (30)$$

Thus, $\sin^2 2\theta^{\text{osc}}_{\nu_\mu,\nu_\tau} = (0.96, 0.91, 0.86, 0.83, 0.81)$

for $m_{\nu_2}/m_{\nu_3} = (1/10, 1/15, 1/20, 1/25, 1/30). \quad (32)$

Both of these predictions are extremely successful.

Note the interesting point that the MSW solution, together with the requirement that $|y|$ should have a natural hierarchical value (as mentioned above), lead to $y$ having the same sign as $\epsilon$; that (it turns out) implies that the two contributions in Eq.(30) must add rather than subtract, leading to an almost maximal oscillation angle\([6]\). The other factor contributing to the enhancement of $\theta^{\text{osc}}_{\nu_\mu,\nu_\tau}$ is, of course, also the asymmetry-ratio which increases $|\theta^{\ell}_{\mu\tau}|$ from 0.25 to 0.437 (see Eq. (24)). We see that one can derive rather plausibly a large $\nu_\mu-\nu_\tau$ oscillation angle $\sin^2 2\theta^{\text{osc}}_{\nu_\mu,\nu_\tau} \geq 0.8$, together with an understanding of hierarchical masses and

\(^7\)This estimate for $M_R$ is retained even if one allows for $\nu_\mu-\nu_\tau$ mixing (see Ref. [6]).
mixings of the quarks and the charged leptons, while maintaining a large hierarchy in the seesaw derived masses \((m_{l_2}/m_{l_3} = 1/10-1/30)\), all within a unified framework including both quarks and leptons. In the example exhibited here, the mixing angles for the mass eigenstates of neither the neutrinos nor the charged leptons are really large, in that \(\theta_{\mu \tau}^0 \simeq 0.437 \sim 23^\circ\) and \(\theta_{\mu \tau}^0 \simeq (0.18-0.31) \approx (10-18)^\circ\), yet the oscillation angle obtained by combining the two is near-maximal. This contrasts with most previous work, in which a large oscillation angle is obtained either entirely from the neutrino sector (with nearly degenerate neutrinos) or almost entirely from the charged lepton sector.

While \(M_R \approx (5-15) \times 10^{14}\) GeV and \(y \approx -1/20\) are better determined, the parameters \(x\) and \(z\) cannot be obtained reliably at present because very little is known about observables involving \(\nu_e\). Taking, for concreteness, \(m_{\nu_e} \approx (10^{-5}-10^{-4}) (1\text{ to few})\) eV and \(\theta_{\mu \tau}^{\text{osc}} \approx \theta_{\mu \tau}^0 \approx 10^{-3} \pm 0.03\) as inputs, we obtain: 
\[
z \sim (1-5) \times 10^{-3}\text{ and } x \sim (1\text{ to few})(10^{-6}-10^{-5}),
\]
with the guidelines of \(|z| \sim |y|/10\) and \(|x| \sim z^2\). This in turn yields: 
\[
\theta_{\mu \tau}^{\text{osc}} \approx \theta_{e\mu}^0 - \theta_{e\mu}^0 \approx 0.06 \pm 0.015.
\]
Note that the mass of \(m_{\nu_e} \approx 3 \times 10^{-3}\) eV, that follows from a natural hierarchical value for \(y \sim -(1/20)\), and \(\theta_{e\mu}\) as above, go well with the small angle MSW explanation\(^8\) of the solar neutrinos puzzle.

It is worth noting that although the superheavy Majorana masses of the RH neutrinos cannot be observed directly, they can be of cosmological significance. The pattern given above and the arguments given in Sec. 3 and in this section suggests that \(M(\nu^R) \approx (5-15) \times 10^{14}\) GeV, \(M(\nu^L) \approx (1-4) \times 10^{12}\) GeV (for \(x \approx 1/20\)); and \(M(\nu^R) \approx (1/2-10) \times 10^9\) GeV (for \(x \approx (1/2-10)10^{-6} > z^2\)). A mass of \(\nu^c \sim 10^9\) GeV is of the right magnitude for producing \(\nu^e R\) following reheating and inducing lepton asymmetry in \(\nu_e R\) decay into \(H^0 + \nu^L\), that is subsequently converted into baryon asymmetry by the electroweak sphalerons [19, 20].

In summary, we have proposed an economical and predictive pattern for the Dirac mass matrices, within the SO(10)/G(224)-framework, which is remarkably successful in describing the observed masses and mixings of all the quarks and charged leptons. It leads to five predictions for just the quark- system, all of which agree with observation to within 10%. The same pattern, supplemented with a similar structure for the Majorana mass matrix, accounts for both the large \(\nu_e-\nu_\tau\) oscillation angle and a mass of \(\nu_\tau \sim 1/20\) eV, suggested by the SuperK data. It also accommodates a small \(\nu_e-\nu_\mu\) oscillation angle relevant for theories of the solar neutrino deficit. Given this degree of success, it makes good sense to study proton decay concretely within this SO(10)/G(224)-framework. The results of this study [6] are presented in the next section.

Before turning to proton decay, it is worth noting that much of our discussion of fermion masses and mixings, including those of the neutrinos, is essentially unaltered if we go to the limit \(\epsilon' \to 0\) of Eq. (28). This limit clearly involves:
\[
m_u = 0, \quad \theta_C \simeq \sqrt{m_d/m_s}, \quad m_{\nu_e} = 0, \quad \theta_{e\mu}^0 = \theta_{e\tau}^0 = 0.
\]
\[
|V_{ub}| \simeq \sqrt{\eta - \epsilon \over \eta + \epsilon} \sqrt{m_d/m_b} (m_s/m_b) \simeq (2.1)(0.039)(0.023) \approx 0.0019
\]
\(^8\)Although the small angle MSW solution appears to be more generic within the approach outlined above, we have found that the large angle solution can still plausibly emerge in a limited region of parameter space, without affecting our results on fermion masses.
All other predictions remain unaltered. Now, among the observed quantities in the list above, $\theta_C \simeq \sqrt{m_d/m_s}$ is a good result. Considering that $m_u/m_t \approx 10^{-5}$, $m_u = 0$ is also a pretty good result. There are of course plausible small corrections which could arise through Planck scale physics; these could induce a small value for $m_u$ through the (1,1)-entry $\delta \approx 10^{-5}$. For considerations of proton decay, it is worth distinguishing between these two variants, which we will refer to as cases I and II respectively.

$$\begin{align*}
\text{Case I} & : \quad \epsilon' \approx 2 \times 10^{-4}, \quad \delta = 0 \\
\text{Case II} & : \quad \delta \approx 10^{-5}, \quad \epsilon' = 0.
\end{align*}$$

(34)

6 Expectations for Proton Decay in Supersymmetric Unified Theories

6.1 Turning to the main purpose of this talk, I present now the reason why the unification framework based on SUSY SO(10) or G(224), together with the understanding of fermion masses and mixings discussed above, strongly suggest that proton decay should be imminent.

Recall that supersymmetric unified theories (GUTs) introduce two new features to proton decay: (i) First, by raising $M_X$ to a higher value of about $2 \times 10^{16}$ GeV, they strongly suppress the gauge-boson-mediated $d = 6$ proton decay operators, for which $e^+\pi^0$ would have been the dominant mode (for this case, one typically obtains: $\Gamma^{-1}(p \to e^+\pi^0)|_{d=6} \approx 10^{36\pm1.5}$ yrs). (ii) Second, they generate $d = 5$ proton decay operators [22] of the form $Q_iQ_jQ_kQ_l/M$ in the superpotential, through the exchange of color triplet Higginos, which are the GUT partners of the standard Higgs(ino) doublets, such as those in the $5 + \overline{5}$ of SU(5) or the 10 of SO(10). Assuming that a suitable doublet-triplet splitting mechanism provides heavy GUT-scale masses to these color triplets and light masses to the doublets, these “standard” $d = 5$ operators, suppressed by just one power of the heavy mass and the small Yukawa couplings, are found to provide the dominant mechanism for proton decay in supersymmetric GUT [46, 47, 48, 49].

Now, owing to (a) Bose symmetry of the superfields in $QQQL/M$, (b) color antisymmetry, and especially (c) the hierarchical Yukawa couplings of the Higgs doublets, it turns out that these standard $d = 5$ operators lead to dominant $\pi K^+$ and comparable $\pi \pi^+$ modes, but in all cases to highly suppressed $e^+\pi^0$, $e^+K^0$ and even $\mu^+K^0$ modes. For instance, for minimal SUSY SU(5), one obtains (with $\tan \beta \leq 20$, say):

$$[\Gamma(\mu^+K^0)/\Gamma(\pi K^+)]_{SU(5)}^{\text{std}} \approx [m_u/m_c \sin^2 \theta] R \approx 10^{-3},$$

(35)

where $R \approx 0.1$ is the ratio of the relevant $|\text{matrix element}|^2 \times \text{(phase space)}$, for the two modes.

It was recently pointed out that in SUSY unified theories based on SO(10) or G(224), which assign heavy Majorana masses to the RH neutrinos, there exists a new set of color triplets and thereby very likely a new source of $d = 5$ proton decay operators [21]. For
instance, in the context of the minimal set of Higgs multiplets\(^9\) \(\{45_H, 16_H, \overline{10}_H \text{ and } 10_H\}\) (see Sec. 5), these new \(d = 5\) operators arise by combining three effective couplings introduced before: i.e., (a) the couplings \(f_{ij} 16_i 16_j \overline{10}_H \overline{10}_H / M\) (see Eq. (7)) that are required to assign Majorana masses to the RH neutrinos, (b) the couplings \(g_{ij} 16_i 16_j 16_H 16_H / M\), which are needed to generate non-trivial CKM mixings (see Eq. (15)), and (c) the mass term \(M_{16} 16_H \overline{10}_H\). For the \(f_{ij}\) couplings, there are two possible \(\text{SO}(10)\)-contractions, and we assume both to have comparable strength\(^10\). In this case, the color-triplet Higgsinos in \(\overline{10}_H\) and \(16_H\) of mass \(M_{16}\) can be exchanged between \(q_i q_j\) and \(q_k q_l\)-pairs. This exchange generates a new set of \(d = 5\) operators in the superpotential of the form

\[
W_{\text{new}} = f_{ij} g_{kl} (16_i 16_j) (16_k 16_l) (\overline{10}_H) (\overline{10}_H) / M ,
\]

which induce proton decay. Note that these operators depend, through the couplings \(f_{ij}\) and \(g_{kl}\), both on the Majorana and on the Dirac masses of the respective fermions. This is why within SUSY \(\text{SO}(10)\) or \(G(224)\), proton decay gets intimately linked to the masses and mixings of all fermions, including neutrinos.

### 6.2 Framework for Calculating Proton Decay Rate

To establish notations, consider the case of minimal SUSY SU(5) and, as an example, the process \(\bar{c} d \to s \nu_\mu\), which induces \(p \to \overline{\nu}_\mu K^+\). Let the strength of the corresponding \(d = 5\) operator, multiplied by the product of the CKM mixing elements entering into wino-exchange vertices, (which in this case is \(\sin \theta_C \cos \theta_C\)) be denoted by \(A\). Thus (putting \(\cos \theta_C = 1\), one obtains:

\[
\hat{A}_\bar{c} d (SU(5)) = (h_{22}^u h_{12}^d / M_{H_C}) \sin \theta_c \simeq (m_c m_s \sin^2 \theta_C / v_\nu^2) (\tan \beta / M_{H_C})
\]

\[
\simeq (1.9 \times 10^{-8}) (\tan \beta / M_{H_C}) \approx (2 \times 10^{-24} \text{ GeV}^{-1}) (\tan \beta / 2) (2 \times 10^{16} \text{ GeV} / M_{H_C}) ,
\]

where \(\tan \beta \equiv v_u / v_d\), and we have put \(v_u = 174\) GeV and the fermion masses extrapolated to the unification-scale – i.e. \(m_c \simeq 300\) MeV and \(m_s \simeq 40\) MeV. The amplitude for the associated four-fermion process \(\bar{d} s \to \overline{\nu}_\mu\) is given by:

\[
A_5 (\bar{d} s \to \overline{\nu}_\mu) = \hat{A}_\bar{c} d \times (2f)
\]

where \(f\) is the loop-factor associated with wino-dressing. Assuming \(m_\tilde{\nu} \ll m_\tilde{\nu}^{} \sim m_\tilde{\nu}\) one gets: \(f \simeq (m_\tilde{\nu} / m_\tilde{\nu}) (\alpha_2 / 4\pi)\). Using the amplitude for \((d u)(s \nu_\ell)\), as in Eq. (38), \((\ell = \mu \text{ or } \tau)\), one then obtains \[47, 48, 49, 6\]:

\[
\Gamma^{-1}(p \to \overline{\nu}_\mu K^+) \approx (2.2 \times 10^{31}) \text{ yrs} \times
\]

\[
\left(\frac{0.67}{A_5}\right)^2 \left[\frac{0.006 \text{ GeV}^3}{\beta_H}\right]^2 \left[\frac{(1/6)}{m_\tilde{\nu} / m_\tilde{\nu}}\right]^2 \left[\frac{m_\tilde{\nu} / 1 \text{ TeV}}{A(\nu)}\right]^2 \left[\frac{2 \times 10^{-24} \text{ GeV}^{-1}}{A(\nu)}\right]^2 .
\]

\(^9\)The origin of the new \(d = 5\) operators in the context of other Higgs multiplets, in particular in the cases where \(126_H\) and \(\overline{10}_H\) are used to break \(B-L\), has been discussed in Ref. [21].

\(^{10}\)One would expect such a general contraction to hold, especially if the \(f_{ij}\) couplings are induced by non-perturbative quantum gravity. Furthermore, the \(f_{ij}\) couplings with the contraction of the pair \((16_i \cdot \overline{10}_H)\), being effectively in \(45\) (rather than in \(1\) of \(\text{SO}(10)\), would be induced also by tree-level exchanges, if these pairs couple to the \(45\)’s in the string tower. Such a contraction would lead to proton decay.
Here $\beta_H$ denotes the hadronic matrix element defined by $\beta_H u_L(\vec{k}) \equiv \epsilon_{\alpha\beta\gamma} \langle 0 | (d_L^\alpha u_L^\beta) u_L^\gamma | p, \vec{k} \rangle$. While the range $\beta_H = (0.003-0.03) \text{ GeV}^3$ has been used in the past [48], given that one lattice calculation yields $\beta_H = (5.6 \pm 0.5) \times 10^{-3} \text{ GeV}^3$ [50], we will take as a plausible range: $\beta_H = (0.006 \text{ GeV}^3)(1/2 - 2)$. Here, $A_S \approx 0.67$ stands for the short distance renormalization factor of the $d = 5$ operator. Note that the familiar factors that appear in the expression for proton lifetime – i.e., $M_{HC}, (1 + \gamma_t)$ representing the interference between the $\tilde{t}$ and $\tilde{c}$ contributions, and $\tan \beta$ (see e.g. Ref.[48]) – are all effectively contained in $\hat{A}(\tau)$. Allowing for plausible and rather generous uncertainties in the matrix element and the spectrum we take:

$$\beta_H = (0.006 \text{ GeV}^3)(1/2 - 2)$$

$$m_{\tilde{d}}/m_{\tilde{q}} = 1/6 (1/2 - 2), \ \text{and} \ m_{\tilde{q}} \approx m_{\tilde{t}} \approx 1 \text{ TeV} (1/\sqrt{2} - \sqrt{2}). \quad (40)$$

Using Eqs. (39-40), we get:

$$\Gamma^{-1}(p \to \tau\bar{\tau} K^+) \approx (2.2 \times 10^{31} \text{ yrs}) [2.2 \times 10^{-24} \text{ GeV}^{-1}/\hat{A}(\tau)]^2 [32 - 1/32]. \quad (41)$$

This relation, as well as Eq. (39) are general, depending only on $\hat{A}(\tau)$ and on the range of parameters given in Eq. (40). They can thus be used for both SU(5) and SO(10).

The experimental lower limit on the inverse rate for the $\tilde{u} K^+$ modes is given by [9],

$$[\sum_{\ell} \Gamma(p \to \tau_{\ell} K^+)]^{-1}_{\text{expt}} > 1.6 \times 10^{33} \text{ yrs}. \quad (42)$$

Allowing for all the uncertainties to stretch in the same direction (in this case, the square bracket = 32), and assuming that just one neutrino flavor (e.g. $\nu_\mu$ for SU(5)) dominates, the observed limit Eq. (42) provides an upper bound on the amplitude\footnote{If there are sub-dominant $\tau_{\ell} K^+$ modes with branching ratio $R$, the right side of Eq. (43) should be divided by $\sqrt{1+R}$.}:

$$\hat{A}(\tau) \leq \sqrt{2} \times 10^{-24} \text{ GeV}^{-1} \quad (43)$$

which holds for both SU(5) and SO(10). For minimal SU(5), using Eq. (37) and $\tan \beta \geq 2$ (which is suggested on several grounds), one obtains a lower limit on $M_{HC}$ given by:

$$M_{HC} \geq 3 \times 10^{16} \text{ GeV} \quad (SU(5)) \quad (44)$$

At the same time, higher values of $M_{HC} > 3 \times 10^{16} \text{ GeV}$ do not go very well with gauge coupling unification. Thus keeping $M_{HC} \leq 3 \times 10^{16}$ and $\tan \beta \geq 2$, we obtain from Eq. (37): $\hat{A}(SU(5)) \geq (4/3) \times 10^{-24} \text{ GeV}^{-1}$. Using Eq. (41), this in turn implies that

$$\Gamma^{-1}(p \to \tau K^+) \leq 1.5 \times 10^{33} \text{ yrs} \quad (SU(5)) \quad (45)$$

This is a conservative upper limit. In practise, it is unlikely that all the uncertainties, including that in $M_{HC}$, would stretch in the same direction to nearly extreme values so as
to prolong proton lifetime. A more reasonable upper limit, for minimal SU(5), thus seems to be: \( \Gamma^{-1}(p \to \tau K^+)(SU(5)) \leq (0.7) \times 10^{33} \) yrs. Given the experimental lower limit (Eq. (42)), we see that minimal SUSY SU(5) is already or almost on the verge of being excluded by proton decay-searches. We have of course noted in Sec. 4 that SUSY SU(5) does not go well with neutrino oscillations observed at SuperK.

Now, to discuss proton decay in the context of supersymmetric SO(10), it is necessary to discuss first the mechanism for doublet-triplet splitting. Details of this discussion may be found in Ref. [6]. A synopsis is presented in the appendix.

### 6.3. Proton Decay in Supersymmetric SO(10)

The calculation of the amplitudes \( \hat{A}_{\text{std}} \) and \( \hat{A}_{\text{new}} \) for the standard and the new operators for the SO(10) model, are given in detail in Ref. [6]. Here, I will present only the results. It is found that the four amplitudes \( \hat{A}_{\text{std}}(\tau K^+), \hat{A}_{\text{std}}(\mu K^+), \hat{A}_{\text{new}}(\tau K^+) \) and \( \hat{A}_{\text{new}}(\mu K^+) \) are in fact very comparable to each other, within about a factor of two, either way. Since there is no reason to expect a near cancellation between the standard and the new operators, especially for both \( \tau K^+ \) and \( \mu K^+ \) modes, we expect the net amplitude (standard + new) to be in the range exhibited by either one. Following Ref. [6], I therefore present the contributions from the standard and the new operators separately. Using the upper limit on \( M_{\text{eff}} \geq 3 \times 10^{18} \) GeV (see Appendix), we obtain a lower limit for the standard proton decay amplitude given by

\[
\hat{A}(\tau K^+)_{\text{std}} \geq \begin{cases} 
(7 \times 10^{-24} \text{GeV}^{-1}) (1/6 - 1/4) & \text{case I} \\
(3 \times 10^{-24} \text{GeV}^{-1}) (1/6 - 1/2) & \text{case II}
\end{cases}
\]  

(46)

Substituting into Eq. (41) and adding the contribution from the second competing mode \( \mu K^+ \), with a typical branching ratio \( R \approx 0.3 \), we obtain

\[
\Gamma^{-1}(\tau K^+)_{\text{std}} \leq \begin{cases} 
(3 \times 10^{31} \text{yrs.}) (1.6 - 0.7) & \text{case I} \\
(6.8 \times 10^{31} \text{yrs.}) (4 - 0.44) & \text{case II}
\end{cases}
\]  

(47)

The upper and lower entries in Eqs. (46) and (47) correspond to the cases I and II of the fermion mass-matrix - i.e. \( \epsilon' \neq 0 \) and \( \epsilon' = 0 \) - respectively, (see Eq. (34)). The uncertainty shown inside the square brackets correspond to that in the relative phases of the different contributions. The uncertainty of (32 to 1/32) arises from that in \( \beta_H, (m_{\tilde{W}}/m_{\tilde{q}}) \) and \( m_{\tilde{q}} \) (see Eq. (40)). Thus we find that for MSSM embedded in SO(10), the inverse partial proton decay rate should satisfy:

\[
\Gamma^{-1}(p \to \tau K^+)_{\text{std}} \leq \begin{cases} 
3 \times 10^{31\pm1.7} \text{yrs.} & \text{case I} \\
6.8 \times 10^{31\pm2.1} \text{yrs.} & \text{case II}
\end{cases}
\]  

(48)

The central value of the upper limit in Eq. (48) corresponds to taking the upper limit on \( M_{\text{eff}} \). The uncertainties of matrix element and spectrum are reflected in the exponents. The uncertainty in the most sensitive entry of the fermion mass matrix - i.e. \( \epsilon' \) - is fully incorporated (as regards obtaining an upper limit on the lifetime) by going from case I to case II. Note that this increases the lifetime by almost a factor of five. Any non-vanishing value of
\(\epsilon'\) would only shorten the lifetime compared to case II. In this sense, the larger of the two upper limits quoted above is rather conservative.

Evaluating similarly the contributions from only the new operators, we obtain:

\[
\Gamma^{-1}(\overline{\nu} K^+)_{\text{new}} \approx (3 \times 10^{31} \text{ yrs})\{16 - 1/1.7\} \{32 - 1/32\} .
\] (49)

Note that this contribution is independent of \(M_{\text{eff}}\). It turns out that it is also insensitive to \(\epsilon'\); thus it is nearly the same for cases I and II. Allowing for a net uncertainty at the upper end by as much as a factor of 20 to 200, arising jointly from the square and the curly brackets, i.e., without going to extreme ends of all parameters, the new operators related to neutrino masses, by themselves, lead to a proton decay lifetime bounded by:

\[
\Gamma^{-1}(\overline{\nu} K^+)_{\text{expected}} \leq (0.6 - 6) \times 10^{33} \text{ yrs. (SO(10) or string G(224))}
\] (50)

It should be stressed that while the standard \(d = 5\) operators would be absent for a string-derived G(224)-model, the new \(d = 5\) operators, related to the Majorana masses of the RH neutrinos and the CKM mixings, would still be present for such a model. Thus our expectations for the proton decay lifetime (as shown in Eq. (50)) and the prominence of the \(\mu^+ K^0\) mode (see below) hold for a string-derived G(224)-model, just as they do for SO(10).

6.4. The Charged Lepton Decay Mode \((p \rightarrow \mu^+ K^0)\)

I now note a distinguishing feature of the SO(10) or the G(224) model presented here. Allowing for uncertainties in the way the standard and the new operators can combine with each other for the three leading modes i.e. \(\overline{\nu}_\tau K^+, \overline{\nu}_\mu K^+\) and \(\mu^+ K^0\), we obtain (see Ref. [6] for details):

\[
B(\mu^+ K^0)_{\text{std+new}} \approx [1\% \text{ to } 50\%] \rho \text{ (SO(10) or string G(224))}
\] (51)

where \(\rho\) denotes the ratio of the squares of relevant matrix elements for the \(\mu^+ K^0\) and \(\overline{\nu} K^+\) modes. In the absence of a reliable lattice calculation for the \(\overline{\nu} K^+\) mode [50], one should remain open to the possibility of \(\rho \approx 1/2\) to 1 (say). We find that for a large range of parameters, the branching ratio \(B(\mu^+ K^0)\) can lie in the range of 20 to 40\% (if \(\rho \approx 1\)). This prominence of the \(\mu^+ K^0\) mode for the SO(10)/G(224) model is primarily due to contributions from the new operators. This contrasts sharply with the minimal SU(5) model, in which the \(\mu^+ K^0\) mode is expected to have a branching ratio of only about \(10^{-3}\). In short, prominence of the \(\mu^+ K^0\) mode, if seen, would clearly show the relevance of the new operators, and thereby reveal the proposed link between neutrino masses and proton decay [21].

6.5. Section Summary

In summary, our study of proton decay has been carried out within the SO(10) or the G(224)-framework\(^{12}\), with special attention paid to its dependence on fermion masses and threshold effects. The study strongly suggests an upperlimit on proton lifetime, given by:

\[
\tau_{\text{proton}} \leq (1/2 - 1) \times 10^{34} \text{ yrs ,}
\] (52)

\(^{12}\)As described in Secs. 3 and 5.
with $\pi K^+$ being the dominant decay mode. Although there are uncertainties in the matrix element, in the SUSY-spectrum, and in certain sensitive elements of the fermion mass matrix, especially $\epsilon'$ (see Eq. (48) for predictions in cases I versus II), this upper limit is obtained by allowing for a generous range in these parameters and stretching all of them in the same direction so as to extend proton lifetime. In this sense, while the predicted lifetime spans a wide range, the upper limit quoted above is quite conservative. In turn, it provides a clear reason to expect that the discovery of proton decay should be imminent. The implication of this prediction for a next-generation detector is emphasized in the next section.

7 Concluding Remarks

The preceding sections show that one is now in possession of a set of facts, which may be viewed as the matching pieces of a puzzle; in that all of them can be resolved by just one idea - that is grand unification. These include: (i) the observed family-structure, (ii) meeting of the three gauge couplings, (iii) neutrino oscillations; in particular the mass of $\nu_e$ (suggested by SuperK), (iv) the intricate pattern of the masses and mixings of all the fermions, including the smallness of $V_{cb}$ and the largeness of $\rho_{\nu_e\nu_e}$, and (v) the need for $B-L$ to implement baryogenesis. All these pieces fit beautifully together within a single puzzle board framed by supersymmetric unification, based on SO(10) or a string-unified G(224)-symmetry.

The one and the most notable piece of the puzzle still missing, however, is proton decay. Based on a systematic study of this process within the supersymmetric SO(10)/G(224)-framework [6], which is clearly favored by the data, I have argued here that a conservative upper limit on the proton lifetime is about $(1/2 - 1) \times 10^{34}$ yrs. So, unless the fitting of all the pieces listed above is a mere coincidence, and I believe that that is highly unlikely, discovery of proton decay should be around the corner. In particular, as mentioned in the Introduction, we expect that candidate events should be observed in the near future already at SuperK. However, allowing for the possibility that proton lifetime may well be near the upper limit stated above, a next-generation detector providing a net gain in sensitivity by a factor five to ten, compared to SuperK, would be needed to produce real events and distinguish them unambiguously from the background. Such an improved detector would of course be essential to study the branching ratios of certain crucial though sub-dominant decay modes such as the $\mu^+K^0$.

The reason for pleading for such improved searches is that proton decay would provide us with a wealth of knowledge about physics at truly short distances ($< 10^{-30}$ cm), which cannot be gained by any other means. Specifically, the observation of proton decay, at a rate suggested above, with $\pi K^+$ mode being dominant, would not only reveal the underlying unity of quarks and leptons but also the relevance of supersymmetry. It would also confirm a unification of the fundamental forces at a scale of order $2 \times 10^{16}$ GeV. Furthermore, prominence of the $\mu^+K^0$ mode, if seen, would have even deeper significance, in that in addition to supporting the three features mentioned above, it would also reveal the link between neutrino masses and proton decay, as discussed in Sec. 6. In this sense, the role of proton decay in probing into physics at the most fundamental level is unique. In view of how valuable such a probe would be and the fact that the predicted upper limit on the proton lifetime is only a factor of three to six higher than the empirical lower limit, the argument
in favor of building an improved detector seems compelling.

To conclude, the discovery of proton decay would undoubtedly constitute a landmark in the history of physics. It would provide the last, missing piece of gauge unification and would shed light on how such a unification may be extended to include gravity.

Acknowledgements: I would like to thank Kaladi S. Babu and Frank Wilczek for a most enjoyable collaboration, and Joseph Sucher for valuable discussions. I would also like to thank the organizers of the NNNN99 workshop, especially Chang Kee Jung, Millind Diwan and Hank Sobel, for arranging a stimulating meeting and also for the kind hospitality. The research presented here is supported in part by DOE grant no. DE-FG02-96ER-41015.

Appendix
A Natural Doublet-Triplet Splitting Mechanism in SO(10)

In supersymmetric SO(10), a natural doublet–triplet splitting can be achieved by coupling the adjoint Higgs $45_H$ to a $10_H$ and a $10'_H$, with $45_H$ acquiring a unification–scale VEV in the $B-L$ direction [51]: $\langle 45_H \rangle = (a, a, a, 0, 0) \times \tau_2$ with $a \sim M_U$. As discussed in Section 2, to generate CKM mixing for fermions we require $\langle 16_H \rangle_d$ to acquire a VEV of the electroweak scale. To ensure accurate gauge coupling unification, the effective low energy theory should not contain split multiplets beyond those of MSSM. Thus the MSSM Higgs doublets must be linear combinations of the SU(2)$_L$ doublets in $10_H$ and $16_H$. A simple set of superpotential terms that ensures this and incorporates doublet-triplet splitting is [6]:

$$W_H = \lambda 10_H 45_H 10'H + M_{10} 10'_H^2 + \lambda' \overline{16}_H 1\overline{6}_H 10_H + M_{16} 16_H 1\overline{6}_H .$$ (A1)

A complete superpotential for $45_H$, $16_H$, $\overline{16}_H$, $10_H$, $10'_H$ and possibly other fields, which ensure that $45_H$, $16_H$ and $\overline{16}_H$ acquire unification scale VEVs with $\langle 45_H \rangle$ being along the $(B-L)$ direction, that exactly two Higgs doublets ($H_u$, $H_d$) remain light, with $H_d$ being a linear combination of $\langle 10_H \rangle_d$ and $\langle 16_H \rangle_d$, and that there are no unwanted pseudoGoldstone bosons, can be constructed. With $\langle 45_H \rangle$ in the $B-L$ direction, it does not contribute to the Higgs doublet mass matrix, so one pair of Higgs doublet remains light, while all triplets acquire unification scale masses. The light MSSM Higgs doublets are

$$H_u = 10_u, \quad H_d = \cos \gamma 10_d + \sin \gamma 16_d ,$$ (A2)

with $\tan \gamma \equiv \lambda' \langle \overline{16}_H \rangle / M_{16}$. Consequently, $\langle 10 \rangle_d = (\cos \gamma) v_d$, $\langle 16 \rangle_d = (\sin \gamma) v_d$, with $\langle H_d \rangle = v_d$ and $\langle 16 \rangle_d$ and $\langle 10 \rangle_d$ denoting the electroweak VEVs of those multiplets. Note that $H_u$ is purely in $10_H$ and that $\langle 10_d \rangle^2 + \langle 16_d \rangle^2 = v_d^2$. This mechanism of doublet-triplet (DT) splitting is rather unique for the minimal Higgs systems in that it meets the requirements of both D-T splitting and CKM-mixing. In turn, it has three special consequences:

(i) It modifies the familiar SO(10)-relation $\tan \beta \equiv v_u / v_d = m_t / m_b \approx 60$ to:

$$\tan \beta / \cos \gamma \approx m_t / m_b \approx 60$$ (A3)
As a result, even low to moderate values of $\tan \beta \approx 3$ to 10 (say) are perfectly allowed in SO(10) (corresponding to $\cos \gamma \approx 1/20$ to 1/6).

(ii) The most important consequence of the DT-splitting mechanism outlined above is this: In contrast to SU(5), for which the strengths of the standard d=5 operators are proportional to $(M_{H_c})^{-1}$ (where $M_{H_C} \sim few \times 10^{16}$ GeV (see Eq. (44)), for the SO(10)-model, they become proportional to $M_{\text{eff}}^{-1}$, where $M_{\text{eff}} = (\lambda \alpha)^2/M_{10} \sim M_{10}^2/M_{10}$. $M_{10}$ can be naturally smaller (due to flavor symmetries) than $M_{U}$ and thus $M_{\text{eff}}$ correspondingly larger than $M_{U}$ by one to two orders of magnitude (see Ref. [6]). Now the proton decay amplitudes for SO(10) in fact possess an intrinsic enhancement compared to those for SU(5), owing primarily due to differences in their Yukawa couplings for the up sector (see Appendix C in Ref. [6]). As a result, these larger values of $M_{\text{eff}} \sim 10^{18}$ GeV are in fact needed for the SO(10)-model to be compatible with the observed limit on the proton lifetime. At the same time, being bounded above (see below), they allow optimism as regards future observation of proton decay.

(iii) $M_{\text{eff}}$ gets bounded above by considerations of coupling unification and GUT-scale threshold effects. Owing to mixing between $10_d$ and $16_d$ (see Eq. (A2)), the threshold correction to $\alpha_3(m_z)$ due to doublet-triplet splitting becomes proportional to $\ln (M_{\text{eff}} \cos \gamma/M_{U})$. Inclusion of this correction and those due to splittings within the gauge and the Higgs multiplets (i.e. $45_H$, $16_H$, and $\bar{16}_H$)\(^{13}\), together with the observed degree of coupling unification allows us to obtain a conservative upper limit on $M_{\text{eff}}$, given by [6]:

\[ M_{\text{eff}} \leq 3 \times 10^{18} \text{ GeV}. \] (A4)

This in turn helps provide an upper limit on the expected proton decay lifetime (see text).

References


\(^{13}\)The correction to $\alpha_3(m_z)$ due to Planck scale physics through the effective operator $F_{\mu\nu}F^{\mu\nu}45_H/M$ vanishes due to antisymmetry in the SO(10)-contraction.


[25] For recent reviews see e.g. P. Langacker and N. Polonsky, Phys. Rev. D 47, 4028 (1993) and references therein.

[26] See e.g., Refs. [25] and [12].


[50] For a recent work, comparing the results of lattice and chiral lagrangian-calculations for the $p \rightarrow \pi^0, p \rightarrow \pi^+$ and $p \rightarrow K^0$ modes, see N. Tatsui et al (JLQCD collaboration), hep-lat/9809151.