PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Unification and Supersymmetry

J. Ellis

Phil. Trans. R. Soc. Lond. A 1983 **310**, 279-292
doi: 10.1098/rsta.1983.0090

Phil. Trans. R. Soc. Lond. A **310**, 279–292 (1983)
Printed in Great Britain

Unification and supersymmetry

By J. Ellis†
CERN, Theory Division, CH-1211 Geneva 23, Switzerland

Current attempts to construct unified theories of fundamental particles and their interactions are described, with emphasis on their ability to understand the values of the fundamental constants. Examples include grand unified theories, which enable one to estimate the fine structure constant, the neutral weak interaction mixing parameter and certain quark masses. Finally, a review will be presented of the prospects offered by supersymmetry for understanding the scale of the weak interactions and for an eventual unification with gravity.

Introduction

My plan in this talk is to lead you further down the primrose path of unification to which you were introduced by Llewellyn Smith (1983) and by Weinberg (1983). Starting from the basic ideas of grand unification which you have already met, we will go on to more recent approaches to unification such as technicolour (Farhi & Susskind 1981), supersymmetry (Fayet & Ferrara 1977) and supergravity (van Nieuwenhuizen 1981). The emphasis throughout will be on ideas for understanding the values of the apparently 'fundamental' constants.

In § 1 we will count the parameters of the standard $SU(3) \times SU(2) \times U(1)$ model and find that there are at least 20 'fundamental' constants to be explained. The philosophy of conventional grand unification (Georgi & Glashow 1974; Georgi *et al.* 1974) is reviewed in § 2, and we find (Ellis & Nanopoulos 1981) that it is consistent only if the fine structure constant lies in the range $\frac{1}{170} < \alpha < \frac{1}{120}$. Section 3 introduces simple models for grand unification and we find that they enable us to calculate successfully the ratios of some of the standard model parameters. For example, charge quantization $|Q_e/Q_p| = 1$ is explained and the neutral weak mixing angle θ_W is calculable (Georgi *et al.* 1974), as well as some quark masses which are related to charged lepton masses (Chanowitz *et al.* 1977; Buras *et al.* 1978). These sections are brief, since many reviews of classical grand unification exist and the topics have already been touched on at this meeting (Llewellyn Smith 1983; Weinberg 1983). In § 4 we will assess critically the significance for grand unification of the recent negative results (Bionta *et al.* 1983; Goldhaber 1983) of a search for baryon decay. Recent calculations (Brodsky *et al* 1983; Isgur & Llewellyn Smith 1983) relating the baryon decay rate to the short-distance behaviour of the proton form factor suggest that the baryon decay amplitude may be rather larger (for a given value of m_X) than previous SU(6) estimates (Isgur & Wise 1982) had suggested, which would deepen the apparent conflict between experiment and conventional minimal grand unified theories (GUTs). However, even before the negative results of this search, theorists had grown dissatisfied with minimal GUTs with their 21 'fundamental' parameters, which do not represent a significant decrease from the 20 'fundamental' parameters of the standard $SU(3) \times SU(2) \times U(1)$ model! Establishing and maintaining a small weak interaction scale (m_W/M_P of order 10^{-17}) is a severe difficulty for conventional GUTs

† On leave of absence at SLAC, P.O. Box 4349, Stanford, California 94305, U.S.A.

with elementary scalar Higgs fields. Attempts to understand the weak interaction scale (the 'hierarchy problem') (Gildener & Weinberg 1976; Gildener 1976) are described in § 5. They include dynamical symmetry breaking, called here technicolour (Farhi & Susskind 1981), which renders m_W calculable, but must be supplemented by epicycles if one is to get non-zero quark and lepton masses. This complication is the source of severe phenomenological difficulties (Dimopoulos & Ellis 1981) which have not yet been overcome. A currently favoured alternative strategy for understanding the weak interaction scale is supersymmetry (Fayet & Ferrara 1977). It seems that one needs local supersymmetry (supergravity) for building realistic models (Alvarez-Gaumé *et al.* 1983; Ibáñez & López 1983; Ellis *et al.* 1983*a*) So far, these models tend to have m_W of order the gravitino mass, although there are models (Ellis *et al.* 1983*a*) in which the weak interaction scale is fixed dynamically by an analogue of dimensional transmutation (Coleman & Weinberg 1973). Finally, in § 6 we mention some possible strategies (Ellis *et al.* 1979; Ellis *et al.* 1983*a*) for understanding the grand unification scale m_X and the magnitude of the gauge coupling at energies of order m_X. Ultimately one may hope (Ellis *et al.* 1980*c*) that all the 'fundamental' constants may be calculable using an underlying $N = 8$ supergravity theory, but so far this is a dream beyond our dynamical understanding.

1. The parameters of the standard model

As Llewellyn Smith (1983) has told you, the standard model of strong, weak and electromagnetic interactions is based on the gauge group $SU(3) \times SU(2) \times U(1)$ and contains three generations of quarks and leptons (u, d, e, ν_e), (c, s, μ, ν_μ) and (t, b, τ, ν_τ). We have three independent gauge couplings g_3, g_2 and g_1 for the three factors of the gauge group. The fine structure constant

$$\alpha = (g_2^2/4\pi)\sin^2\theta_W, \tag{1}$$

where θ_W describes mixing between the neutral SU(2) and U(1) currents:

$$\sin^2\theta_W = \tfrac{3}{5}g_1^2/(g_2^2+\tfrac{3}{5}g_1^2). \tag{2}$$

You have heard that QCD with massless quarks has no free parameters: this is true as long as there is no external scale on which to measure g_3. We will take the Planck mass

$$M_P \equiv G_n^{-\frac{1}{2}} \approx 1.2\times 10^{19}\,\text{GeV}, \tag{3}$$

as our fundamental physical scale. It is then meaningful to ask what is the value of g_3 at a specified energy scale, say $E = 10^{-18} M_P$. The gauge interactions of $SU(3) \times SU(2) \times U(1)$ also require for their specification two non-perturbative CP-violating vacuum parameters θ_3, θ_2. The QCD angle θ_3 is in principle observable via the neutron electric dipole moment

$$d_n \approx 3\times 10^{-16}\,\theta_3\, e\,\text{cm}, \tag{4}$$

and from the present experimental upper limit (Altarev *et al.* 1981) on d_n we know that

$$\theta_3 < 2\times 10^{-9}. \tag{5}$$

Understanding the smallness of θ_3 is a major theoretical puzzle. The SU(2) vacuum angle θ_2 is practically unobservable because non-perturbative weak interactions are negligible: in principle θ_2 could be $O(1)$, though this seems highly implausible. Leaving the gauge sector we encounter six quark masses and three lepton masses as 'fundamental' parameters. In the standard model these are derived from underlying Yukawa interactions along with the three charged weak interaction mixing angles which generalize the 4-quark Cabibbo mixing angle to the case of six quarks, and a single CP-violating phase δ (Kobayashi & Maskawa 1973) which is blamed for the observed CP-violation in K^0 decays. Finally, in the standard model there are two weak boson masses to be specified: $m_{W\pm}$ and the Higgs boson mass m_H. The strength of the Fermi weak interaction is derived from the 'fundamental' parameters g_2 and m_W:

$$G_F/2^{\frac{1}{2}} = g_2^2/8m_W^2 + (\text{radiative corrections}). \tag{6}$$

Table 1. Parameter counts

standard SU(3) × SU(2) × U(1) model	minimal SU(5) GUT	type
3: g_3, g_2, g_1	1: g_5	gauge couplings
2: θ_3, θ_2	1: θ_5	vacuum angles
6	6	quark masses
3	0	lepton masses
3: θ_i	3: θ_i	charged weak mixing angles
1: δ	3: δ_j	CP-violating phases
2: m_W, m_H	7	boson masses
20	21	total

The above parameters ('fundamental' constants) are listed in table 1: there are a total of 20 of them. However, even this list is incomplete as we have assumed the standard electroweak quantum numbers for quarks, leptons and Higgs fields. We do not know why left-handed fermions should sit in doublets of SU(2) while right-handed fermions should be singlets of SU(2). Another egregious mystery is the choice of weak U(1) hypercharges which have arranged themselves so as to respect charge quantization:

$$|Q_e/Q_p| = 1 + O(10^{-20}). \tag{7}$$

There is no explanation for this quantization within the standard model.

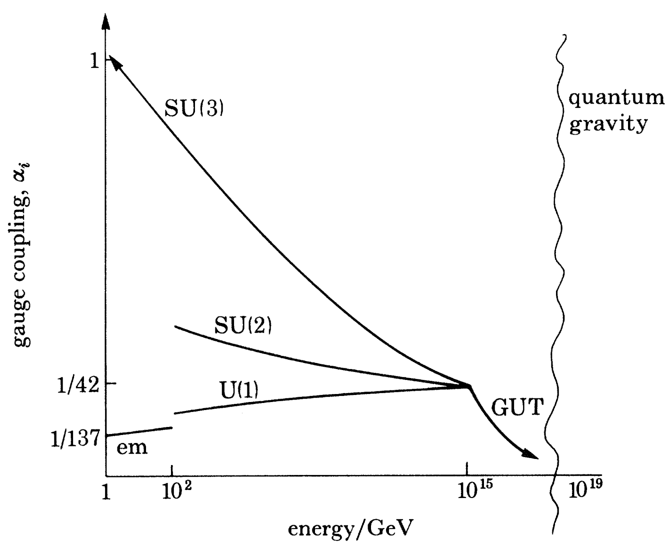

Figure 1. The approach of the SU(3) × SU(2) × U(1) gauge couplings to be unified at a scale $m_X = O(10^{15})$ GeV well below the Planck mass of 10^{19} GeV at which quantum gravity effects become important.

2. The philosophy of grand unification

This is indicated in figure 1 and has already been explained to you by Llewellyn Smith (1983). Given the absurd assumption that there is a grand desert with no prior oases of new particle or interaction thresholds, asymptotic freedom drives the strong SU(3) coupling down to meet the SU(2) and U(1) weak couplings at an energy scale of order 10^{15} GeV, to be identified with the masses of superheavy gauge vector bosons m_X. This unification scale is astronomically high because of the logarithmically slow evolution of the gauge couplings:

$$m_X = m_B \exp\,(O(1)/\alpha). \tag{8}$$

It is important for the consistency of the whole GUT philosophy that m_X be less than 10^{19} GeV, so that the neglect of gravity is a reasonable first approximation, while m_X must be larger than about 10^{14} GeV if baryons are not to decay more rapidly than experiment allows. For m_X (equation (8)) to lie within this range, we must have (Ellis & Nanopoulos 1981)

$$\tfrac{1}{170} < \alpha < \tfrac{1}{120} \tag{9}$$

as a consistency condition for the GUT philosophy. Happily enough, $\alpha = \frac{1}{137}$ lies within the range (equation (9)), and we can go on to look at specific GUT models.

3. Simple models

We must look for simple non-Abelian groups of rank $R \geqslant 4$ in order to be able to include the SU(3) × SU(2) × U(1) group of the standard model. Georgi & Glashow (1974) found that the only acceptable group of rank 4 was SU(5) which contains 24 gauge bosons. Half of these are the familiar photon, eight gluons, recently discovered $W^{\pm}$ (Arnison *et al.* 1983*a*; Banner *et al.* 1983) and the even more recent Z^0 (Arnison *et al.* 1983*b*). Then there are 12 superheavy gauge bosons X and Y which will mediate new hyperweak interactions violating baryon number, B, conservation. The gauge bosons mediate interactions between three generations of fermions, each of which contains 15 helicity states assigned unaesthetically to a reducible $\underline{\bar{5}}+\underline{10}$ representation of SU(5). The lightest $\underline{\bar{5}}$ representation is

$$\underline{\bar{5}} = \begin{pmatrix} \bar{d}_R \\ \bar{d}_Y \\ \bar{d}_B \\ e^- \\ \nu_e \end{pmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\}\text{strong SU(3)} \\ \left.\begin{matrix} \\ \\ \end{matrix}\right\}\text{weak SU(2)} \end{matrix} \Bigg\}\ \text{X, Y hyperweak interactions,} \tag{10}$$

where we have indicated explicitly the strong interactions acting on the first three entries, the SU(2) weak interactions acting on the last two, and the hyperweak interactions mediated by X and Y bosons coupling together the first three and the last two indices. We see that the X and Y couple quarks to leptons, as well as quarks to antiquarks in the $\underline{10}$ representation not shown, and hence engender B violation and baryon decay. The simplest SU(5) model requires in addition two multiplets of Higgs fields, a $\underline{24}\ \phi$ with a vacuum expectation value $O(10^{15})$ GeV to break SU(5) to SU(3) × SU(2) × U(1), and a $\underline{\bar{5}}$ H with a vacuum expectation value $O(10^2)$ GeV to break SU(2) × U(1) to $U(1)_{em}$ and generate m_W, m_q and m_l.

Since there are a denumerable infinity of simple gauge groups, there are as many GUT models.

The next smallest one after SU(5) is SO(10) (Georgi 1975; Fritzsch & Minkowski 1975) which contains 45 gauge bosons, thereby offering baryons more ways to decay, three $\underline{16}$s of fermions, including a candidate for a right-handed neutrino, and at least three irreducible representations $\underline{16}+\underline{45}+\underline{10}$ of Higgses. Since this model introduces no fundamental new principles, we will concentrate on SU(5) as a bellwether GUT.

All GUTs predict charge quantization because they embed the U(1) of electromagnetism in a simple group, which means that charges are related by Clebsch–Gordan coefficients. The sum of the charges in every GUT representation must vanish, for example in the $\underline{\bar{5}}$ of SU(5) we find from (10) that

$$\left.\begin{aligned} Q_e = -1,\ 3Q_{\bar{d}}+Q_e = 0 \\ \Rightarrow Q_d = -\tfrac{1}{3} \Rightarrow Q_u = +\tfrac{2}{3} \Rightarrow \\ Q_p = 2Q_u+Q_d = +1, \end{aligned}\right\} \tag{11}$$

in accord with the experimental constraint (7).

GUTs also predict the 'fundamental' parameter $\sin^2\theta_W$ (equation (2)) which is $\frac{3}{8}$ in the GUT symmetry limit $g_2 = g_1$, but gets renormalized in simple models as indicated in figure 1 to (Georgi *et al.* 1974; Marciano & Sirlin 1981; Llewellyn Smith *et al.* 1981)

$$\sin^2\theta_W = 0.216 \pm 0.002, \tag{12}$$

to be compared (successfully) with the experimental value

$$\sin^2\theta_W = 0.215 \pm 0.012, \tag{13}$$

as discussed by Llewellyn Smith (1983).

Another successful prediction of a 'fundamental' parameter which he did not mention is that of the b quark mass in terms of the τ lepton mass (Chanowitz *et al.* 1977). Generally, quark and lepton masses are related by Clebsch–Gordan coefficients in GUTs, and in minimal SU(5) we have $m_b = m_\tau$ in the GUT symmetry limit. This gets renormalized analogously to $\sin^2\theta_W$, resulting in the physical prediction (Buras *et al.* 1978):

$$m_\tau = 1.78\,\text{GeV} \Rightarrow m_b \approx 5\,\text{GeV}, \tag{14}$$

if there are only six quarks in three generations (Nanopoulos & Ross 1979). In all fairness, it should be confessed that there are analogous predictions for m_s and m_d which are controversial and wrong respectively, but these may be modified without doing violence to the successful prediction (equation (14)).

The phenomenological successes (equations (7), (12) and (14)) constitute the only practical reasons so far for believing in grand unification, apart from its possible aesthetic appeal.

4. Baryon number violating interactions

After all this foreplay, let us get down to the nitty-gritty of GUTs, namely the prediction of baryon decay. The strength of the new interactions

$$G_X/2^{\frac{1}{2}} = g_X^2/8m_X^2, \tag{15}$$

yields a $\Delta B \neq 0$ amplitude $O(m_X^{-2})$ and hence a decay rate $O(m_X^{-4})$ and a lifetime

$$\tau_{p,n} = (m_X^4/m_B^5) \times O(1). \tag{16}$$

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY
PHILOSOPHICAL TRANSACTIONS OF

We will return shortly to the estimation of the $O(1)$ coefficient in (16), for the moment we just emphasize the great sensitivity to m_X which is estimated to be (Goldman & Ross 1980; Ellis *et al.* 1980*b*; Llewellyn Smith *et al.* 1981)

$$m_X = (1\text{–}2) \times 10^{15} \times \Lambda_{\overline{MS}} \tag{17}$$

in minimal SU(5), where as discussed by Llewellyn Smith (1983) a favoured range of $\Lambda_{\overline{MS}}$ is

$$100\,\mathrm{MeV} < \Lambda_{\overline{MS}} < 200\,\mathrm{MeV}, \tag{18}$$

though it could be as large as 400 MeV. Using (16)–(18) conventional estimates (Llewellyn Smith 1983) yield

$$\tau_{p,n} = 10^{29\pm 2}\,\text{years} \tag{19}$$

The favoured decay modes in minimal SU(5) are

$$\left.\begin{array}{l} p \rightarrow e^+\pi^0,\ e^+\omega,\ e^+\rho^0 \quad \text{and} \quad \mu^+K^0, \\ n \rightarrow e^+\pi^- \quad \text{and} \quad e^+\rho^-. \end{array}\right\} \tag{20}$$

Until recently, the predictions (19) and (20) looked quite healthy, there being several lower limits on the baryon lifetime of order 2×10^{30} years, but two reports (Krishnaswamy *et al.* 1982; Battistoni *et al.* 1982) of candidates for baryon decay, including one possible $p \rightarrow \mu^+K^0$. However, as discussed at this meeting by Goldhaber (1983), the I.M.B. collaboration (Bionta *et al.* 1983) has recently established that

$$\tau(p \rightarrow e^+\pi^0) > 10^{32}\,\text{years}, \tag{21}$$

which is very embarrassing for conventional GUTs. They have one event compatible with $p \rightarrow \mu^+K^0$, but it could very well be a background neutrino interaction.

Does the result (21) rule out GUTs? There are many possible baryon decay modes that the I.M.B. collaboration has not yet searched for, and it clearly does not exclude models which do not predict the decay mode $p \rightarrow e^+\pi^0$, but the minimal conventional SU(5) described in § 3 looks rather sick. We (Brodsky *et al.* 1983) have recently re-evaluated the baryon decay rate to be expected for a given value of m_X, i.e. the $O(1)$ coefficient in expression (16), in an attempt to answer the question at the head of this paragraph. We related the short distance baryon decay amplitude to knowledge about baryon wave functions at short distances gleaned from the proton magnetic form factor at large momentum transfers and from $J/\psi \rightarrow p\bar{p}$ decay. We found a much larger baryon decay rate than previous non-relativistic SU(6) and bag model calculations. This discrepancy may mean that the three-quark wavefunction overlap at short distance which controls the baryon decay rate in the chiral limit cannot be related easily to the one- and two-quark wavefunctions which may be known reliably from non-relativistic SU(6). Alternatively, it may mean that the proton form factor at $Q^2 = O(10)\,\mathrm{GeV}^2$ is not dominated by the short-distance three-quark baryon wave-function, which has accordingly been grossly over-estimated in the past (Isgur & Llewellyn Smith 1983, personal communication). This will undoubtedly become a controversy among practitioners of QCD. If we accept at face value our normalization of the baryon decay rate using form factor data, we infer from the limit (21) that

$$G_X < O(10^{-32})\,\mathrm{GeV}^{-2} \Rightarrow m_X > 2 \times 10^{15}\,\mathrm{GeV}, \tag{22}$$

corresponding to $\Lambda_{\overline{MS}} > 1\,\mathrm{GeV}$ according to the minimal SU(5) GUT relation (17). This makes the minimal SU(5) GUT look very sick indeed, even if it is 'not dead yet'.

Before abandoning this simple model, it is salutary to record (see table 1) how many 'funda-

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

mental' constants it contains. The gauge sector is certainly simpler than in the standard model: 1 gauge coupling g_5 instead of 3, and 1 non-perturbative vacuum parameter θ_5 instead of 2. There are still six quark masses, but now the three charged lepton masses are no longer independent of them. There are still three charged weak mixing angles, but now there are three CP-violating phases instead of one. The two new phases only appear in X and Y boson interactions and could play a role in cosmological baryosynthesis. There are now many more parameters needed to specify the boson masses, namely seven parameters in the minimal SU(5) Higgs potential. This model therefore has a total of 21 'fundamental' constants, which is not a significant improvement on the 20 of the standard model! Furthermore, the model predicts the wrong value of m_d, does not generate enough baryons in the early Universe, and probably predicts too short a baryon lifetime.

In fact, even before this latest experimental setback, the smart money had already moved out of stocks in minimal SU(5), as we see in the next section.

5. Attempts to understand the weak interaction scale

Since m_W at 80 GeV (Arnison *et al.* 1983*a*; Banner *et al.* 1983) is so much larger than the mass of any other known 'elementary' particle, it may seem at first sight strange to ask why m_W is so small. But m_W *is* very small on the scales of gravitation or of grand unification:

$$m_W/M_p = O(10^{-17}), \quad m_W/m_X \leqslant O(10^{-13})? \tag{23}$$

In spontaneously broken gauge theories m_W must be of the same order as the light Higgs boson m_H, which is awkward since the Higgs mass is notoriously unstable. We find

$$\delta m_H^2 = O(M_P^2), \tag{24}$$

from propagation through space–time foam at the Planck scale (Hawking *et al.* 1980), while couplings between the light Higgs H and the heavy Higgs ϕ in GUTs give

$$\delta m_X^2 = O(m_X^2), \tag{25}$$

from propagation through the GUT vacuum. Even if we set these to zero (how? why?), the Higgs mass is still destabilized by radiative corrections:

$$\delta m_H^2 = O(\alpha^n)\,(M_P^2 \text{ or } m_X^2). \tag{26}$$

This is the so called 'hierarchy problem' (Gildener & Weinberg 1976; Gildener 1976): we must understand what symmetry protects the light W and H from feedthroughs from the large mass scales. We now discuss two alternative strategies for such a 'solution' of the hierarchy problem. One may dissolve the offending diagrams by making H composite on a distance scale $x = O(1/1\,\text{TeV})$ and invoking dynamical symmetry breaking as in technicolour models. Alternatively one may cancel boson and fermion loops against each other as in supersymmetric theories.

(*a*) *Technicolour*

We postulate (Weinberg 1976, 1979; Susskind 1979) a complete new set of gauge interactions which become strong and confine unseen technifermions on a new distance scale $O(1/1\,\text{TeV})$. The previously elementary Higgs H is now replaced by a composite spinless techni-pion π_T which is a techni-fermion $\bar{F}F$ bound state, analogous to the conventional π which is a $(\bar{q}q)$ bound state:

$$H \to \pi_T = (\bar{F}F) \leftrightarrow \pi = (\bar{q}q). \tag{27}$$

The Higgs vacuum expectation value is replaced by a vacuum condensate as in QCD:

$$\langle 0|\,\mathrm{H}\,|0\rangle \rightarrow \langle 0|\,\overline{\mathrm{F}}\mathrm{F}\,|0\rangle \leftrightarrow \langle 0|\,\bar{\mathrm{q}}\mathrm{q}\,|0\rangle, \tag{28}$$

which breaks weak gauge symmetry spontaneously:

$$m_{\mathrm{W}} = g_2 \times O(\langle 0|\,\overline{\mathrm{F}}\mathrm{F}\,|0\rangle)^{\frac{1}{3}}, \tag{29}$$

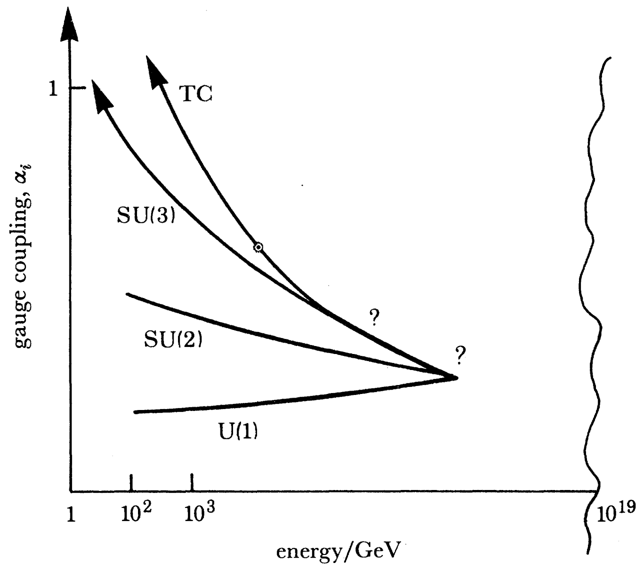

FIGURE 2. A sketch of technicolour (TC) interactions which get strong at a scale $O(1)$ TeV and may be unified with the other interactions at higher energies.

thanks to massless techni-pions being eaten by the $\mathrm{W}^{\pm}$ and Z^0. One can imagine that the technicolour interaction is unified with the others at some high energy scale as seen in figure 2, in terms of which the scale of $\langle 0|\,\mathrm{F}\overline{\mathrm{F}}\,|0\rangle$ and hence m_{W} (equation (29)) is determined dynamically. Thus m_{W} and m_{H} would no longer be 'fundamental' constants.

This is a very economical scenario for generating m_{W} and m_{Z}, but to obtain non-zero masses for quarks and leptons we must add epicycles to this elegant theory. We need new extended technicolour interactions (Dimopoulos & Susskind 1979; Eichten & Lane 1980) mediated by new heavy gauge bosons E:

$$m_{\mathrm{q,\,l}} = \langle 0|\,\overline{\mathrm{F}}\mathrm{F}\,|0\rangle / m_E^2. \tag{30}$$

These additional interactions cause problems, since there are related gauge boson exchanges which mediate flavour-changing interactions at levels far above experimental upper limits (Dimopoulos & Ellis 1981). Moreover, realistic theories contain many uneaten techni-pions that acquire calculable masses from conventional strong, weak and electromagnetic interactions, but not from extended technicolour interactions (Binétruy *et al.* 1981, 1982). None of these have been seen by experiment, which is disastrous in the case of the colourless charged techni-pions $\mathrm{P}^{\pm}$ whose masses were calculated (Dimopoulos 1980; Chadha & Peskin 1981 *a, b*) to be

$$m_{\mathrm{P}^{\pm}} \leqslant 15\,\mathrm{GeV}. \tag{31}$$

This pair of disasters, coupled with the unattractive nature of the complicated extended technicolour interactions, has recently led to a general abandonment of technicolour.

(b) *Supersymmetry*

This is a new type of symmetry (Wess & Zumino 1974) in which fermions are connected to bosons:

$$Q\,|\mathrm{F}\rangle = |\mathrm{B}\rangle, \quad Q\,|\mathrm{B}\rangle = |\mathrm{F}\rangle, \tag{32}$$

by spinorial charges Q_α. As might be expected for fermionic objects, these charges Q_α obey an anticommutation algebra:

$$\{Q^i_\alpha, Q^{+\dot{\alpha}}_j\} = -2(\sigma^\mu)^{\dot{\alpha}}_\alpha\,\delta^i_j\,P_\mu, \tag{33}$$

where P_μ is the energy–momentum operator. In writing (33) we have sneaked in a new internal index $i = 1, 2, \ldots, N$ to label several different supersymmetry charges Q^i_α. The case $N = 1$ is simple supersymmetry, while $N > 1$ theories are said to possess extended supersymmetry. How large can N be? Renormalizable gauge theories are restricted to helicities $|\lambda| \leqslant 1$. Therefore at most four changes of spin $\frac{1}{2}$ by supersymmetry charges Q^i_α are permitted:

$$\lambda = +1 \underset{Q}{\rightarrow} +\tfrac{1}{2} \underset{Q}{\rightarrow} 0 \underset{Q}{\rightarrow} -\tfrac{1}{2} \underset{Q}{\rightarrow} -1; \tag{34}$$

and hence $N \leqslant 4$ for gauge theories. Supergravity theories with $|\lambda| \leqslant 2$ for the graviton are allowed to have $N \leqslant 8$ (van Nieuwenhuizen 1981). In most of what follows we will restrict ourselves to simple $N = 1$ supersymmetric theories, though some speculations about $N = 8$ supergravity will be advanced. The basic supermultiplets of $N = 1$ theories are the following

$$\text{gauge: } \lambda = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}; \quad \text{chiral: } \lambda = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}; \tag{35}$$

together with the graviton–gravitino $(2, \frac{3}{2})$ supermultiplet.

Table 2. Supersymmetric particles

particle	spin	sparticle	spin
quark: q	$\frac{1}{2}$	squark: $\tilde{\mathrm{q}}$	0
lepton: l	$\frac{1}{2}$	slepton: $\tilde{\mathrm{l}}$	0
gluon: g	1	gluino: $\tilde{\mathrm{g}}$	$\frac{1}{2}$
photon: γ	1	photino: $\tilde{\gamma}$	$\frac{1}{2}$
$\mathrm{W}^\pm$	1	wino: $\tilde{\mathrm{W}}$	$\frac{1}{2}$
Z^0	1	zino: $\tilde{\mathrm{Z}}$	$\frac{1}{2}$
Higgs: H	0	shiggs: $\tilde{\mathrm{H}}$	$\frac{1}{2}$

Unfortunately, no known particle can be the supersymmetric partner of any other, so we must at least double the number of known particles by the addition of unseen partners as seen in table 2. All the charged particles must have masses large enough to have avoided production and detection in $\mathrm{e}^+\mathrm{e}^-$ collisions:

$$m_{\tilde{\mathrm{q}}}, m_{\tilde{\mathrm{l}}}, m_{\tilde{\mathrm{W}}\pm}, m_{\tilde{\mathrm{H}}\pm} \gtrsim O(17)\,\mathrm{GeV}. \tag{36}$$

The neutral ones could be rather lighter. For example, the best limit on the gluino mass from its absence in hadron–hadron collisions is (Bergsma *et al.* 1983)

$$m_{\tilde{\mathrm{g}}} \gtrsim O(2)\,\mathrm{GeV}, \tag{37}$$

whereas particle physics offers no lower bound on the mass of the photino $\tilde{\gamma}$.

How heavy could these supersymmetric particles be? An answer is provided by attempts to stabilize the gauge hierarchy. The correspondence (equation (32)) between bosons and fermions with identical couplings enforces systematic cancellations between loop diagrams which ensure that

$$\delta m_{\mathrm{H}}^2 = O(\alpha)\,|m_{\mathrm{B}}^2 - m_{\mathrm{F}}^2|. \tag{38}$$

This is acceptably small, i.e. $O(m_{\mathrm{W}}^2) = O(10^2\,\mathrm{GeV})^2$ if

$$|m_{\mathrm{B}}^2 - m_{\mathrm{F}}^2| \lesssim O(1)\,\mathrm{TeV}^2. \tag{39}$$

Thus the particles must be very light on the Planck scale, and potentially accessible to the next generation of particle accelerators. There are other phenomenological implications of supersymmetric theories beyond the existence of many new particles. For example, in minimal supersymmetric SU(5) GUTs (Dimopoulos & Georgi 1981; Sakai 1982) the prediction (12) for $\sin^2\theta_{\mathrm{W}}$ is modified (Ibáñez & Ross 1982; Einhorn & Jones 1982) to

$$\sin^2\theta_{\mathrm{W}} = 0.236 \pm 0.002, \tag{40}$$

which appears less successful. (Though the C.D.H.S. collaboration may soon announce a new, higher measurement of $\sin^2\theta_{\mathrm{W}}$.) A more dramatic modification is that of the GUT predictions (equations (19) and (20)) for baryon decay: in minimal supersymmetric GUTs (Dimopoulos *et al.* 1982; Ellis *et al.* 1982*d*)

$$\mathrm{p} \rightarrow \bar{\nu}\mathrm{K}^+, \quad \bar{\nu}\mathrm{K}^{*+}; \quad n \rightarrow \bar{\nu}\mathrm{K}^0, \quad \bar{\nu}\mathrm{K}^{*0}, \tag{41}$$

with a less certain estimate of the total lifetime, thanks to our ignorance of the spectrum of supersymmetric particles. There is no longer a *prima facie* conflict with the negative results of the I.M.B. experiment.

So far, supersymmetric theories offer no clear explanation of the origin of m_{W}, but only allow a small value of m_{W} to be stabilized against radiative corrections. The first attempts (Dimopoulos & Georgi 1981; Sakai 1982) to construct supersymmetric theories neglected gravity and only used global supersymmetry. Spontaneously broken versions of these models came to grief, either because they had anomalies and/or did not break supersymmetry and/or had weak interactions becoming strong at energies much less than M_{P} (the so-called D theories) (Farrar & Weinberg 1983) or else required baroque spectra of unseen particles with unaesthetic symmetries imposed just to break supersymmetry (the so called F theories) (Ellis *et al.* 1982*a*, *b*). Therefore, recent model-building has focused on $N = 1$ supergravity theories with local supersymmetry. These theories embody the super-Higgs effect (Cremmer *et al.* 1979, 1983) which offers a new mechanism of spontaneous supersymmetry breaking not available in globally supersymmetric theories. In general they have a breaking scale

$$m_{\tilde{\mathrm{q}}}, m_{\tilde{\mathrm{l}}}, m_{\tilde{\mathrm{W}}}, m_{\tilde{\mathrm{H}}} = O(m_{\mathrm{gravitino}}), \tag{42}$$

in the sparticle spectrum. If one neglects radiative corrections, the scale of weak gauge symmetry breaking is of the same magnitude

$$m_{\mathrm{W}} = O(m_{\mathrm{gravitino}}), \tag{43}$$

and this could well also be the case in theories where gauge symmetry breaking is induced by radiative corrections in the supergravity theory (Ellis *et al.* 1983*b*; Alvarez-Gaumé *et al.* 1983; Ibáñez & López 1983; Ellis *et al.* 1983*a*). However, it is also possible to construct models (Ellis *et al.* 1983*a*) with symmetry breaking by radiative corrections in which the weak gauge symmetry breaking scale is determined dynamically by dimensional transmutation, as described in more detail in the next section. In this case it is possible that

$$m_{\mathrm{W}} \gg m_{\mathrm{gravitino}}, \tag{44}$$

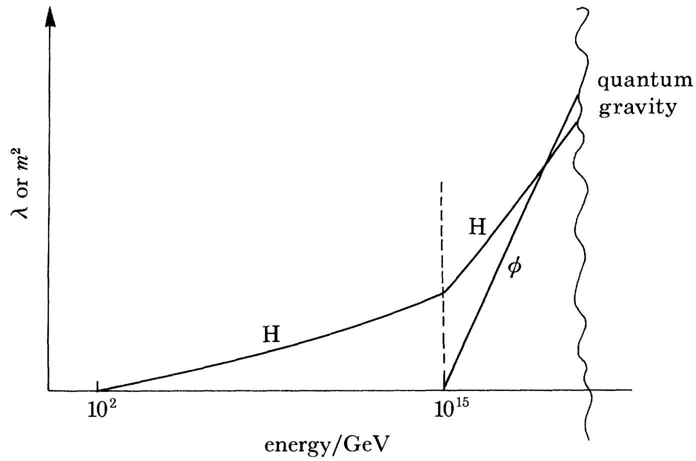

FIGURE 3. The coupling (mass) of a large Higgs representation ϕ evolves more rapidly than that of a small Higgs representation H. The grand unification symmetry may be broken when $\lambda_\phi(m_\phi)$ goes to zero, while weak SU(2) is broken when $\lambda_H(m_H)$ vanishes.

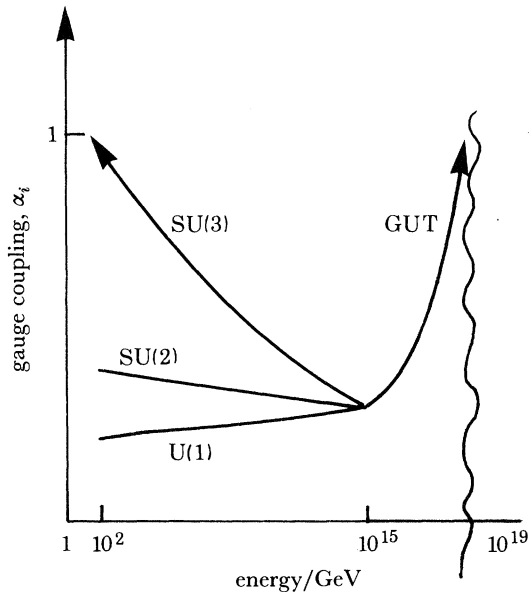

FIGURE 4. If there are enough particles with masses between 10^{15} and 10^{19} GeV, the strong coupling may decrease from $O(1)$ at M_P down to $O(\frac{1}{20}$ to $\frac{1}{40})$ at m_X.

though phenomenological considerations tell us that the supersymmetry breaking scale and hence $m_{\text{gravitino}}$ cannot be much less than 20 GeV. Models of weak gauge symmetry breaking by radiative corrections require the existence of at least one heavy fermion, and the most natural candidate would be the t quark. It would need to have a mass

$$m_t \gtrsim O(60)\text{ GeV}, \tag{45}$$

in these radiative scenarios.

It is an unfortunate feature of all these supergravity models that they require a light gravitino of mass $\lesssim O(m_W)$. There is as yet no clear idea how such a small mass parameter could emerge from a theory whose natural dynamical scale is $O(M_P)$. Thus supersymmetric theories have so far only lengthened, not shortened our list of 'fundamental' parameters.

6. Attempts to understand the grand unification scale

In the simple models of grand unification discussed up till now, including supersymmetric ones, there is a grand unification scale $m_X \ll M_P$. It is possible to push m_X up to M_P only at the expense of introducing large numbers of additional low-mass fields (oases in the desert). Assuming that m_X is significantly less than M_P, it is interesting to speculate why m_X is so near to M_P compared with m_W on a logarithmic scale, and yet so far: $m_X/M_P = O(10^{-4})$? Possible scenarios for understanding the value of m_X are provided by the idea of dimensional transmutation alluded to earlier. In its original form due to Coleman & Weinberg (1973), the spontaneous breaking of weak gauge symmetry in a theory with zero Higgs mass at the tree level occurred at a mass scale where the renormalization group equations drove to zero a dimensionless parameter, namely a quartic Higgs self-coupling λ. If one imagines the initial value of this parameter being determined to be $O(\alpha)$ by dynamics at the Planck scale, then one finds that

$$m_W = M_P \exp\left(-O(1)/\alpha\right), \tag{46}$$

which is reminiscent of (8). In supersymmetric theories λ is unrenormalized, but its role can be usurped (Ellis *et al.* 1983*a*) by a combination m of the supersymmetry breaking Higgs masses in the theory, with the result (44). These ideas can be extended (Ellis *et al.* 1979, 1983*a*) to GUTs with two Higgs representations ϕ and H, the former large ($\underline{24}$ of SU(5)?) and responsible for the initial GUT beaking at 10^{15} GeV, while the latter is smaller ($\underline{5}$ of SU(5)?) and responsible for weak symmetry breaking at 10^2 GeV. One can imagine specifying a theory with $\lambda_{24} = O(\lambda_5)$ ($m_{24} = O(m_5)$) at the Planck scale. As one comes down to lower mass scales, $\lambda_{24}(m_{24})$ will change more rapidly because of the larger Casimir coefficients associated with larger group representations and calculations (Ellis *et al.* 1979, 1983*a*) indicate that it could easily vanish at $O(10^{-4})\,M_P$ as required to fix m_X satisfactorily (see figure 3). Meanwhile, $\lambda_5(m_5)$ is non-zero and evolves even more slowly at lower mass-scales as indicated in figure 3, because the H representation is split in mass resulting in even smaller Casimir coefficients. Eventually $\lambda_5(m_5)$ will also vanish and thereby generate m_W, but because of the different group-theoretical coefficients it is very likely that

$$m_W/m_X \ll m_X/M_P = O(10^{-4}) \ll 1, \tag{47}$$

as desired.

It would also be nice to understand the gauge coupling α_X at the grand unification scale, which is about $\frac{1}{42}$ in minimal GUTs but about $\frac{1}{24}$ in supersymmetric GUTs. It is natural to suppose that $\alpha_X = O(1)$ at the Planck mass, which could facilitate full unification with gravity, possibly in an $N = 8$ supergravity theory as we shall speculate in a moment. If there are very many heavy particles with masses between m_X and M_P, their effect on the GUT renormalization group equations can be so large as to reverse the normal asymptotic freedom trend for α_X to increase with decreasing energy, and instead make α_X decrease as the energy scale decreases as shown in figure 4. If the contributions of these conjectured heavy particles in the renormalization group equations for α_X have about three times the magnitude of the gauge boson couplings driving asymptotic freedom, then α_X can (Ellis *et al.* 1982*c*) decrease from $O(1)$ at the Planck scale to $O(\frac{1}{20}$ to $\frac{1}{40})$ at the grand unification scale, as desired.

It is appropriate to speculate at the end of this talk about the possible eventual unification of all interactions at the Planck mass. The most natural candidate theory is supergravity, most probably the largest version with $N = 8$. We have already discussed the use of $N = 1$ supergravity, but this was as a phenomenological framework (Ellis *et al.* 1983*b*) for low energy physics at energy scales much less than M_P. If one takes a more fundamental approach and regards the $N = 8$ extended supergravity theory as an underlying theory of all elementary particle interactions (Ellis *et al.* 1980*a*; Ellis *et al.* 1980*c*), one encounters problems since all the known particles cannot be among the elementary states in the $N = 8$ graviton supermultiplet (Gell-Mann 1977). Perhaps some or all of the particles we know are not in fact 'elementary' but are actually composites of these $N = 8$ supergravity 'preons' (Cremmer & Julia 1979). While this is an attractive conjecture, our ignorance of supergravity dynamics does not yet allow us to put this idea on any sort of calculational basis, and there are even arguments that it may fail (Davis *et al.* 1983). If such a strategy could be made to work, it would offer the prospect of a theory with 0 (or perhaps 1) free parameter. This would at last be a significant reduction in the number of 'fundamental' constants.

References

Altarev, I. S. *et al.* 1981 *Physics Lett.* B **102**, 13.
Alvarez-Gaumé, L., Polchinski, J. & Wise, M. B. 1983 Harvard preprint HUTP-82/A063.
Arnison, G. *et al.* 1983*a* *Physics Lett.* B **122**, 103.
Arnison, G. *et al.* 1983*b* *Physics Lett.* B **126**, 398.
Banner, M. *et al.* 1983 *Physics Lett.* B **122**, 476.
Battistoni, G. *et al.* 1982 *Physics Lett.* B **118**, 461.
Bergsma, F. *et al.* 1983 *Physics Lett.* B **121**, 429.
Binétruy, P., Chadha, S. & Sikivie, P. 1981 *Physics Lett.* B **107**, 425.
Binétruy, P., Chadha, S. & Sikivie, P. 1982 *Nucl. Phys.* B **207**, 505.
Bionta, R. M. *et al.* 1983 *Phys. Rev. Lett.* **91**, 27.
Brodsky, S. J., Ellis, J., Hagelin, J. S. & Sachrajda, C. T. 1983 SLAC-PUB-3141.
Buras, A. J., Ellis, J., Gaillard, M. K. & Nanopoulos, D. V. 1978 *Nucl. Phys.* B **135**, 66.
Chadha, S. & Peskin, M. E. 1981*a* *Nucl. Phys.* B **185**, 61.
Chadha, S. & Peskin, M. E. 1981*b* *Nucl. Phys.* B **187**, 541.
Chanowitz, M. S., Ellis, J. & Gaillard, M. K. 1977 *Nucl. Phys.* B **128**, 506.
Coleman, S. & Weinberg, E. 1973 *Phys. Rev.* D **7**, 1888.
Cremmer, E., Ferrara, S., Girardello, L. & van Proeyen, A. 1983 *Nucl. Phys.* B **212**, 413.
Cremmer, E. & Julia, B. 1979 *Nucl. Phys.* B **159**, 141.
Cremmer, E., Julia, B., Scherk, J., Ferrara, S., Girardello, L. & van Nieuwenhuizen, P. 1979 *Nucl. Phys.* B **147**, 105.
Davis, A. C., Macfarlane, A. J. & van Holten, J. W. 1983 *Physics Lett.* B **125**, 151.
Dimopoulos, S. 1980 *Nucl. Phys.* B **168**, 69.
Dimopoulos, S. & Ellis, J. 1981 *Nucl. Phys.* B **182**, 505.
Dimopoulos, S. & Georgi, H. 1981 *Nucl. Phys.* B **193**, 150.
Dimopoulos, S., Raby, S. & Wilczek, F. 1982 *Physics Lett.* B **112**, 133.
Dimopoulos, S. & Susskind, L. 1979 *Nucl. Phys.* B **155**, 237.
Eichten, E. & Lane, K. D. 1980 *Physics Lett* B. **90**, 125.
Einhorn, M. B. & Jones, D. R. T. 1982 *Nucl. Phys.* B **196**, 475.
Ellis, J., Gaillard, M. K., Maiani, L. & Zumino, B. 1980*a* *Unification of the fundamental particle interactions* (ed. S. Ferrara, J. Ellis & P. van Nieuwenhuizen). New York: Plenum Press.
Ellis, J., Gaillard, M. K., Nanopoulos, D. V. & Rudaz, S. 1980*b* *Nucl. Phys.* B **176**, 61.
Ellis, J., Gaillard, M. K., Peterman, A. & Sachrajda, C. T. 1979 *Nucl. Phys.* B **164**, 253
Ellis, J., Gaillard, M. K. & Zumino, B. 1980*c* *Physics Lett.* B **94**, 343.
Ellis, J., Gaillard, M. K. & Zumino, B. 1982*c* *Acta phys. pol.* B **13**, 253.
Ellis, J., Hagelin, J. S., Nanopoulos, D. V. & Tamvakis, K. A. 1983*a* *Physics Lett.* B **125**, 275.
Ellis, J., Ibáñez, L. E. & Ross, G. G. 1982*a* *Physics Lett.* B **113**, 283.
Ellis, J., Ibáñez, L. E. & Ross, G. G. 1982*b* *Nucl. Phys.* B **221**, 29.
Ellis, J. & Nanopoulos, D. V. 1981 *Nature, Lond.* **292**, 436.
Ellis, J., Nanopoulos, D. V. & Rudaz, S. 1982*d* *Nucl. Phys.* B **202**, 43.
Ellis, J., Nanopoulos, D. V. & Tamvakis, K. A. 1983*b* *Phys. Lett.* B **121**, 123.
Farji, E. & Susskind, L. 1981 *Physics Rep.* C **74**, 277.
Farrar, G. & Weinberg, S. 1983 *Phys. Rev.* D **27**, 2732.
Fayet, P. & Ferrara, S. 1977 *Physic Rep.* C **32**, 249.

Fritzsch, H. & Minkowski, P. 1975 *Ann. Phys.* **93**, 193.
Gell-Mann, M. 1977 (Unpublished.)
Georgi, H. 1975 *Particles and Fields* 1974 (ed. C. E. Carlson). New York: A.I.P.
Georgi, H. & Glashow, S. L. 1974 *Phys. Rev. Lett.* **32**, 438.
Georgi, H., Quinn, H. & Weinberg, S. 1974 *Phys. Rev. Lett.* **33**, 451.
Gildener, E. 1976 *Phys. Rev.* D **14**, 1667.
Gildener, E. & Weinberg, S. 1976 *Phys. Rev.* D **13**, 3333.
Goldhaber, M. 1983 *Phil. Trans. R. Soc. Lond.* A **310**, 225.
Goldman, T. J. & Ross, D. 1980 *Nucl. Phys.* B **171**, 273.
Hawking, S., Page, D. N. & Pope, C. 1980 *Nucl. Phys.* B **170** (FS1), 283.
Ibáñez, L. E. & López, C. 1983 *Physics Lett.* B **126**, 54.
Ibáñez, L. E. & Ross, G. G. 1982 *Physics Lett.* B **105**, 439.
Isgur, N. & Wise, M. B. 1982 *Physics Lett.* B **117**, 179.
Kobayashi, M. & Maskawa, T. 1973 *Prog. theor. Phys.* **49**, 652.
Krishnaswamy, M. R. *et al.* 1982 *Physics Lett.* B **115**, 349.
Llewellyn Smith, C. H. 1983 *Phil. Trans. R. Soc. Lond.* A **310**, 253.
Llewellyn Smith, C. H., Ross, G. G. & Wheater, J. F. 1981 *Nucl. Phys.* B **177**, 263.
Marciano, W. J. & Sirlin, A. 1981 *Phys. Rev. Lett.* **46**, 163.
Nanopoulos, D. V. & Ross, D. A. 1979 *Nucl. Phys.* B **157**, 273.
Sakai, N. 1982 *Z. Phys.* C **11**, 153.
Susskind, L. 1979 *Phys. Rev.* D **20**, 2619.
van Nieuwenhuizen, P. 1981 *Physics Rep.* C **68**, 189.
Weinberg, S. 1976 *Phys. Rev.* D **13**, 974.
Weinberg, S. 1979 *Phys. Rev.* D **19**, 1277.
Weinberg, S. 1983 *Phil. Trans. R. Soc. Lond.* A **310**, 249.
Wess, J. & Zumino, B. 1974 *Nucl. Phys.* B **70**, 39.

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF