GRAND UNIFIED THEORIES - II

J. Ellis
CERN, Geneva, Switzerland

INTRODUCTION

In the previous lectures[1,2] you have seen how the inadequacies of the standard $SU(3) \times SU(2) \times U(1)$ model of the strong, weak, and electromagnetic interactions motivate the construction of grand unified theories (GUTs) which are sketched in Fig. 1. The strategy then adopted was to assign the fundamental fermions -- quarks and leptons -- to "generations" forming economical representations of a simple grand unifying group G. The observed disparity of the strong and weak coupling constants implies that the group G must be broken down at a very high energy scale. The properties of the running coupling constants in renormalizable field theories enable one to estimate[3] this breaking scale to be around 10^{15} GeV in simple models[4] such as SU(5). Although the grand unification symmetry is very strongly broken, it can still be used to make predictions for parameters observable at low energies. In addition to the quantization of electromagnetic charge ($|Q_e|/|Q_p| = 1$) which flows from the embedding of electromagnetic U(1) into a simple group, it is possible to make predictions for the neutral weak mixing parameter $\sin^2 \theta_W$ and for the ratios of certain quark and lepton masses. While the predictions for $\sin^2 \theta_W$ and m_b/m_τ are spectacularly successful, there are clouds over the predictions for the masses of other quarks[1,2].

The real tests of GUTs lie, however, with their predictions for new interactions, which can lead to qualitatively new phenomena such as baryon decay or neutrino masses and oscillations. In these lectures we will see how some of these new interactions may have consequences in terrestrial experiments. They may also have important implications for cosmology[5], which forms the subject of the third of these lectures.

1. BARYON NUMBER VIOLATION

1.1 Qualitative introduction

We have already seen in previous lectures that in general GUTs one expects baryon number B to be violated owing to the introduction of direct quark-lepton and quark-antiquark transitions. This should not shock us: whilst formerly B conservation was sacred, it is no longer, thanks to the dogma of the gauge age which asserts that

"the only exact symmetries are gauge symmetries"

of which examples are electromagnetic U(1) and colour SU(3). Baryon and lepton numbers are not protected by such an exact gauge symmetry, and it is therefore to be expected that they will be violated at some level. In fact, two mechanisms for $\Delta B \neq 0$ processes are already known. One involves non-perturbative gravitational phenomena[6], where it is known from the "no-hair" theorem that the only conserved quantum numbers that can rigorously be associated with a black hole are those corresponding to long-range (gauge) fields: mass, angular momentum, and electromagnetic charge. We can therefore imagine a black hole catalysing a $\Delta B \neq 0$ process:

$$p + BH(m,J,Q) \to BH(m',J+\tfrac{1}{2},Q) \to BH(m,J,Q) + e^+ . \tag{1.1}$$

Obviously such a process would be unconscionably slow if we needed an astrophysical black hole. However, Zel'dovich[6] has pointed out that we could get an effective $\Delta B \neq 0$ interaction from a virtual black hole with a mass of the order of the Planck mass of 10^{19} GeV, which could cause baryons to decay with a lifetime

$$\tau_B \approx (10^{45} \text{ to } 10^{50}) \text{ years} . \tag{1.2}$$

Furthermore, 't Hooft[7] has pointed out that instantons of the standard SU(3) × SU(2) × U(1) theory can cause $\Delta B = N_G$ reactions which would allow, for example,

$$N + N \to \bar{N} + \text{leptons} , \tag{1.3}$$

but the rate for this is expected to be even slower than the black hole process (1.2). Fortunately baryons are expected to decay much faster in GUTs.

In minimal SU(5) [4] we have X and Y bosons interacting with the fundamental fermions in the following way:

$$\left.\begin{aligned} &\frac{1}{\sqrt{2}} g X_{i\mu} \left[-\bar{d}_{iR} \gamma^\mu e^+_R + \epsilon_{ijk} \bar{u}^c_{k_L} \gamma^\mu u_{jL} + \bar{d}_{iL} \gamma^\mu e^+_L \right] + \binom{\text{herm.}}{\text{conj.}} \\ &\frac{1}{\sqrt{2}} g Y_{i\mu} \left[\bar{d}_{iR} \gamma^\mu \nu^c_{e_R} + \epsilon_{ijk} \bar{u}^c_{k_L} \gamma^\mu d_{jL} - \bar{u}_{iL} \gamma^\mu e^+_L \right] + \binom{\text{herm.}}{\text{conj.}} \end{aligned}\right\} \tag{1.4}$$

if we neglect the grand unified analogue of generalized Cabibbo mixing, to which we return in Section 1.4. The exchanges of X and Y bosons give rise to an effective four-fermion interaction at low energies $\ll m_{X,Y}$:

$$\frac{1}{4}\mathcal{L}_{GU} = \frac{1}{\sqrt{2}} G_{GU} \begin{bmatrix} (\epsilon_{ijk} \bar{u}^c_{k_L} \gamma_\mu u_{jL})(2\bar{e}^+_L \gamma^\mu d_{iL} + \bar{e}^+_R \gamma^\mu d_{iR}) & \\ -(\epsilon_{ijk} \bar{u}^c_{k_L} \gamma_\mu d_{jL})(\bar{\nu}^c_e \gamma^\mu d_{iR}) & + (\text{herm. conj.}) \end{bmatrix} , \tag{1.5}$$

where the effective coupling G_{GU} is very analogous to the conventional Fermi coupling G_F mediated by $W^\pm$ exchange:

$$\frac{G_{GU}}{\sqrt{2}} = \frac{g^2}{8m_X^2} = \frac{g^2}{8m_Y^2} , \quad \text{cf} \quad \frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} . \tag{1.6}$$

We would get from the interaction (1.5) a decay amplitude

$$A \propto \frac{1}{m_X^2} \tag{1.7}$$

and thence a decay rate

$$\Gamma \propto |A|^2 \propto \frac{1}{m_X^4} \tag{1.8}$$

so that the baryon lifetime τ_B is $\propto m_X^4$. But τ_B has the mass-energy dimension -1, so dimensional analysis tells us that it must be scaled by some inverse mass to the fifth power, say the nucleon mass,

$$\tau_B = C\left(m_X^4/m_B^5\right) , \tag{1.9}$$

where C is a model-dependent coefficient to be calculated.

Looking at the minimal SU(5) results [(1.4) to (1.9)], several questions leap to mind. Do we expect the $\Delta B \neq 0$ interaction in general GUTs to resemble (1.5) in minimal SU(5)? In particular, should we always anticipate a four-fermion interaction? and is B-L always conserved as in (1.5)? These points will be addressed in Section 1.2. If we have a general form (1.5), so that the baryon lifetime is proportional to m_X^4 (1.9), we need to know the value of m_X as precisely as possible -- a point discussed in Section 1.3. In order to be sure that baryons are not trying to decay into $\bar{t}$ quarks or τ leptons, we need to study the generalized Cabibbo mixing structure that we neglected in writing down the effective interaction (1.4), which is the subject of Section 1.4. Next we need to know the value of the coefficient C in (1.9), and hence determine the total decay rate and branching ratios (Section 1.5). Finally, we should ask whether there are any other possible manifestations of $\Delta B \neq 0$ interactions, and neutron-antineutron oscillations are discussed in Section 1.6.

1.2 The form of the effective Lagrangian for baryon decay

Since the Lagrangian has mass dimension 4, any operator appearing as an effective Lagrangian which has dimension d > 4 must have a coefficient of order M^{4-d}. We will be generating such effective interactions from the exchanges of heavy particles, so we will evidently concentrate on the operators of lowest dimension d. Since SU(3) × SU(2) × U(1) is a good low-energy symmetry, we should classify[8,9] all possible interactions using invariance under this group. A simple tool for this task has been provided by Weinberg[8] in the guise of F parity. This quantity has the following assignments:

$$\underline{F=+1}: \quad q, \ell \qquad\qquad \underline{F=-1}: \quad \bar{q}, \bar{\ell}, \text{gauge bosons, Higgses, derivatives} \tag{1.10}$$

which are arrived at in the following way. Lorentz invariance can be implemented by classifying particles in representations (a,b) (where 2a and 2b are integers) and writing down interactions in which a and b are separately conserved. This implies in particular that $(-1)^{2a}$ and $(-1)^{2b}$ must be multiplicatively conserved. On the other hand, weak SU(2) invariance requires that $(-1)^{2I}$ be conserved. Weinberg[8] defines F parity as the particular combination

$$F \equiv (-1)^{2a+2I} \tag{1.11}$$

of these conserved quantities. Once F conservation is assumed, it is easy to implement invariance under the full SU(3) × SU(2) × U(1) group. Using the definition (1.11) we easily arrive at the assignments of (1.10):

$$\left.\begin{array}{ll} f_L: & a = I = \frac{1}{2}, b = 0 \\ f_R: & a = I = 0, b = \frac{1}{2} \end{array}\right\} F = +1 \qquad \left.\begin{array}{l} \bar{f}_L: a = 0, b = I = \frac{1}{2} \\ \bar{f}_R: a = \frac{1}{2}, b = I = 0 \end{array}\right\} F = -1 ,$$

$$\left.\begin{array}{ll} W, Z, & a = b = \frac{1}{2} \\ g, \gamma, \partial_\mu & I = 0 \text{ or } 1 \end{array}\right\} F = -1 \qquad H: \left.\begin{array}{l} a = b = 0 \\ I = \frac{1}{2} \end{array}\right\} F = -1 . \tag{1.12}$$

We are now in a position to look for the $\Delta B \neq 0$ operators of lowest possible dimension:

$$\underline{d=6} \qquad \left.\begin{array}{l} qqq\ell \\ qqq\bar{\ell} \end{array}\right\} \text{ have } \left\{\begin{array}{l} F = +1 : \text{allowed}, \\ F = -1 : \text{disallowed}. \end{array}\right. \tag{1.13}$$

We can construct[8,9] several different d = 6 operators of the allowed form, using the first generation of fermions,

$$(\epsilon_{ijk} \bar{u}^c_{k_L} \gamma_\mu u_{j_L})(\bar{e}^+_L \gamma^\mu d_{i_L}) \tag{1.14a}$$

$$(\epsilon_{ijk} \bar{u}^c_{k_L} \gamma_\mu)(u_{j_L} \bar{e}^+_R - d_{j_L} \bar{\nu}^c_{e_R}) \gamma^\mu d_{i_R} , \tag{1.14b}$$

and operators with similar particle content which cannot be generated by the exchanges of vector bosons. As examples, the minimal SU(5) effective Lagrangian (1.5) contains 2 × (1.14a) and 1 × (1.14b), whilst SO(10) with its extra gauge bosons can yield a more general combination of (1.14a) and (1.14b). There are some common properties of the allowed $qqq\ell$ operators (1.13) which are quite striking. The first is that B and L are violated in such a way that $\Delta B = \Delta L$, and hence

$$p, n \rightarrow (\ell^+, \bar{\nu}) + X \ , \not\rightarrow (\ell^-, \nu) + X \quad ; \tag{1.15}$$

B-L conservation is in fact an *exact* global symmetry of minimal SU(5), but only an *accidental*, approximate feature of more general models such as SO(10). Another striking feature of (1.13) is that

$$\Delta S / \Delta B \leq 0 \quad \Rightarrow \quad p, n \not\rightarrow K^- + X \ . \tag{1.16}$$

We now turn to the next lowest dimension operators:

$$\underline{d=7} \qquad \left.\begin{array}{l} qqq\bar{\ell}B \\ qqq\ell B \end{array}\right\} \text{ have } \left\{\begin{array}{l} F = +1 : \text{allowed} , \\ F = -1 : \text{disallowed} . \end{array}\right. \tag{1.17}$$

In equation (1.17), B stands for any gauge boson, Higgs boson, or space-time derivative. The allowed interaction $qqq\bar{\ell}B$ has $\Delta B = -\Delta L$, in contrast to (1.13). If our GUT only has one large mass-scale m_X, then we would expect the coefficient of (1.17) to be suppressed relative to that of (1.13) by a factor of $(m_W$ or $\Lambda_{QCD})/m_X$. Thus we expect B-L to be a good symmetry of baryon decays unless there is some intermediate mass-scale in between m_W and m_X. The prediction (1.15) is therefore a very general test of the GUT philosophy which planners of experiments should try to carry out.

It is often asked whether baryons can decay into three leptons or antileptons[10]. To do this requires an operator of dimension at least as large as

$$\underline{d=9} \qquad qqq\,\ell\ell\ell \quad \text{have} \quad F=+1 : \text{ allowed} , \tag{1.18}$$

which requires a very low mass-scale $M \sim 10^5$ GeV to scale its coefficient if it is to yield a baryon lifetime close to the present experimental limit. The decay of a baryon into three leptons requires

$$\underline{d=10} \qquad qqq\,\bar{\ell}\bar{\ell}\bar{\ell}B \quad \text{have} \quad F=+1 : \text{ allowed} , \tag{1.19}$$

and hence an even lower "large" mass-scale if it is to be competitive.

In the next few sections we will focus on the preferred form (1.13) of $\Delta B \neq 0$, B-L conserving interaction, and study it in detail. In Section 1.6 we will change gear to consider neutron-antineutron oscillations[11] which can only be mediated by higher-dimensional operators.

1.3 Determination of the grand unification mass

The basic idea of computing the unification mass m_X was given in the previous lectures[1,2]. In this section we mention a few of the subtleties that go into a more precise calculation, but refer you to the original literature[2] for more details. Roughly speaking, the SU(3), SU(2), and suitably normalized U(1) couplings are all equal at energies $Q \gg m_X$, and vary at lower energies according to the renormalization group so that, for example, the SU(3) and SU(2) couplings approach each other according to

$$\frac{1}{\alpha_3(Q)} - \frac{1}{\alpha_2(Q)} = \frac{11}{12\pi} \ln Q^2/m_X^2 . \tag{1.20}$$

It can then be deduced from the rate of approach that

$$\left(\frac{m_X}{\Lambda_{QCD}}\right) = \exp\left\{ \frac{O(1)}{\alpha} + \cdots \right\} . \tag{1.21}$$

Our job in this section is to determine precisely the O(1) terms in (1.21), and the leading dotted terms which are

$$\exp(\cdots) = \exp\left(O(1) \times \ln\alpha + O(1) + \cdots\right) , \tag{1.22}$$

as can be deduced from the solutions to the two-loop renormalization group equations for the coupling constants α_i.

If we assume the conventional generation structure for all particles with masses $\ll m_X$, there is nevertheless one small modification to be made to the leading approximation (1.20) to the rate of approach of α_1 and α_2. This is due to the inclusion of N_H light Higgs doublets:

$$\frac{1}{\alpha_3(Q)} - \frac{1}{\alpha_2(Q)} = \left(\frac{11 + N_H/2}{12\pi}\right) \ln Q^2/m_X^2 + \cdots . \tag{1.23}$$

Setting N_H equal to the minimal number of one reduces the grand unification mass by a factor of O(2) compared with what one would estimate on the basis of (1.20). This now correctly determines the O(1) coefficient in the exponent of equation (1.21).

Several effects contribute to the expression (1.22). One effect uses the two-loop approximation to the renormalization group equations between the weak interaction threshold $O(M_W)$ and the grand unification threshold $O(m_X)$: their inclusion has the effect of decreasing the estimate of m_X by O(4). Another effect is that of carefully taking into account the grand unification threshold[12]. The couplings defined in momentum space do not come together strictly at the point $Q = m_X$, but rather approach each other gradually as indicated in Fig. 2. Looked at from the low-energy side, the momentum scale at which the coupling constants seem to come together is actually rather larger than m_X, meaning that a naive estimate of m_X would be too high by a factor of 2 or 3. An analogous phenomenon[12] occurs at the weak interaction threshold. We might naively use the full SU(2) renormalization group equation for α_2 immediately we pass $Q = m_W$, in a sort of θ-function approximation. However, the true evolution of α_2 in momentum space looks rather complicated, as indicated in Fig. 3, and a better place to start the SU(2) evolution is somewhat above $Q = m_W$. This means that α_2 at $Q > m_W$ is somewhat larger than had been naively thought, and hence that the naive estimate of m_X is further reduced by a factor of the order of 2 or 3. However, the most significant reduction in m_X comes from a very banal source[13,14], namely the variation in the fine structure constant between the Thompson limit ($Q = 0$) where it is measured to be $1/137$, and the region of $Q \approx m_W$ where it is embedded into the SU(2) × U(1) weak interaction theory. The dominant renormalization of α in this range comes from the vacuum polarization graphs of Fig. 4, which change α^{-1} by about 9:

$$\left.\frac{1}{\alpha}\right|_{Q \approx m_W} = \left.\frac{1}{\alpha}\right|_{Q=0} - \left(9 \pm \tfrac{1}{2}\right) = 128 \pm \tfrac{1}{2} \quad , \tag{1.24}$$

where the main uncertainty comes from not knowing how to treat exactly the strong corrections to the vacuum polarization (Fig. 4) for the light quarks. This change again pushes up the weak coupling constants at large Q, and hence decreases the estimate of m_X by almost an order of magnitude[13,14]. As a result of all these corrections, the current best estimate of m_X is 2 orders of magnitude smaller than early estimates. The principal residual uncertainty is in the value of the QCD Λ parameter, which must be extracted either from perturbative QCD analyses of experimental data or else from non-perturbative calculations. Values are most often quoted for $\Lambda_{\overline{MS}}$, which is related to m_X by

$$\left(\frac{m_X}{\Lambda_{\overline{MS}}}\right) = (1 \text{ to } 2) \times 10^{15} \tag{1.25}$$

in minimal SU(5) and also GUTs which reduce to it at energies $Q \lesssim m_X$. The $\Lambda_{\overline{MS}}$ parameter in equation (1.25) is that corresponding to four operational flavours $m_q \ll Q$ in two-loop QCD calculations. The remaining uncertainty in (1.25) reflects possible higher-order effects [the dots in equation (1.22)], the possible effects of varying m_H, uncertainty in the precise value (1.24) of $\alpha(m_W)$ to use, etc.

Grand unified theories with a symmetry-breaking pattern which does not take them through SU(5) at any stage, e.g. patterns of SO(10)-breaking such as[1,2]

$$SO(10) \to SU(4)\times SU(2)\times SU(2) \to \cdots \to SU(3)\times U(1) \quad , \tag{1.26}$$

may well have values of the grand unification mass rather different from (1.25), but for the next two sections we will stick with the simplest schemes.

1.4 Mixing angles in baryon decay

Now that we have the form (1.14) of the dominant B-violating interaction in simple GUTs, and know (1.25) the grand unification mass to expect in such theories, we should address the question of generalized Cabibbo mixing which was neglected in writing down (1.14) and (1.5). In particular we hope to exclude the possibility that baryons would want to try to decay into heavier quarks or leptons, which would obviously greatly suppress their decay rates! To analyse this problem in the minimal SU(5) model, we must return[15] to the fermion-fermion Higgs couplings[1,2] which take the form

$$\mathcal{L}_Y = \frac{1}{\sqrt{2}}\left(\chi^+\right)^{\alpha\beta}\gamma_0\mathcal{M}_1\left[H_\alpha\psi_\beta - H_\beta\psi_\alpha\right] - \frac{1}{4}\epsilon^{\alpha\beta\gamma\delta\epsilon}\chi_{\alpha\beta}\mathcal{M}_2 H_\gamma\chi_{\delta\epsilon} \quad . \tag{1.27}$$

We can always choose[1,2] a basis in generation space such that the matrix $\mathcal{M}_1$ is diagonal. Because of Fermi statistics, the matrix $\mathcal{M}_2$ must be symmetric in generation space, and can therefore be diagonalized by a transformation U_1 on the $\underline{10}$ representations of fermions χ:

$$\mathcal{M}_2 = U_1^T\mathcal{M}_2^D U_1 \quad . \tag{1.28}$$

If we want to, we can remove the phase factors $e^{i\phi_{ij}}$ from the first row and column of U_1, thus putting it in the standard form[16,17] of the Kobayashi-Maskawa matrix, by making further transformations U_2 and U_3:

$$U_1 = U_3 U U_2 \ : \ U_2 = \begin{pmatrix} e^{i\phi_{11}} & & 0 \\ & e^{i\phi_{21}} & \\ 0 & & \ddots \ e^{i\phi_{n1}} \end{pmatrix}, \quad U_3 = e^{-i\phi_{11}}\begin{pmatrix} e^{i\phi_{11}} & & 0 \\ & e^{i\phi_{12}} & \\ 0 & & \ddots \ e^{i\phi_{1n}} \end{pmatrix} . \tag{1.29}$$

If we absorb the matrix U_2 into the definition of the χ, and make the corresponding transformation on the ψ so as to keep $\mathcal{M}_1$ real and diagonal, the second term of (1.27) becomes

$$\mathcal{L}_Y \ni -\frac{1}{4}\epsilon^{\alpha\beta\gamma\delta\epsilon}\chi_{\alpha\beta}U^T U_3^2\mathcal{M}_2^D U H_\gamma\chi_{\delta\epsilon} \quad . \tag{1.30}$$

The charge $+\frac{2}{3}$ quark mass matrix comes from the $1 \leq \alpha,\beta,\delta,\epsilon \leq 4$ pieces of this interaction. We can diagonalize it by going to a new basis,

$$U\chi_{\alpha\beta} \equiv \chi'_{\alpha\beta} \ : \ 1 \leq \alpha,\beta \leq 4 \quad , \tag{1.31}$$

for the top left-hand 4×4 submatrix of the SU(5) matrix χ. The mass matrix is still not real, a problem we can cure by making extra phase transformations

$$U_5^2 \, \chi'_{\alpha\beta} = \chi''_{\alpha\beta} \quad : \quad 1 \leq \alpha, \beta \leq 3 \tag{1.32}$$

on the top left-hand 3 × 3 submatrix of the SU(5) matrix χ'.

It is clear from the assignments[1,2] of quarks to the matrix χ that the rotation U (1.31) rotates the left-handed charge +⅔ quarks relative to the left-handed charge -⅓ quarks, and can therefore be identified with the Kobayashi-Maskawa matrix[16,17]. The phase rotations (1.32) modify the relative phases of left-handed antiquarks relative to the charge +⅔ quarks, and hence do not show up in the conventional SU(2) weak interactions. On the other hand, they will show up in interactions involving the X and Y bosons, which couple together the first three and the last two indices of SU(5).

The most important implication of this analysis is the close connection that exists between generalized Cabibbo angles and grand unified mixing angles. This means, in particular, that baryon decay cannot be "Cabibbo-rotated away". If we incorporate mixing between the first two generations into the SU(5) effective interaction (1.5), we find in minimal SU(5)

$$\begin{aligned} \frac{1}{4}\mathcal{L}_{GU} = e^{i\phi}\frac{1}{\sqrt{2}}G_{GU}\Big[& (\epsilon_{ijk}\bar{u}^c_{k_L}\gamma_\mu u_{j_L})\Big\{\Big[(1+\cos^2\theta_c)\bar{e}^+_L + \sin\theta_c\cos\theta_c\bar{\mu}^+_L\Big]\gamma^\mu d_{i_L} + \\ & + \Big[(1+\sin^2\theta_c)\bar{\mu}^+_L + \sin\theta_c\cos\theta_c\bar{e}^+_L\Big]\gamma^\mu s_{i_L} + \bar{e}^+_R\gamma^\mu d_{i_R} + \bar{\mu}^+_R\gamma^\mu s_{i_R}\Big\} \\ & - \Big[\epsilon_{ijk}\bar{u}^c_{k_L}\gamma_\mu(d_{j_L}\cos\theta_c + s_{j_L}\sin\theta_c)\Big]\Big[\bar{\nu}^c_{e_R}\gamma^\mu d_{i_R} + \bar{\nu}^c_{\mu_R}\gamma^\mu s_{i_R}\Big]\Big] \\ & + (\text{herm. conj.}) \end{aligned} \tag{1.33}$$

Notice the occurrence of the new phase parameter, which does not of course affect baryon decay rates. It is evident from (1.33) that one has Cabibbo-favoured decay modes (into e^+ + non-strange, μ^+ + strange) and disfavoured modes (e^+ + strange, μ^+ + non-strange) whose ratios can be predicted[15] quantitatively:

$$\left.\begin{aligned} \frac{\Gamma(N \to \mu^+ + \text{non-strange})}{\Gamma(N \to e^+ + \text{non-strange})} &= \frac{\sin^2\theta_c\cos^2\theta_c}{(1+\cos^2\theta_c)^2 + 1}, \\ \frac{\Gamma(N \to e^+ + \text{strange})}{\Gamma(N \to \mu^+ + \text{strange})} &= \frac{\sin^2\theta_c\cos^2\theta_c}{(1+\sin^2\theta_c)^2 + 1}. \end{aligned}\right\} \tag{1.34}$$

If baryon decay is ever found, it will be very interesting and important to check the predictions (1.34), as they go right to the guts of our GUTs. Baryon decay may be our only window opening directly onto 10^{15} GeV physics. The predictions (1.33) and (1.34) clearly rest very strongly on our assumed assignments[1,2] of fermions to different generations, the only evidence for which comes indirectly from the patchy successes of the renormalized predictions[1,2] for $m_b \leftrightarrow m_\tau$, $m_s \leftrightarrow m_\mu$ and $m_d \leftrightarrow m_e$.

1.5 Estimate of baryon decay rates

Armed with the full effective Lagrangian (1.33), with its normalization fixed by (1.25), we are now in a position to calculate the total decay rates of protons and bound neutrons, as well as individual branching ratios. Our basic problem is to calculate the blob in Fig. 5. This calculation can be factored into two parts, one of which is the reliable computation of short distance enhancement effects, and the other is the unreliable estimate of hadronic matrix elements of operators renormalized at large distances.

The basic $qqq\ell$ interaction takes place over a distance scale $\Delta x \sim 1/m_X \sim 10^{-28}$ cm, whereas the conventional non-leptonic decay technology of hadronic physics, which we will borrow for the second part of the calculation, can in principle be used to estimate the matrix elements of operators renormalized at $\Delta x \sim 1/1$ GeV $\sim 10^{-14}$ cm. To get from the short scale to the long one we must calculate the exchanges at distances 10^{-28} cm $< \Delta x < 10^{-14}$ cm of gluons, $W^{\pm}$, Z^0, photons, etc., as shown in Fig. 6. This can be done using the renormalization group and the anomalous dimensions of the $qqq\ell$ operators. The leading-order enhancement in amplitude due to gluon exchange is[18]

$$A_3 = \left[\frac{\alpha_3(1\,\mathrm{GeV})}{\alpha_5(m_X)}\right]^{\frac{2}{11-\frac{4}{3}N_G}} , \tag{1.35a}$$

while that from $W^{\pm}$, Z^0, and γ exchange is[15]

$$A_2 A_1 = \left[\frac{\alpha_2(100\,\mathrm{GeV})}{\alpha_5(m_X)}\right]^{\frac{27}{86-4N_G}} \times \begin{cases} \left[\dfrac{\alpha_1(100\,\mathrm{GeV})}{\alpha_5(m_X)}\right]^{\frac{-69}{6+20N_G}} & \text{for operator (1.14a)} \\ \left[\dfrac{\alpha_1(100\,\mathrm{GeV})}{\alpha_5(m_X)}\right]^{\frac{-33}{6+20N_G}} & \text{for operator (1.14b)} \end{cases} \tag{1.35b}$$

The overall amplitude enhancement factor is about 3½ to 4, resulting in a decrease in the calculated decay rate by O(15).

The trickier part of the calculation is that of the hadronic matrix elements. We believe that the dominant contribution to the blob of Fig. 5 comes from the quark-quark annihilation diagram of Fig. 7. We then calculate the overlap of two quarks in the initial nucleon and the probability that the produced antiquark will combine with the spectator quark into a given meson system using either naive old-fashioned non-relativistic SU(6),

$$|p_\uparrow\rangle = |u_\uparrow(u_\uparrow d_\downarrow - u_\downarrow d_\uparrow)\rangle, \text{ etc.} , \tag{1.36}$$

or else the ultra-relativistic kinematics of the MIT bag model. Even different authors using the same calculational scheme get widely differing answers, varying over[19]

$$\tau_{p,n} = (0.6 \text{ to } 25) \times 10^{30} \text{ years} \times \left(\frac{m_X}{5\times10^{14}\,\mathrm{GeV}}\right)^4 . \tag{1.37}$$

If we take $\Lambda_{\overline{MS}}$ from the latest analyses of deep inelastic scattering data or lattice QCD calculations,

$$\Lambda_{\overline{MS}} = (100 \text{ to } 200)\ \mathrm{MeV} , \tag{1.38}$$

we obtain from (1.25) the estimate

$$m_X = (1 \text{ to } 4) \times 10^{14} \text{ GeV} \tag{1.39}$$

in minimal SU(5). Substituting into the estimated range (1.37), we finish up with

$$\tau_{p,n} = 1 \times 10^{27} \text{ years to } 1 \times 10^{31} \text{ years} . \tag{1.40}$$

Since the present experimental limit[20] on the baryon lifetime is

$$\tau_{p,n} \gtrsim (1 \text{ to } 2) \times 10^{30} \text{ years} , \tag{1.41}$$

and experiments now coming into operation should reach a sensitivity to $\tau_{p,n} \sim 10^{33}$ years[21], we can hope that the prediction (1.40) will soon be confirmed or refuted.

The same methods of Fig. 7 and either non-relativistic or ultra-relativistic kinematics can be used to estimate branching ratios into individual mesonic final states. Some typical results for minimal SU(5) are shown in the table, together with the values expected in a preferred intermediate model[22]. We see from this table that decays into relatively easily detectable modes such as $p \to e^+\pi^0$ or $n \to e^+\pi^-$ are expected to have relatively large branching ratios. This is good news for our experimental colleagues, and we wish them luck with their tests of the prediction (1.40).

Nucleon decay branching ratios in minimal SU(5)

Decay mode		Non-relativistic model	Preferred "recoil" model	Relativistic model
p	$e^+\omega$	21	25	26
	$e^+\rho^0$	2	7	11
	$e^+\pi^0$	36	40	38
	$e^+\eta$	7	2	0
	$\bar{\nu}\rho^+$	1	3	4
	$\bar{\nu}\pi^+$	14	16	15
	μ^+K^0	18	8	5
	$\bar{\nu}_\mu K^+$	0	0	1
n	$\bar{\nu}\omega$	5	5	5
	$\bar{\nu}\rho^0$	1	1	2
	$\bar{\nu}\pi^0$	8	7	7
	$\bar{\nu}\eta$	2	0	0
	$e^+\rho^-$	6	12	19
	$e^+\pi^-$	79	72	68
	$\bar{\nu}_\mu K^0$	1	3	1

1.6 Neutron-antineutron oscillations?

There has recently been some interest[23] in this phenomenon[11], which would require a $\Delta B = 2$, $\Delta L = 0$ transition. It can only be obtained from an operator containing at least six quarks, and the simplest possibility is for dimension

$$\underline{d=9} \qquad qqqqqq \text{ have } F=+1 : \text{ allowed} \tag{1.42}$$

according to the general analysis of Section 1.2. Since this operator has dimension 9, its coefficient $G_{n\bar{n}}$ must have the dimensions M^{-5}. We can make an order of magnitude analysis of n-$\bar{n}$ oscillations as follows: consider the mass matrix

$$(n, \bar{n}) \begin{pmatrix} m_n + \frac{\Delta E}{2} & \delta m \\ \delta m & m_n - \frac{\Delta E}{2} \end{pmatrix} \begin{pmatrix} n \\ \bar{n} \end{pmatrix} \tag{1.43}$$

where we have neglected any possible CP violation and included, via the terms $\pm\Delta E/2$, a possible difference in the neutron and antineutron energy levels due to the environment in which they are embedded. Possible sources of this "environmental" pollution are stray magnetic fields or interactions inside a nucleus. We can estimate the order of magnitude of

$$\delta m = [C = O(1)?] \times m_P^6/M^5 \tag{1.44}$$

in the same way as we did for the baryon lifetime (1.9). It seems that the experimental sensitivity[24] for oscillation times t is

$$t \lesssim O(10^8) \text{ s.} \tag{1.45}$$

Comparing this with equation (1.44) for δm, and setting $\delta m \sim t^{-1}$, we see that such experiments might be sensitive to

$$M \sim 10^6 \text{ GeV} \tag{1.46}$$

but probably to rather smaller M in many cases, since any simple model (see, for example, Fig. 8) will tend to have powers of small coupling constants in the denominator of (1.44), so that the uncalculated coefficient C may be rather less than unity. Corresponding as it does to an oasis in the desert, an energy scale such as (1.46) is not found in minimal GUTs, although it could be present in a non-minimal GUT such as SO(10) broken through the chain (1.26).

In fact, no $\Delta B = 2$ term of the type (1.44) could ever arise in the minimal SU(5) model, because it conserves the combination B-L. We can set up[25] non-minimal SU(5) models with a $\underline{15}$ of Higgses in which $\delta m \neq 0$, as shown in Fig. 8a, but they give very small values. The amplitude from Fig. 8a,

$$A \sim \frac{V g_{15} g_5^2}{m_{15}^2 m_5^4} , \tag{1.47}$$

where the indices on the couplings g and masses m indicate the dimensionality of the corresponding Higgs representations. From the lower limit (1.41) on the lifetime of the proton we know that

$$\frac{g_5^2}{m_5^2} \lesssim 10^{-30} \text{ GeV}^{-2} , \tag{1.48}$$

and plausible values for the other parameters in (1.47) eventually give[25] unobservable oscillation times:

$$t \gtrsim 10^{20} \text{ years !} \tag{1.49}$$

We can do better in non-minimal SO(10) broken down in the manner

$$\begin{aligned} SO(10) &\underset{\underline{54}}{\rightarrow} SU(2)_L \times SU(2)_R \times SU(4) \underset{\underline{45}}{\rightarrow} SU(2)_L \times SU(2)_R \times U(1)_{B-L} \times SU(3)_c \\ &\underset{\underline{126}}{\rightarrow} SU(2)_L \times U(1) \times SU(3)_c \underset{\underline{10}}{\rightarrow} SU(3)_c \times U(1)_{em} \quad , \end{aligned} \tag{1.50}$$

in which case the $\Delta B = 2$ transition is associated with the vacuum expectation value of the 126 of Higgses H_R. Figure 8b gives

$$A \sim \frac{\lambda g^3_{H_R} \langle 0|H_R|0\rangle}{m^6_{H_R}} \quad , \tag{1.51}$$

and if

$$\lambda = O(\alpha^2) \;,\; g_{H_R} = O(\alpha), \langle 0|H_R|0\rangle = O(10^3 \,GeV), m_{H_R} = O(10^4 \,GeV) \tag{1.52}$$

it has been estimated[25] that

$$t \sim (10^5 \text{ to } 10^6) \quad s. \tag{1.53}$$

Let us close with some remarks about the phenomenology of n-$\bar{n}$ oscillations. If they exist, we should also observe other $\Delta B = 2$ transitions such as $N + N \to$ pions in nuclei. The rate for such interactions is bounded below by the same experiments that establish the limit (1.41) on the baryon lifetime. We can then estimate[26] that

$$t \sim \frac{1}{\delta m} \approx \frac{1}{[\Gamma(N+N \to \text{pions}) \times O(1) GeV]^{\frac{1}{2}}} \gtrsim (10^6 \text{ to } 10^7) s . \tag{1.54}$$

so that there is a relatively narrow window between this and the accessible bound (1.45). To reach this sensitivity we must be very careful about the environmental pollution ΔE of equation (1.43). As an example, if we suppose that the energy difference between a neutron and a neutron in a nucleus might be $\Delta E = O(10)$ MeV, this is clearly catastrophically larger than the $\delta m \sim 10^{-30}$ GeV corresponding to $t \sim 10^8$ s, rendering oscillations inside a nucleus unobservable. In an external magnetic field B, $\Delta E = \mu_n B$, where μ_n is the neutron's magnetic moment, and for the Earth's magnetic field of O(1) gauss we find

$$\Delta E = O(10^{-11}) eV \quad , \tag{1.55}$$

which is still too large. However, if we shield the Earth's magnetic field by a factor $O(10^{-3})$ so that

$$\mu_n B \tau \ll 1 \quad , \tag{1.56}$$

where τ is the typical flight time of a neutron, of the order of 10^{-1} or 10^{-2} s in practicable experiments, then the environmental pollution no longer kills the n-$\bar{n}$ oscillations and we can expect

$$\frac{\#\ \text{antineutrons}}{\#\ \text{neutrons}} = \frac{1}{4}(\delta m \tau)^2 . \tag{1.57}$$

Several experiments[24] are now being set up to look for this phenomenon, although it should be emphasized that it is considerably more speculative than the search for baryon decay, which occurs in minimal GUTs with many less uncertainties as to the rate.

2. NEUTRINO MASSES AND OSCILLATIONS

2.1 Neutrino masses in GUTs

Neutrinos and their masses are clearly different from other fermions. We know that the neutrino masses are much smaller than those of the corresponding leptons and quarks:

$$m_{\nu_e}/m_e \lesssim O(10^{-4}) , \quad m_{\nu_\mu}/m_\mu \lesssim \frac{1}{200} , \quad m_{\nu_\tau}/m_\tau < \frac{1}{10} . \tag{2.1}$$

The most sensitive experiments on the neutrino masses come from the endpoint of tritium $\rightarrow$ $\rightarrow$ $^3He + e^- + \bar{\nu}_e$ decay[27], the $\pi \rightarrow \mu\nu_\mu$ decay[28], and the shape of the spectrum in $\tau \rightarrow \mu\nu\bar{\nu}$ decay[29]. A Russian group has recently reported[30] a positive result of a non-zero $\bar{\nu}_e$ mass from observations of tritium decay:

$$14\ \text{eV} < m_{\bar{\nu}_e} < 46\ \text{eV} , \tag{2.2}$$

and it will be exciting to see if this result is confirmed by other experiments, a point we will return to in Section 2.3.

Another peculiarity of the neutrino is that only left-handed neutrinos ν_L have ever been observed. All data on weak charged currents are consistent with their being entirely left-handed, which means that only ν_L beams are available for neutral-current studies. In contrast, the quarks and charged leptons are known to have both right- and left-handed helicity states. This enables them to acquire a "Dirac" mass m^D through couplings to Higgs fields of the type

$$g_{Hf\bar{f}} \langle 0|H_{\substack{\Delta I=\frac{1}{2}\\ \Delta L=0}}|0\rangle \bar{f}_R f_L \Rightarrow m_f^D = g_{Hf\bar{f}} \langle 0|H_{\substack{\Delta I=\frac{1}{2}\\ \Delta L=0}}|0\rangle . \tag{2.3}$$

Whether or not the right-handed neutrino ν_R exists, we have a problem to solve. If a ν_R exists, why are the neutrino masses (2.1) so much smaller than the other fermion masses? If a ν_R does not exist, can neutrinos acquire any mass at all?

Let us answer the second question first -- in the affirmative. Recall that the antiparticle of a right-handed fermion is a left-handed fermion, hence we could in general imagine replacing $\overline{f_R}$ in (2.3) by $\overline{f_L^c}$. The only difference between $\overline{f_L^c}$ and f_L is in the Lorentz structure: hereafter we will lazily write $\overline{f_L^c}$ as f_L as we are only interested in the internal symmetry properties. Replacing $\overline{f_R}$ by f_L could not give masses to the quarks or charged leptons, because $q_L q_L$ and $\ell_L \ell_L$ have non-trivial transformation properties under the exact gauge symmetry group SU(3) × U(1), and hence are not legal mass terms. However, since the neutrino has zero electric charge and is colourless, a Majorana mass term of the type

$$m^M \nu_L^T C \nu_L = m^M \bar{\nu}_L^c \nu_L = m^M \, "\nu_L \nu_L" \tag{2.4}$$

is quite legal *if we allow ourselves to violate lepton number conservation,* and can generate an interaction with weak isospin $\Delta I = 1$. The standard SU(3) × SU(2) × U(1) model[16] has lepton number conservation built into it, and only has $I = \frac{1}{2}$ Higgs fields. However, as we saw in the previous lectures, GUTs in general violate L conservation and have Higgs fields with $I \neq \frac{1}{2}$, for example in the $\underline{24}$ of minimal SU(5). Hence we might anticipate finding non-zero Majorana masses m^M (2.4) in GUTs.

There exist several possible mechanisms for generating such Majorana masses in GUTs. The simplest possibility would just be

$$\tilde{g}_{H\nu\nu} \langle 0| H_{\substack{\Delta I=1 \\ \Delta L=2}} |0\rangle \, "\nu_L \nu_L" \Rightarrow m^M = \tilde{g}_{H\nu\nu} \langle 0| H_{\substack{\Delta I=1 \\ \Delta L=2}} |0\rangle \, . \tag{2.5}$$

This possibility does not arise in minimal SU(5), where the only Higgs representation with an $I = 1$ component is the $\underline{24}$ ϕ which does not couple to pairs of $\underline{\bar{5}}$ + $\underline{10}$ fermions. However, this possibility does exist in SO(10) models containing a $\underline{126}$ of Higgs fields, which contain an $I = 1$ piece and can couple to pairs of fermions in the $\underline{16}$. Another mechanism for generating a Majorana mass (2.4) is via a 2 Higgs - 2 fermion interaction[8,31] such as

$$\frac{1}{M} \left(H_{I=\frac{1}{2}} \nu_L \right) \left(H_{I=\frac{1}{2}} \nu_L \right) \quad , \tag{2.6}$$

where we have indicated explicitly the inverse power of a mass M which enters since the interaction term has dimension 5. The interaction (2.6) is formally unrenormalizable and so normally disallowed in gauge theory. However, we know that gravity is not renormalizable in the usual sense, and quantum effects are of O(1) at the Planck mass, so we could imagine a term like (2.6) with $M \sim m_P \sim 10^{19}$ GeV maybe emerging from a theory of quantum gravity[31]. It could perhaps be generated by virtual black holes, which have no reason to conserve the global quantum number L, just as they had no reason to conserve baryon number B, as discussed in Section 1.1! Somewhat more prosaically, an interaction (2.6) could be generated by the exchange of a conventional superheavy fermion of mass M as illustrated in Fig. 9. This would give

$$m^M = O\left(\frac{\langle 0 | H_{I=\frac{1}{2}} | 0 \rangle^2}{M} \right) \quad , \tag{2.7}$$

and putting in $\langle 0|H_{I=\frac{1}{2}}|0\rangle = O(m_W)$, $M = O(m_X)(\times 10^{\pm 4}?)$ we get

$$m^M = O\left(m_W^2 / m_X \right) \ll m_{q,l} \quad , \tag{2.8}$$

consistent with the limits (2.1). [In point of fact, one would expect the same order of magnitude (2.8) from the alternative source (2.5) -- see Ref. 23.]

We have crossed a Rubicon in drawing Fig. 9. We have introduced a new neutral fermion field: why could it not be the long-lost ν_R? and then why could we not have a simple "Dirac" mass term (2.3)

$$g_{H\bar{\nu}\nu} \langle 0| H_{I=\frac{1}{2}} |0\rangle \bar{\nu}_R \nu_L \Rightarrow m^D = g_{H\bar{\nu}\nu} \langle 0| H_{I=\frac{1}{2}} |0\rangle \,? \tag{2.9}$$

The simplest way to permit (2.10) is for the ν_R to have $I = 0$. In this case it would be simplest to assign it to a singlet of SU(5), and such an object is found, for example, in the $\underline{16}$ of fermions used in SO(10) models[1,2]. But if such a term (2.9) exists, why is it not the case that

$$m_\nu \approx m^D = O(m_W) \;, \text{ or at least } O(m_q, m_l) \;? \tag{2.10}$$

The answer is provided by the observation[32] that once we have ν_R as well as ν_L, we have to consider the most general form of mass matrix whose internal symmetries are shown explicitly below:

$$(\nu_L \;, \bar{\nu}_R) \begin{pmatrix} m^M & m^D \\ m^D & M^M \end{pmatrix} \begin{pmatrix} \nu_L \\ \bar{\nu}_R \end{pmatrix} . \tag{2.11}$$

We have already met the top left and off-diagonal elements of (2.11): they are expected to be of order $O(m_W^2/m_X)$ (2.8) and $O(m_W)$ (2.10), respectively. The new element in (2.11) is in the bottom right-hand corner. If the ν_R is a singlet of SU(5) as we hypothesized above, then a Majorana $\bar{\nu}_R\bar{\nu}_R$ mass term is SU(5) invariant and hence can be as large as it likes, and quite possibly much larger than m_X. If we now diagonalize the mass matrix (2.11) we find two mass eigenstates:

$$\nu_m \equiv \nu_L + O(m_W/m_X)\bar{\nu}_R \;:\; m_{\nu_m} = O(m_W^2/m_X) \;, \tag{2.12a}$$

$$N \equiv \bar{\nu}_R + O(m_W/m_X)\nu_L \;:\; m_N = O(m_X) \;, \tag{2.12b}$$

where we should allow several orders of magnitude latitude in the estimates (2.12). For example, perhaps

$$m_{\nu_m} = O\left(\frac{m_l^2 \text{ or } m_q^2}{m_X}\right) \;? \tag{2.13a}$$

or perhaps

$$m_N = O(m_X \times 10^{\pm 4}) = O(m_P \text{ or } 10^{11}\,\text{GeV}) \;? \tag{2.13b}$$

In writing (2.11), (2.12), and (2.13) the generation degree of freedom has been neglected. However, it should be borne in mind that there is no particular reason why the mass eigenstates ν_m should be same as the weak eigenstates ν_e, ν_μ, ν_τ, ..., and we will return to the implications of this fact in Section 2.2.

So far we have made a general analysis punctuated by illustrations in different GUTs. Let us now look systematically at what happens in some popular models. As observed before, in *minimal SU(5)* there is only a ν_L [so we do not need the matrix (2.11)] and B-L is strictly conserved, so that one would naively expect $m_\nu = 0$. But the point has already been made that it is only a global symmetry that protects the neutrino mass, and we may expect these to be broken at the Planck mass, even if not before. Therefore we could anticipate[31] finding an interaction of the type (2.6) with $M \sim m_P$, and hence a Majorana mass (2.7) of order

$$m_\nu = O\left(\frac{m_W^2}{m_P}\right) = O(10^{-5})\,eV \quad , \tag{2.14}$$

which might turn out to be a sort of lower bound on the neutrino mass, as we shall see. In *non-minimal SU(5)* models we can easily introduce ν_R as singlets of SU(5), introduce new Higgses such as a $\underline{15}$ of SU(5), and violate B-L conservation. Since we already have a $\underline{5}$ of Higgs, and

$$\underline{5} \times \underline{1} = \underline{5} \quad , \tag{2.15}$$

we can easily generate an m^D for (2.11). We can get an m^M from a $\underline{15}$ of Higgs, since

$$\underline{5} \times \underline{5} = \underline{10} + \underline{15} \tag{2.16}$$

and the $\underline{15}$ contains a weak isotriplet of Higgses whose vacuum expectation value is naturally of order (m_W^2/m_X). We can get an M^M term from an SU(5) invariant interaction, and hence go down the standard route (2.11) to (2.13) and easily get[32]

$$m_\nu = O(1)\,eV\,. \tag{2.17}$$

Even the *minimal SO(10)* model[1,2] has a ν_R, and B-L is necessarily violated, since it is a broken generator:

$$\begin{array}{l} SO(10) \supset SU(4) \times SU(2)_L \times SU(2)_R \\ \qquad\qquad \hookrightarrow \supset U(1)_{B-L} \end{array} \tag{2.18}$$

[see also equation (1.50)]. Once again we can get contributions to all entries in the neutrino mass matrix (2.11), using the SO(10) Higgs representations shown below:

$$(\nu_L, \bar{\nu}_R)\begin{pmatrix} \underline{126}_{15} & \underline{10}_5 \\ \underline{10}_{\bar{5}} & \underline{126}_1 \end{pmatrix}\begin{pmatrix} \nu_L \\ \bar{\nu}_R \end{pmatrix} \quad , \tag{2.19}$$

where we have indicated by subscripts the SU(5) representation content of the contributing Higgs fields. At this point you might object that minimal SO(10) [1,2] actually does not contain an explicit $\underline{126}$ of Higgses. However, it has been pointed out[33] that even in this case one can get an "effective" $\underline{126}$ from higher-order diagrams as in Fig. 10. In this case we get

$$M^M = O\left(\frac{\alpha}{\pi}\right)^2 m_X \quad , \tag{2.20}$$

and the off-diagonal elements m^D are naturally of order m_ℓ or m_q, so that

$$m_{\nu_m} \approx \frac{m_\ell^2 \text{ or } m_q^2}{O\left(\frac{\alpha}{\pi}\right)^2 m_X} \lesssim O(10)\,\text{eV}\ , \tag{2.21}$$

which sets an upper bound to our range of neutrino masses expected in GUTs.

It is possible to get masses larger than (2.21) at the expense of introducing an intermediate energy scale between m_W and m_X, for example the scheme (1.50) discussed in connection with n-$\bar{\text{n}}$ oscillations. And it is of course possible to get neutrino masses without invoking GUTs at all. Hence neutrino masses do not constitute such a crucial and specific test of GUT ideas as does baryon decay. However, GUTs suggest the following general prejudices which are experimentally interesting:

a range: $$O(10^{-5})\,\text{eV} \lesssim m_\nu \lesssim O(10)\,\text{eV} \tag{2.22a}$$

a hierarchy: $$m_{\nu_1} : m_{\nu_2} : m_{\nu_3} \approx m_{e,u}^{1\text{ or }2} : m_{\mu,c}^{1\text{ or }2} : m_{\tau,t}^{1\text{ or }2} \tag{2.22b}$$

mixing: $$\nu_1\ , \nu_2\ , \nu_3 \neq \nu_e\ , \nu_\mu\ , \nu_\tau\ , \tag{2.22c}$$

whose implications we shall now investigate.

2.2 Phenomenology of neutrino oscillations[34)]

The conventional charged currents of neutrinos:

$$J_\mu = \bar{e}\gamma_\mu(1-\gamma_5)\nu_e + \bar{\mu}\gamma_\mu(1-\gamma_5)\nu_\mu + \bar{\tau}\gamma_\mu(1-\gamma_5)\nu_\tau \tag{2.23}$$

automatically produce, either in decays or scattering, weak interaction eigen-neutrinos ν_ℓ. These are in general related to mass eigen-neutrinos ν_m by a unitary mixing matrix

$$\nu_\ell = U_{\ell m}\,\nu_m \tag{2.24}$$

which will normally generate oscillations. Let us consider a beam which is initially of pure ν_e, and has a well-defined momentum $\underline{p}$. [We could have considered a well-defined energy E, or a more general wave-packet, but the bottom line (2.28) would be the same[35)].] As time passes and the beam propagates[34)], the different mass eigen-neutrinos propagate with slightly different phase factors:

$$|\nu_e\rangle \to U_{e1}|\nu_1\rangle e^{i(\underline{p}\cdot\underline{x}-E_1 t)} + U_{e2}|\nu_2\rangle e^{i(\underline{p}\cdot\underline{x}-E_2 t)} + \cdots\ , \tag{2.25}$$

where

$$E_i \approx |\underline{p}| + \frac{m_i^2}{2|\underline{p}|} + \cdots\ . \tag{2.26}$$

For simplicity, we will henceforward truncate equation (2.25) at $|\nu_2\rangle$, and assume

$$E_i \approx |\underline{p}| \equiv E \gg m_i \tag{2.27}$$

so that the neutrinos propagate with essentially the speed of light and we can neglect the dots in equation (2.26). The different phase factors in (2.25) mean that an initial ν_e eventually becomes a superposition of ν_e, ν_μ, and ν_τ. For example, the persistence probability

$$P(\nu_e \to \nu_e) = \left| \sum_{i,j} U_{ei} \exp\left[\frac{i(m_i^2 - m_j^2)R}{2E} \right] U_{ej}^* \right|^2 , \tag{2.28}$$

where R is the distance of propagation, is clearly less than unity in general. It is convenient to define oscillation lengths

$$L_{ij} \equiv \frac{4\pi E}{|m_i^2 - m_j^2|} , \tag{2.29}$$

which in convenient units becomes

$$L_{ij} \begin{Bmatrix} \text{metres} \\ \text{km.} \end{Bmatrix} = \frac{2 \cdot 5 \times E \begin{Bmatrix} \text{MeV} \\ \text{GeV} \end{Bmatrix}}{|m_i^2 - m_j^2| (\text{eV}^2)} . \tag{2.30}$$

In the simple case of two neutrino species there is just one mixing angle θ:

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{2.31}$$

and the transition probability

$$P(\nu_1 \to \nu_2) = \frac{1}{2} \sin^2 2\theta \left(1 - \cos \frac{2\pi R}{L_{12}} \right) . \tag{2.32}$$

The latest experimental results on neutrino oscillations, as well as neutrino masses, are discussed in the lectures of Perkins[21]).

3. GRAND UNIFIED THEORIES AND COSMOLOGY

3.1 Introduction

Particle physics and cosmology have grown closer in the past few years, and most of the new developments concern GUTs. The two most dramatic aspects of the interaction between high-energy physics and cosmology involve the very early Universe and the origin of matter on the one hand, and the possible influence of neutrinos on the future destiny of the Universe on the other hand. In both these physical problems the collaboration between macroscopic and microscopic physics is a two-way street. GUTs can explain the origin of the present predominance of matter over antimatter in the Universe, but only provided that the GUTs have certain crucial properties. Neutrinos may dominate the future evolution of the Universe if they are sufficiently massive, and from its present expansion we can already deduce upper limits on the neutrino masses which are barely compatible with the theoretical ideas presented in the previous lecture.

The interfaces between high-energy physics and cosmology are growing so rapidly that these notes contain no attempt at a comprehensive review of current progress. Rather, an attempt will be made to bring out the relevant basic physical principles of the two disciplines in such a way that future developments can be placed in context. Subjects barely touched upon in this lecture include grand unified monopoles and phase transitions in the early Universe, nucleosynthesis limits on the number of neutrino species, and the possible influence of massive neutrinos on galaxy formation and dynamics. The interested reader should consult Refs. 5 and 36 for more details on these subjects.

3.2 An introduction to Big Bang cosmology

The Big Bang is by now the standard model of cosmology[37], and the particle physicist whose interests lie mainly elsewhere may like to be reminded of the reasons why the Big Bang theory is so favoured. There are three main phenomenological arguments for this theory. The most obvious one is the Hubble expansion: all distant galactic clusters and quasars seem to be receding from us and from each other at velocities proportional to the distances separating them in a ratio

$$H_0 = 100 h_0 \text{ km./s. per megaparsec of separation}, \tag{3.1}$$

where observations now suggest that

$$0.4 \lesssim h_0 \lesssim 1 \tag{3.2}$$

Furthermore, the present distribution and expansion of galactic clusters in the Universe seem to be essentially homogeneous and isotropic when viewed on a sufficiently coarse scale. A naive extrapolation to the distant past suggests that there should have been some sort of singularity then, when all the presently diffuse matter was compressed together, and there are indeed theorems[38] that a singularity lurks in the past light-cone of each one of us. These theorems do not prove that all of us come from the same singularity, but this is the case with the Robertson-Walker-Friedmann solutions[37] which are the only ones known that can adequately describe the present large-scale structure of the Universe. There are two apparent "footprints" of this earlier epoch when the matter in it was much more compressed: the 3 K microwave background radiation, and the abundances of light nuclei. Any thermodynamic system heats up when compressed, and the Universe is no exception. When it was about a millionth of its present age it would have been sufficiently hot for all the matter in it to have been ionized and in thermal equilibrium with a bath of photons. As the Universe cooled, the electrons and nuclei would have combined to form atoms and molecules, leaving the photons thermally isolated. These photons should still be present today in a thermal Boltzmann distribution at a lower temperature cooled by the universal expansion. Indeed, such photons have apparently been observed[39] -- the 3 K microwave background radiation with a photon density

$$n_\gamma \approx 400 \left(\frac{T_0}{2.7^\circ\text{K}}\right)^3 \text{ per cm}^3 \tag{3.3}$$

for which no convincing alternative explanation has been found. Going back earlier in history, the Universe should have been sufficiently compressed and hot early in its evolution for nuclear reactions to have been common. These should have combined protons and neutrons into light nuclei, mainly ^{4}He, which should have survived the subsequent Hubble expansion and cooling to be visible in the Universe today. Indeed, throughout the Universe, about 25% to 30% of the mass density seems to be in the form of ^{4}He, and only a few percent can be easily explained by astrophysical mechanisms such as nuclear reactions in first-generation stars. Thus it is believed that there is a universal primordial helium abundance Y of about 20% to 25%, which is taken as the third piece of evidence for the Big Bang[37].

If we accept this theory, we now need to know more about its mathematical formalism. General relativity provides two independent equations to govern the expansion[37]:

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G_N}{3}\rho - \frac{k}{R^2} \tag{3.4a}$$

and

$$\left(\frac{\ddot{R}}{R}\right) = -4\pi G_N\left(p + \frac{\rho}{3}\right) . \tag{3.4b}$$

In these equations R is an overall scale factor, G_N is Newton's constant $\approx 10^{-38}$ $\mathrm{GeV}^{-2} \equiv$ $\equiv (m_P)^{-2}$, ρ is the energy density, p the pressure, and k the curvature parameter which is -1 for an asymptotically open Universe, +1 for an asymptotically closed Universe, and 0 for a borderline "flat" Universe. We can neglect k during the early Universe, which interests us at the moment. For radiation or other ultra-relativistic matter in thermal equilibrium,

$$3p = \rho = \frac{1}{2}\, g(T)\rho_\gamma(T) \approx 3T\,n(T) \tag{3.5}$$

where T is the temperature, and g(T) is the total number of helicity states excited at the temperature T:

$$g(T) = (\#\text{ boson helicities}) + \frac{7}{8}(\#\text{ fermion helicities}) . \tag{3.6}$$

According to conventional thermodynamics, the density of photons $\rho_\gamma(T)$ is

$$\rho_\gamma(T) = \frac{\pi^2}{30}T^4 \approx 0.3\,T^4 \tag{3.7}$$

and n(T) (the number density of particles at the temperature T) is also specified. For non-relativistic matter of different types i,

$$p = 0 \quad , \quad \rho = \sum_i m_i n_i(T) \tag{3.8}$$

where the number density

$$n_i(T) \approx T^{3/2} e^{-m_i/T} \tag{3.9}$$

if species i is in thermal equilibrium, but

$$n_i(T) = c_i T^3 \tag{3.10}$$

if the number of i particles is conserved [as approximately (?) for baryons during the present epoch] and the Universe undergoes adiabatic expansion

$$RT = \text{constant} . \tag{3.11}$$

This is the expected expansion law if $\rho(T)$ is dominated by relativistic or non-relativistic matter. However, there is a third possible source of energy density: the vacuum which could yield

$$\rho = -\langle 0|\mathcal{L}|0\rangle = +\langle 0|\mathcal{H}|0\rangle , \tag{3.12}$$

which acts like a cosmological term Λ in Einstein's equations and would give an exponential expansion if it were dominant in (3.4a). If the Universe is radiation dominated, however, the above equations yield

$$\left.\begin{array}{l} R \propto t^{\frac{1}{2}} \\ t(s.)\,[T(\mathrm{MeV})]^2 = 2{\cdot}42\,[g(T)]^{-\frac{1}{2}} \end{array}\right\} , \tag{3.13a}$$

whereas in a matter-dominated Universe

$$R \propto t^{2/3} \tag{3.13b}$$

and the adiabatic property (3.11) is obeyed in both cases.

We have already been talking about thermal equilibrium and will need to discuss it again in connection with baryon number generation. Crudely speaking, equilibrium will be established if the particle interaction rates are faster than the expansion rate. The collision rate for a particle of type i is

$$\Gamma_i(T) = \sum_j n_j(T) \langle \sigma v \rangle_{ij,T} , \tag{3.14}$$

where $\langle \sigma v \rangle_{ij,T}$ is the thermal average of the cross-section times velocity for ij collisions. If

$$E \equiv \Gamma_i(T)/(\dot{R}/R) > 1 , \tag{3.15}$$

then we should expect thermal equilibrium and Boltzmann distributions for all the particles *unless* there are chemical potentials corresponding to some conservation law. We will study later to what extent the condition (3.15) may have applied to elementary particle interactions in the very early Universe.

3.3 The baryon-antibaryon asymmetry of the Universe

It is generally believed that the Universe is globally asymmetric between matter and antimatter[40]. Clearly we and the contents of our immediate neighbourhood are predominantly made of matter -- otherwise we would see the products of matter-antimatter annihilation all around us, in the form of X-rays, γ-rays, energetic particles, etc. In fact, the upper limits on matter-antimatter annihilation in the Universe are so strong that it seems no substantial concentrations of antimatter can exist in our local cluster of galaxies. Antiprotons have recently been discovered[41] among cosmic rays, but their flux relative to protons is small enough to be consistent with them being the products of secondary collisions involving primary particles made of matter alone. There is therefore no need to postulate any large source of antimatter "out there". Physicists have not been able to find a plausible mechanism in a matter-antimatter symmetric Universe for separating to distances of millions of megaparsecs the concentrations of matter and antimatter. It is therefore generally believed that the "local" predominance of matter betokens a global asymmetry throughout the Universe[42].

If this is so, we are naturally led to ask how much matter there is dispersed in the Universe. Clearly the matter density ρ_m in the neighbourhood of the writer and reader is locally high, but the mean averaged over intergalactic space is rather low[36]:

$$\rho_m \approx 10^{-31\pm1}\ \mathrm{g/cm^3} \tag{3.16}$$

There are considerable uncertainties in ρ_m owing to the possible existence of invisible non-luminous matter and the overall scale of the Universe, which is known only approximately [cf. the uncertainty (3.2) in the Hubble expansion rate (3.1)]. Multiplying the matter density (3.16) by Avogadro's number $\sim 6 \times 10^{23}$, we get a baryon number density

$$n_B = O(10^{-7}) \text{ per cm}^3 \tag{3.17}$$

This is appreciably less than the density of photons in the 3 K background radiation (3.3) with a ratio

$$\frac{n_B}{n_\gamma} = O(10^{-9\frac{1}{2}}) \quad . \tag{3.18}$$

We might therefore conclude that the Universe contains rather little matter. However, from another point of view, the Universe contains a surprisingly large amount of matter. *If* the Universe had been *matter-antimatter symmetric* when it was compressed to a temperature comparable to the nucleon mass (10^{13} K $\leftrightarrow \sim 1$ GeV), then nucleon-antinucleon annihilation[40]) would have been so efficient that we would nowadays find

$$\frac{n_B}{n_\gamma} \lesssim O(10^{-20}) \tag{3.19}$$

much smaller than the deduced density (3.18). It therefore seems that the Universe must have been *matter-antisymmetric* at a temperature $\gtrsim 1$ GeV. At this epoch it is convenient to describe strongly interacting matter in terms of quark and gluon degrees of freedom, and the present baryon density (3.18) translates into a primordial quark-antiquark asymmetry

$$\frac{n_q - n_{\bar{q}}}{n_q + n_{\bar{q}}} \approx O(10^{-10}) \quad . \tag{3.20}$$

How and why could this asymmetry have arisen?

Of course it could have been put in as an initial condition by whoever pushed the button to inaugurate the Big Bang, but more scientific explanations are possible. It would be attractive to understand the asymmetry (3.20) in terms of the known elementary particle interactions and their behaviour in the very early Universe. The general requirements for successfully generating a baryon asymmetry in this way were set out in a pioneering paper by Sakharov[43]) in 1967. Clearly one needs interactions that violate baryon number conservation if one is to change an initially symmetric B = 0 state into an asymmetric B $\neq$ 0 state. Furthermore, these B violating interactions must also violate charge conjugation C invariance, since a C transformation would interchange the quark and antiquark densities n_q and $n_{\bar{q}}$, and so would enforce $n_q = n_{\bar{q}}$ if C were exact. The C violation alone is not enough -- the interactions must also violate CP. This is because, unlike C, a P transformation leaves n_q and $n_{\bar{q}}$ unchanged, so that a combined CP transformation also enforces $n_q = n_{\bar{q}}$ if it is exact. That C is violated in the conventional weak interactions has been known since the experiment of Mme Wu in 1957. That CP is also violated has been known since the experiment of Cronin and Fitch in 1964. The GUTs discussed in earlier lectures also violate these symmetries as well as B conservation. The baryon-number changing currents (1.4) contain both vector and axial pieces, and hence violate C. Since the grand unified mixing angles are so closely related to the Cabibbo-Kobayashi-Maskawa matrix[17]), they will also violate CP in general, and the specific SU(5) model discussed in previous lectures has extra sources of CP violation as well [cf. the $e^{i\phi}$ factor in equation (1.33)]. Therefore, GUTs seem well placed for generating baryon number. However, there is a further condition they must obey, namely that they drop out of thermal equilibrium in the early Universe. The reason is that in an equilibrium state we lose sense of the arrow of time, and the CPT theorem sacred to quantum field theory then ensures that the states are CP invariant and hence have $n_q = n_{\bar{q}}$. To show that GUTs can indeed generate a baryon asymmetry, we must check that their interactions go out of equilibrium in an appropriate way.

3.4 A GUT mechanism for baryon generation[44]

We must consider the rates for various elementary particle interactions in the early Universe when the temperature $T \sim 10^{28}$ K $\leftrightarrow 10^{15}$ GeV. For convenience we will drop numerical factors of order 1 and use Planck units in which 10^{19} GeV $\equiv$ unity. Dimensional arguments lead one to expect that the high-temperature cross-sections due to the exchanges of particles of negligible mass are approximately

$$\sigma \sim \frac{\alpha^2}{T^2} \rightarrow \frac{\alpha^2 T^2}{m_H^4} \tag{3.21}$$

when the exchanged particle has a heavy mass $m_H >> T$. One might wonder whether infrared singularities associated with massless particle exchange might alter the first of the forms (3.21), but it seems likely that Debye screening restores the naive dimensional form. The interaction rate $\Gamma(3.14) \approx T^3\sigma$ becomes

$$\Gamma \sim \alpha^2 T \rightarrow \frac{\alpha^2 T^5}{m_H^4} \tag{3.22}$$

for "massless" and massive particle exchange, respectively. When compared with the expansion rate $(\dot{R}/R)$ we get from equation (3.15),

$$E \equiv \frac{\text{Interaction Rate}}{\text{Expansion Rate}} = \frac{\Gamma R}{\dot{R}} \sim \frac{\alpha^2}{T} \rightarrow \frac{\alpha^2 T^3}{m_H^4}, \tag{3.23}$$

and $2 \leftrightarrow 2$ interactions are in equilibrium if $E > 1$. We should also consider the rates for $1 \leftrightarrow 2$ processes involving heavy particles -- the decays and inverse decays of Fig. 11. The rate for these is expected to be

$$\Gamma \sim \frac{\alpha\, m_H^2}{\sqrt{T^2 + m_H^2}}\, S(T) \tag{3.24}$$

where the conventional decay rate $\Gamma_0 \sim \alpha m_H$ has been Lorentz-dilated to take acount of the high velocity with which even a heavy particle may travel at high temperature T. The factor S(T) is a fudge factor which may be a Boltzmann factor $\sim e^{-m_H/T}$ for $2 \rightarrow 1$ heavy-particle production and also $1 \rightarrow 2$ decay if the heavy particle H still has an equilibrium distribution, or S(T) will be $\sim e^{-\Gamma t}$ if the particle decays out of equilibrium.

Qualitative features of the equilibrium ratios E (3.23) for different interactions are shown in Fig. 12. (Modestly) we only discuss $T < 1 \equiv 10^{19}$ GeV: above this temperature, quantum gravitational phenomena are presumably O(1) and not (yet) manageable, whereas below this temperature they apparently decline as some power of T and are out of thermal equilibrium.

Conventional strong, weak, and electromagnetic interactions presumably follow the first form for E in (3.23) and are hence in equilibrium at "low" temperatures $\lesssim 10^{15}$ GeV. It may well be that these and all other grand unified interactions were out of equilibrium when $T > 10^{15}$ GeV, as indicated in Fig. 12: but this may require too much confidence in the dimensional cross-section of (3.22). At temperatures below the grand unification scale $m_X \approx 10^{15}$ GeV, the second formula for E in (3.23) suggests that $2 \leftrightarrow 2$ grand unified interactions mediated by X gauge-boson exchange are certainly out of equilibrium. We see from Fig. 12 that this is also true for $1 \leftrightarrow 2$ interactions involving X bosons, and also probably for $2 \leftrightarrow 2$ and $1 \leftrightarrow 2$ interactions involving superheavy Higgs bosons H_X, whose masses we arbitrarily set $\simeq 10^{14}$ GeV for illustrative purposes. It therefore seems that when T fell below 10^{14} GeV or so, the stage may have been set for baryon generation by out-of-equilibrium grand unification interactions.

A possible specific mechanism[44]) for baryon generation is via the out-of-equilibrium decays of some superheavy boson, a vector X boson or a Higgs boson H_X. If we consider the decays dominant in simple models such as SU(5) or SO(10),

$$H_X \text{ or } X \rightarrow q + q \text{ or } \bar{q} + \bar{\ell} , \tag{3.25}$$

then CPT invariance guarantees that the total decay rates of heavy particles and antiparticles will be identical:

$$\Gamma_{tot}(X) = \Gamma(X \to qq) + \Gamma(X \to \bar{q}\bar{\ell}) = \Gamma_{tot}(\bar{X}) = \Gamma(\bar{X} \to \bar{q}\bar{q}) + \Gamma(\bar{X} \to q\ell) , \tag{3.26}$$

but individual decay rates may differ if C and CP are violated:

$$B \equiv \frac{\Gamma(X \to \bar{q}\bar{\ell})}{\Gamma_{tot}(X)} \neq \bar{B} \equiv \frac{\Gamma(\bar{X} \to q\ell)}{\Gamma_{tot}(\bar{X})} . \tag{3.27}$$

It seems plausible that primordially, or as a result of the equilibrium period hinted at in Fig. 12, the densities $n_X = n_{\bar{X}}$ before the heavy particles start to decay. Then the C- and CP-violating asymmetry (3.27) will generate a net quark asymmetry

$$\frac{n_q - n_{\bar{q}}}{n_q + n_{\bar{q}}} \approx \left(\frac{g_X}{g(T)}\right)(\bar{B} - B) \approx (10^{-1} \text{ to } 10^{-2})(\bar{B} - B) \tag{3.28}$$

and hence a net baryon asymmetry and eventually the present non-zero n_B/n_γ ratio. Numerical results[45]) of a typical such calculation are shown in Fig. 13. The lowest order in which such a decay asymmetry $\Delta B \equiv \bar{B} - B$ can be generated is fourth order, via a final-state interaction interference diagram of the type shown in Fig. 14:

$$\Delta B = \frac{\Delta\Gamma}{\Gamma} \approx O\left(\frac{1}{\pi}\right) \frac{\mathrm{Im}\, \mathrm{Tr}(abcd^*)}{\sum_a \mathrm{Tr}(aa^*)} , \tag{3.29}$$

which is probably $\lesssim 10^{-3}$ in most plausible models. When substituted into (3.28) this is amply sufficient to get the required baryon asymmetry (3.18), although many more considerations must be taken into account if this argument is to be made quantitative. However, at least a qualitative mechanism for generating the observed baryonic matter in the Universe has been found.

In simple models the CP-violating decay asymmetry ΔB is larger in Higgs H_X decay than in vector X decay. Actually, the lowest-order suitable diagrams in minimal SU(5) are eighth order[46], and would yield

$$\Delta B \lesssim \frac{1}{10}\left(\frac{\alpha}{\pi}\right)^3 \frac{m_b^4 m_t m_c}{m_W^6} \lesssim 10^{-14}\left(\frac{m_t}{m_W}\right) \quad ? \tag{3.30}$$

which is far too small to give the observed asymmetry (3.18). The CP-violating asymmetry in vector X-boson decay would be $O(\alpha/\pi)$ smaller still, so we conclude that a larger GUT than minimal SU(5) is necessary. Such a non-minimal model -- even SU(5) with two $\underline{5}$ representations of Higgs would suffice -- would contain even more undetermined parameters than the minimal model and make it difficult to perform a quantitative calculation. For this reason it is also unlikely that we can relate the baryon asymmetry directly to the CP violation seen in the K_0 - $\bar{K}_0$ system, e.g. to the sign of the CP violating ε parameter. In fact, even in minimal SU(5) there is no direct connection to this observed measure of CP violation, as the dominant sources of the (wrong) estimate (3.29) are the extra phases [cf. the $e^{i\phi}$ factor in (1.33)] which are not observable in low-energy weak interactions. However, it seems that in most GUTs, diagrams similar to those generating the baryon number also contribute to the (as yet unobserved) CP-violating neutron electric dipole moment[47]. We can then speculate on a possible order of magnitude lower bound on the dipole moment d_n:

$$d_n \gtrsim 2\times 10^{-18}\left(\frac{n_B}{n_\gamma}\right) \gtrsim 5\times 10^{-28}\ e\text{-cm} , \tag{3.31}$$

which is about three orders of magnitude below the present experimental upper limit. In many GUTs the actual value of d_n would even be considerably larger than (3.31).

It is also worth noticing that the calculation of the baryon number is rather sensitive to the masses and couplings of the superheavy decaying bosons: the results of typical calculations[45] for different masses m_X and couplings α_X of the relevant superheavy particles are shown in Fig. 15. What can happen is that if m_X is low enough (or α_X large enough), 2 $\leftrightarrow$ 2 collisions (cf. Fig. 12) can dilute or wash out the asymmetry generated in superheavy particle decays. If we require the dilution factor to be at most a factor of 10^3, then Fig. 15 indicates that

$$\left.\begin{aligned} m_X &\gtrsim 10^{16}\ \text{GeV} \quad \text{if } \alpha_X = 0.1 , \\ &\text{or } 10^{13}\ \text{GeV} \quad \text{if } \alpha_X \sim 10^{-3} . \end{aligned}\right\} \tag{3.32}$$

It seems more likely that an asymmetry from vector X-boson decay would be diluted than an asymmetry from Higgs H_X decay, since the Higgs coupling strength can be expected to be smaller.

For all the reasons listed above it is not yet possible to make a quantitative calculation of the baryon number in a plausible GUT, but the existence of a qualitative mechanism for the origin of the matter asymmetry seems established.

3.5 Neutrinos and the future of the Universe

We finish this lecture by mentioning two interfaces between neutrino physics and cosmology, with particular emphasis on the implications of neutrino masses for the future evolution of the Universe. We have seen in previous lectures that GUTs make no firm prediction for the total number of fermion generations, although the semi-quantitative success[1,2] of the prediction of the bottom quark mass in terms of the τ lepton mass suggests that there may only be three generations. At present, experimental restrictions on the total number of generations can only come from restrictions on the possible number of (almost) massless neutrino types. The best current limit comes from the upper limit on $K^+ \to \pi^+ \nu\bar{\nu}$ decay, which suggests[2] that

$$N_\nu \lesssim O(10^{-5}) \quad , \tag{3.33}$$

although much better limits will eventually be available from those on $Z^0 \to \nu\bar{\nu}$ decays. A cosmological limit on the number of neutrino types comes from the success of the primordial nucleosynthesis calculations mentioned in Section 3.2 as one of the three main reasons for believing in the standard Big Bang cosmology. The point is that, other things being equal, adding into the primordial soup an extra neutrino with mass $\lesssim O(0.1)$ MeV would have increased the Universe's expansion rate at temperatures $\sim 10^9$ K, thereby increasing the number of decaying neutrons still available for capture by protons to make light nuclei such as ^{4}He. Calculations[48] indicate that the effect of adding in a light two-component neutrino is to add to the ^{4}He abundance Y an amount

$$\Delta Y \approx 0.01 \ . \tag{3.34}$$

Since the primordial ^{4}He abundance was probably $\lesssim 25\%$, this indicates that we get a much more stringent upper bound on N_ν than equation (3.32). The precise number we get depends on our guess about the present matter density in the Universe. It seems[48] that in order to get right the abundances of deuterium and ^{3}He, $n_B/n_\gamma > 3 \times 10^{-10}$ in which case $N_\nu < 3$ or 4.

Cosmology can also be used to give interesting limits on the masses of neutrinos. At a temperature around 1 MeV all neutrinos would have decoupled from matter, in the same way as the photons did at the time of recombination. Then in the same way as those photons are now detectable as the 3 K microwave background radiation with the number density (3.3), so also the neutrinos should be all around us in comparable numbers. Their number density would be so high compared with the baryons (3.17) that they might outweigh the baryons even if the sum of their masses were as small as a few eV. The most reliable way of using this physical observation seems to be to use the age of the Universe, which is sensitive to the total neutrino mass density:

$$t_0 \approx \frac{2 \times 10^{10}}{2h_0 + \sqrt{\dfrac{\sum_\nu m_\nu}{19\,\mathrm{eV}}}} \ \text{years} \tag{3.35}$$

if microwave background temperature is taken as 2.65 K. A lower limit on t_0 comes from the cosmochronology of ^{232}Th/^{238}U and ^{187}Re/^{187}Os, which suggests[49]

$$t_0 \gtrsim 8.7 \times 10^9 \ \text{years} \ . \tag{3.36}$$

Plugging this into (3.35) we find that if $h_0 \gtrsim 0.4$, as suggested earlier in (3.2), then

$$\sum_{\nu} m_{\nu} \lesssim 43\,\mathrm{eV}\,. \tag{3.37}$$

This is far more restrictive than the experimental upper limits (2.1) on the ν_μ and ν_τ masses and is comparable with the experimental limit (value) of the ν_e mass (2.2), while the GUT estimates (2.22) of the neutrino masses are compatible with the cosmological bound (3.37). If the experimental finding of a ν_e mass $\simeq$ 30 eV is correct, then a very interesting situation arises.

A neutrino mass $\sim$ 30 eV is sufficient to close the Universe, entailing a future halt to the Hubble expansion and its eventual reversal to fall back into an anti-Big Bang some 10^{11} years hence. On the other hand, if the sum of the neutrino masses were less than $\sim$ 30 eV, then the expansion would continue indefinitely. As time progressed, according to Lecture I eventually (after 10^{31} years or so) all the matter in the Universe would decay into leptons and photons, implying a rather dismal future for our Universe. In fact, things would be more exciting than just watching matter ebb away: in addition, black holes would form as gravitational inhomogeneities, grow, and collapse, and these black holes would in turn decay over extremely long time-scales through quantum mechanical Hawking radiation.

Whether the Universe continues its present expansion for ever, or falls back into another singularity, it seems clear that GUTs will have a lot to say about the distant future of the Universe, as well as about its origin.

Acknowledgements

I would like to thank many of the students at the School for interesting discussions, in particular T. Haudt, M. Jeżabek, M. Kowalski, K. Lang and S. Tkaczyk.

REFERENCES

1) M.B. Voloshin, these Proceedings.

2) J. Ellis, preprint LAPP TH-48/CERN TH-3174 (1981) and references therein.

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

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

5) For a review of GUTs and cosmology, see J. Ellis, M.K. Gaillard and D.V. Nanopoulos, *in* Unification of the fundamental particle interactions, eds. S. Ferrara, J. Ellis and P. Van Nieuwenhuizen (Plenum Press, New York, 1980), p. 461.

6) Ya.B. Zel'dovich, Phys. Lett. 59A, 254 (1976).

7) G. 't Hooft, Phys. Rev. Lett. 37, 8 (1976); Phys. Rev. D 14, 3432 (1976).

8) S. Weinberg, Phys. Rev. Lett. 43, 1566 (1979); Phys. Rev. D 22, 1694 (1980).

9) F.A. Wilczek and A. Zee, Phys. Rev. Lett. 43, 1571 (1979).
H.A. Weldon and A. Zee, Nucl. Phys. B173, 269 (1980).

10) J.C. Pati and A. Salam, Phys. Rev. Lett. 31, 661 (1973); Phys. Rev. D 8, 1240 (1973) and D 10, 275 (1974).

11) V.A. Kuzmin, Zh. Eksp. Teor. Fiz. 13, 335 (1970).

12) D.A. Ross, Nucl. Phys. B140, 1 (1978).

13) T. Goldman and D.A. Ross, Phys. Lett. 84B, 208 (1979).

14) W.J. Marciano, Phys. Rev. D 20, 274 (1980).

15) J. Ellis, M.K. Gaillard and D.V. Nanopoulos, Phys. Lett. 88B, 320 (1980).

16) C. Jarlskog, these Proceedings.

17) M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49, 652 (1973).

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

19) J. Ellis, M.K. Gaillard, D.V. Nanopoulos and S. Rudaz, Nucl. Phys. B176, 61 (1981).

20) J. Learned, F. Reines and A. Soni, Phys. Rev. Lett. 43, 907 (1979).
M.L. Cherry et al., Phys. Rev. Lett. 47, 1507 (1981).
M.R. Khrishnaswamy et al., Phys. Lett. 106B, 339 (1981).

21) D.H. Perkins, these Proceedings.

22) G. Kane and G. Karl, Phys. Rev. D 22, 1808 (1980).

23) For a recent review, see R.E. Marshak, Virginia Polytechnic Institute preprint VPI-HEP-81/3 (1981).

24) M. Baldo-Ceolin et al., ILL-Grenoble research proposal 03-05-027 (1980).
G.R. Young et al., Oak Ridge proposal (1980).
NADIR proposal, Pavia (1981).
R. Ellis et al., LAMPF proposal 647 (1981).

25) R.E. Marshak, R.N. Mohapatra and Riazuddin, Virginia Polytechnic Institute preprint VPI-HEP-80/7 (1980).

26) R.N. Mohapatra and R.E. Marshak, Phys. Lett. 94B, 183 (1980).
K.G. Chetyrkin, M.V. Kazarnovsky, V.A. Kuzmin and M.E. Shaposhnikov, Phys. Lett. 99B, 358 (1981).
Riazuddin, Virginia Polytechnic Institute preprint VPI-HEP-81/4 (1981).

27) K.E. Bergkvist, Nucl. Phys. B39, 317 (1972).

28) M. Daum et al., Phys. Lett. 74B, 126 (1978).

29) J. Kirkby, Proc. Int. Symp. on Lepton and Photon Interactions at High Energies, Batavia, 1979, eds. T.B.W. Kirk and H.D.I. Abarbanel (FNAL, Batavia, 1980), p. 107.

30) V.A. Lyubimov et al., Yad. Fiz. 32, 30 (1980) and Phys. Lett. 94B, 266 (1980).

31) R. Barbieri, J. Ellis and M.K. Gaillard, Phys. Lett. 90B, 249 (1980).

32) M. Gell-Mann, P. Ramond and R. Slansky, unpublished (1979).
R. Slansky, Talk presented at the Sanibel Symposium, CalTech preprint CALT-68-709 (1979).
T. Yanagita, Proc. Workshop on the Unified Theory and the Baryon Number in the Universe (KEK, Japan, 1979).

33) E. Witten, Phys. Lett. 91B, 81 (1980).

34) B. Pontecorvo, Zh. Eksp. Teor. Fiz. 33, 549 (1957); 34, 247 (1958); 53, 1717 (1967).
Z. Maki, M. Nakagawa and S. Sakata, Prog. Theor. Phys. 28, 870 (1962).
S. Bilenky and B. Pontecorvo, Phys. Rep. 41C, 225 (1978).

35) B. Kayser, Phys. Rev. D 24, 110 (1981).

36) G. Steigman, *in* Unification of the fundamental particle interactions, eds. S. Ferrara, J. Ellis and P. Van Nieuwenhuizen (Plenum Press, New York, 1980), p. 495.

37) See, for example, S. Weinberg, Gravitation and cosmology (Wiley, New York, 1972).

38) S. Hawking and G.F.R. Ellis, The large scale structure of space time (Cambridge University Press, Cambridge, 1973).

39) A.A. Penzias and R.R. Wilson, Ap. J. 142, 419 (1965).

40) G. Steigman, Ann. Rev. Astron. and Astrophys. 14, 339 (1976).

41) R.L. Golden et al., Phys. Rev. Lett. 43, 1196 (1979).

42) See, however, F. Stecker, Nature 273, 493 (1978).

43) A.D. Sakharov, Pis'ma Zh. Eksp. Teor. Fiz. 5, 32 (1967).

44) A.Yu. Ignatiev, N.V. Krasnikov, V.A. Kuzmin and A.N. Tavkhelidze, Phys. Lett. 76B, 436 (1978).
M. Yoshimura, Phys. Rev. Lett. 41, 381 (1978); E42, 746 (1979).
S. Dimopoulos and L. Susskind, Phys. Rev. D 18, 4500 (1978).
D. Toussaint, S.B. Treiman, F. Wilczek and A. Zee, Phys. Rev. D 19, 1036 (1979).
J. Ellis, M.K. Gaillard and D.V. Nanopoulos, Phys. Lett. 80B, 360 (1979); E82B, 464 (1979).
S. Weinberg, Phys. Rev. Lett. 42, 850 (1979).
A.D. Sakharov, Zh. Eksp. Teor. Fiz. 76, 1172 (1979).
S. Dimopoulos and L. Susskind, Phys. Lett. 81B, 416 (1979).
M. Yoshimura, Phys. Lett. 88B, 294 (1979).
D.V. Nanopoulos and S. Weinberg, Phys. Rev. D 20, 2484 (1979).

45) E.W. Kolb and S. Wolfram, Phys. Lett. 91B, 217 (1980) and Nucl. Phys. B172, 244 (1980); see also
J.N. Fry, K.A. Olive and M.S. Turner, Phys. Rev. D 22, 2953 and 2977 (1980).

46) J. Ellis, M.K. Gaillard and D.V. Nanopoulos, ref. 44.

47) J. Ellis, M.K. Gaillard, D.V. Nanopoulos and S. Rudaz, Phys. Lett. 99B, 101 (1981).

48) G. Steigman, Bartol Research Foundation preprint BA-81-20 (1981) and references therein.

49) D.N. Schramm and E.M.B. Symbalisty, Reps. Prog. Phys. (in press) (1981).

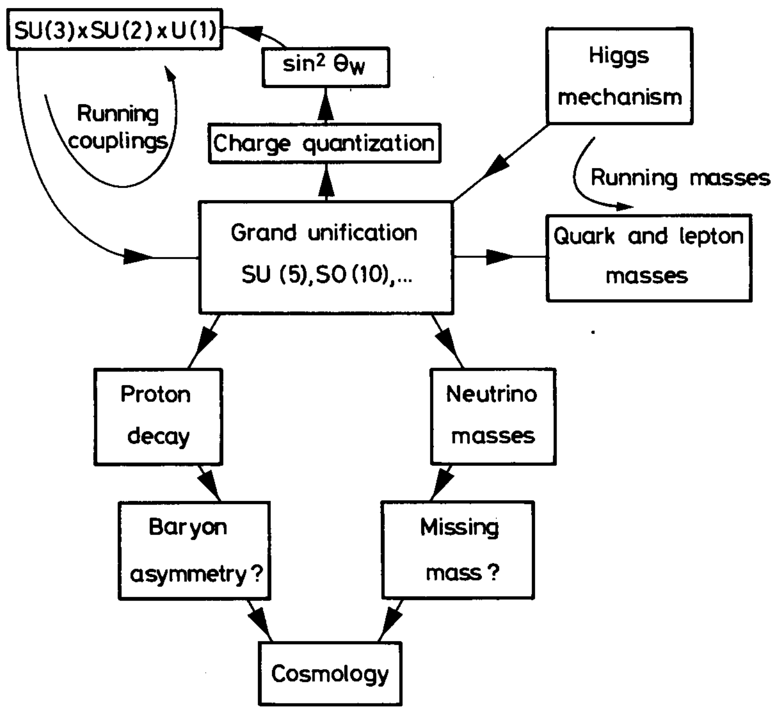

Fig. 1 Sketch of the implications and ramifications of GUTs and their connections with cosmology

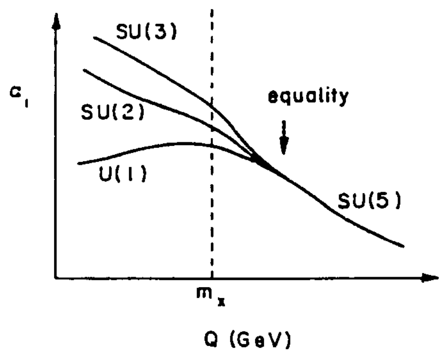

Fig. 2 In a momentum space renormalization scheme the coupling constants in a GUT do not come together exactly at $Q = m_X$

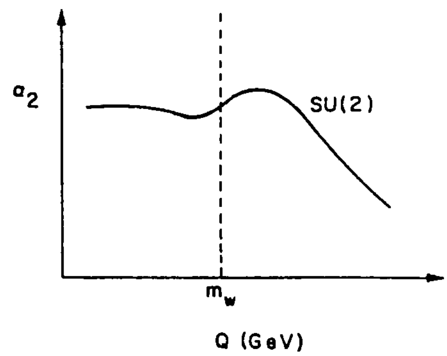

Fig. 3 The SU(2) coupling constant in momentum space evolves in a non-trivial manner close to the W threshold

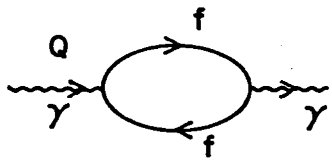

Fig. 4 Dominant contribution to the renormalization of the effective value of α between $Q = 0$ and $Q = O(m_W)$

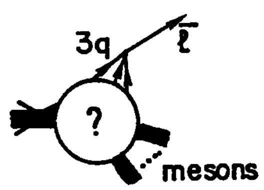

Fig. 5 The strong interaction problem in baryon decay

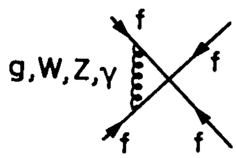

Fig. 6 The effective four-fermion operator responsible for baryon decay gets renormalized by g, W, Z, and γ exchanges at short distances $1/m_X < \Delta x < 1/1$ GeV

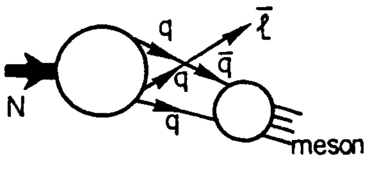

Fig. 7 The usual model for baryon decay into an anti-lepton and a meson

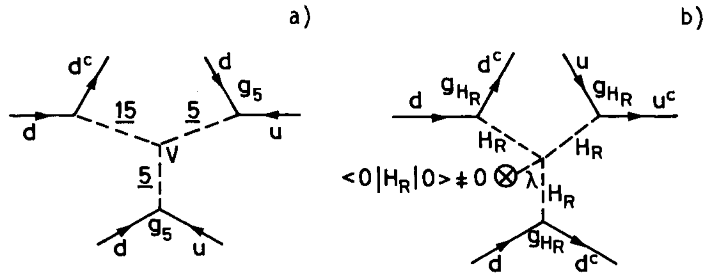

Fig. 8 Diagrams which could yield n-$\bar{n}$ oscillations (a) in a non-minimal SU(5) model, and (b) in an SO(10) model[25]

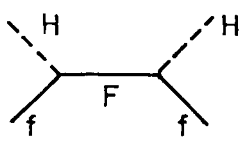

Fig. 9 Diagram which may contribute to an effective $(Hf)^2$ interaction (2.6)

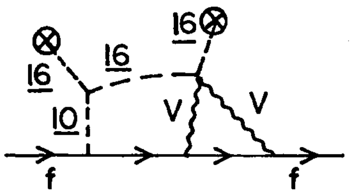

Fig. 10 Diagram generating an effective 126 Higgs fields in the minimal SO(10) model[33]

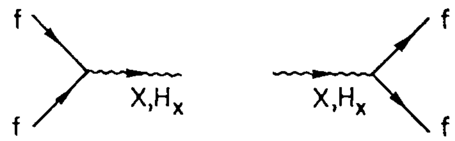

Fig. 11 Decay and inverse decay $1 \leftrightarrow 2$ particle reactions

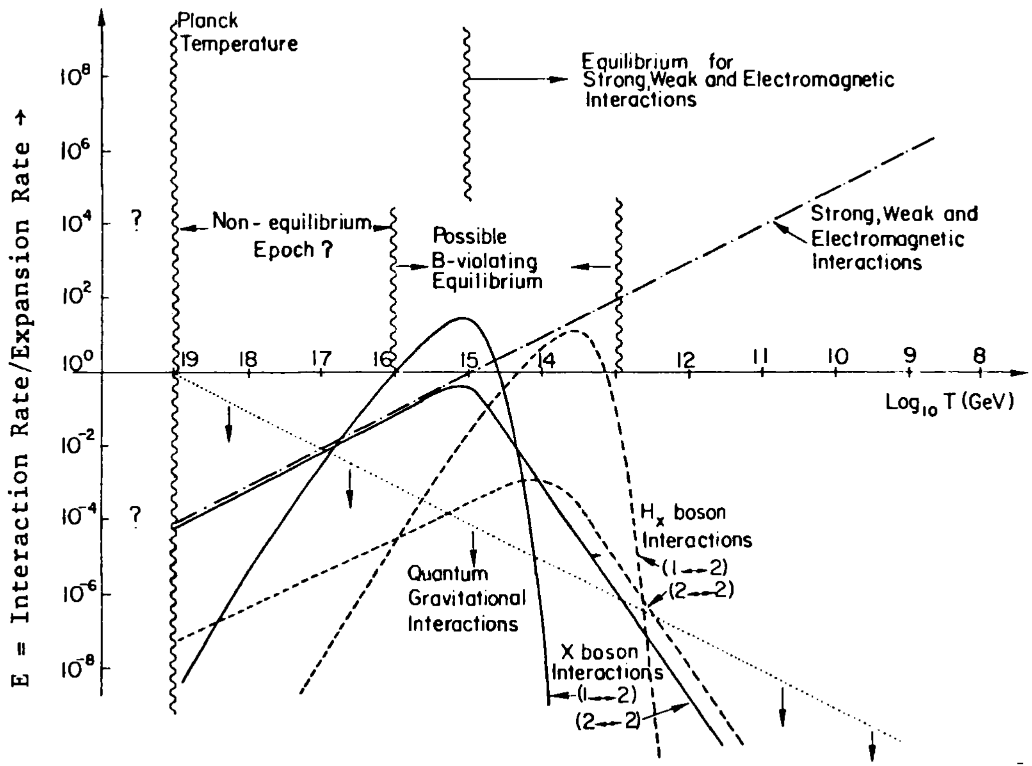

Fig. 12 Qualitative features[5] of the equilibrium ratios E ≡ Interaction Rate/Expansion Rate for different interactions in the early Universe. The different curves are explained in the text.

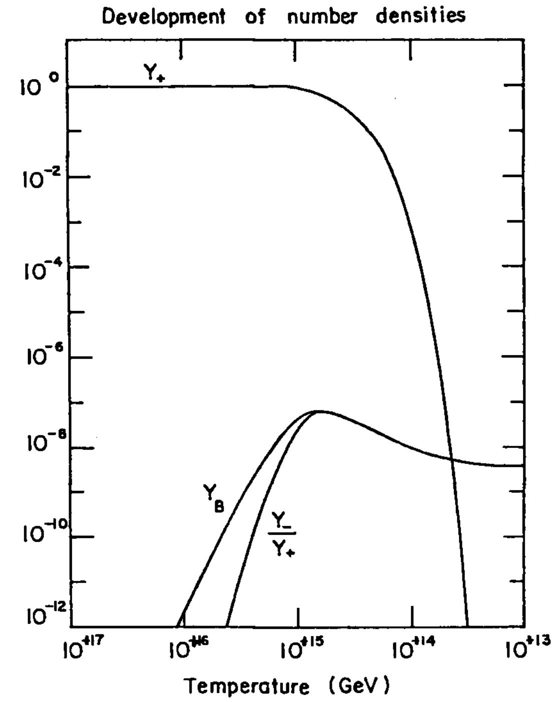

Fig. 13 Results of a typical calculation showing how a net quark asymmetry can be built up through heavy-particle interactions in the early Universe

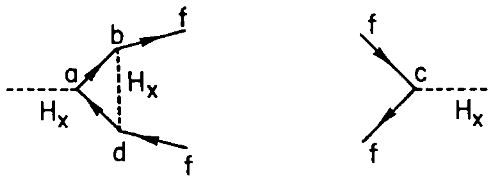

Fig. 14 Example of a fourth-order interference which may contribute to the C- and CP-violating decays of heavy bosons

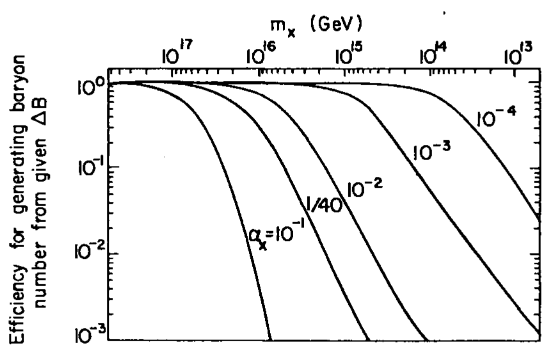

Fig. 15 Dependence of the baryon asymmetry generated for different values of the decaying superheavy boson and its coupling strength α_X