group suggested (with Pati) was the non-Abelian $SU_c(4) \times SU_L(2) \times SU_R(2)$ with two (rather than three) coupling constants, inbuilt left-right symmetry of the electroweak and with $SU_c(4)$ of colour. This $SU_c(4)$ contains $[SU_c(3) \times U_{B-L}(1)]$ as a subgroup[3]) and combines quarks and leptons of each family (B and L are the baryon and lepton numbers) into one multiplet $(\underset{\sim}{4},\underset{\sim}{2},\underset{\sim}{2})$ consisting of $\underset{\sim}{16}$ (rather than $\underset{\sim}{15}$) two-component fermions (i.e. in addition to ν_L, ν_R also exist). This was the first suggestion of quark-lepton unification. The second suggestion (with one coupling constant) is the postulated symmetry SU(5). Here the 15 quarks and leptons of one family are united in a $\underset{\sim}{5} + \overline{\underset{\sim}{10}}$ of SU(5). There is also the symmetry group SO(10) which can contain both $SU_c(4) \times SU_L(2) \times SU_R(2)$ as well as SU(5) and which describes one single family of $\underset{\sim}{16}$ fermions. The maximal gauge group with $\underset{\sim}{16}$ fermions per family but with a vastly larger number of gauge mesons is SU(16) of which SO(10) would be a subgroup.

Consider the SU(5) grand unifying model. With two Higgs multiplets (a $\underset{\sim}{24}$ and a $\underset{\sim}{5}$) spontaneous symmetry breaking at the tree diagram level will induce the following chain of symmetries:

$$SU(5) \underset{\underset{\sim}{24}}{\rightarrow} SU_c(3) \times SU_L(2) \times U(1) \underset{\underset{\sim}{5}}{\rightarrow} SU_c(3) \times U_{em}(1).$$

The second breaking $\left(SU_c(3) \times SU_L(2) \times U(1) \underset{\underset{\sim}{5}}{\rightarrow} SU_c(3) \times U_{em}(1)\right)$ is assumed to occur at m_W. What is the scale (the so-called grand unification scale) of the first breaking $SU(5) \times SU_c(3) \times SU_L(2) \times U(1)$? This upper scale will determine this stretch of the desert.

To illustrate the ideas involved, consider just the unification of the electroweak sector of the theory $SU_L(2) \times U(1)$, starting with two couplings $\alpha/\sin^2\theta$ and $\alpha/\cos^2\theta$ (both evaluated by experiments involving energy and momentum transfers of order m_W) into a non-Abelian unifying symmetry G, which need not be specified. In this scenario the symmetry G is assumed to break down to SU(2) x U(1), around M_U through the Higgs mechanism. Now the renormalization group tells us that both α and $\sin^2\theta$ are functions of energy; that $\alpha/\sin^2\theta$ decreases logarithmically with energy and $\alpha/\cos^2\theta$ increases with it. M_U is the energy where the two curves will meet.

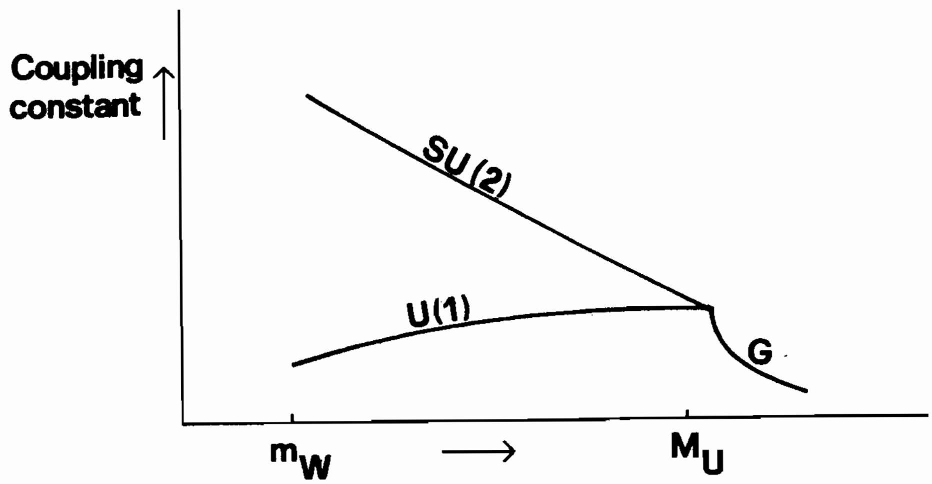


Fig.1

In Fig. 1 are plotted the evolution curves of the two coupling constants. Ignoring contributions from the Higgs particles, these curves meet at M_U given by the formula

$$\ell n(M_U/m_W) = \frac{3}{\alpha(m_W)} \; \frac{\sin^2\theta(M_U) - \sin^2\theta(m_W)}{\cos^2\theta(M_U)}. \qquad (1)$$

To compute M_U we now make two assumptions: 1) There are no new gauge forces to interrupt the evolution curves for the coupling constants shown in Fig. 1 and 2) that we have some theoretical criterion for determining $\sin^2\theta(M_U)$ at the upper energy M_U - assuming that we know $\sin^2\theta(m_W)$ evaluated at the lower scale m_W from experiment.

Now it can be shown that at the upper scale M_U,

$$\sin^2\theta(M_U) = \frac{9\,N_q + 3\,N_\ell}{20N_q + 12N_\ell}$$

where N_q and N_ℓ are the numbers of the fundamental quark and leptonic SU(2) doublets with masses below M_U (assuming that these are the only types of SU(2) multiplets which can exist and that each of the three colours counts once). For the known families, it happens that $N_q = N_\ell$[4]), thus $\sin^2\theta(M_U) = \frac{3}{8}$, if no other types of multiplets are discovered to upset this.

Using (1) with $\sin^2\theta(M_U) = \frac{3}{8}$ as input, and with an empirical value of $\sin^2\theta(m_W) = 0.23$, one would obtain for the unifying mass $M_U \approx 1.3 \times 10^{13}$ GeV. This is a high value for the unifying mass.

To obtain it, we made several assumptions: (1) that there is a desert of new gauge forces up to the mass scale M_U; (2) that there is a desert of intermediate mass scales which the Higgs particles may introduce and (3) that there is a desert of new fermions right up to the (large) mass M_U, such as may shift $\sin^2\theta(M_U)$ from its (unrenormalized) value $\frac{3}{8}$. Making these extrapolations from what we know from experiments below 40 GeV, up to the energy scale M_U, we find that our theory tells us that the desert of new gauge forces, the desert of new Higgs and the desert of new "unconventional" fermions should stretch all the way up to inordinately high energy scales $M_U \sim 10^{13}$ GeV.

To take one counter example of new types of forces which may invalidate this scenario, even within the context of uniting just $SU_L(2) \times U(1)$, remark that we have ignored the very likely experimental possibility that <u>three-family-universality</u> may not hold up to M_U ($\approx 10^{13}$ GeV). Let us relax this assumption; assume this universality is a low-energy phenomenon, i.e. that it holds only up to a mass scale $M < M_U$. Assume that the starting symmetry is $G_e \times G_\mu \times G_\tau$ with a $e \to \mu \to \tau$ discrete symmetry built in to guarantee a unique coupling constant and that each G_i ($i = e,\mu,\tau$) breaks to $[SU(2) \times U(1)]_i$ at M_i. There is the lower breaking stage (which restores e, μ, τ universality) and the emergence of the diagonal-sum symmetry $[SU(2) \times U(1)]^{e+\mu+\tau}$ at the scale M. For this simple scenario, one can now show that the analogue of formula (1) reads:

$$\ell n \frac{M_e M_\mu M_\tau}{M^2 m_W} = \frac{3\pi}{\alpha(m_W)} \frac{\sin^2\theta(M) - \sin^2\theta(m_W)}{\cos^2\theta(M)}.$$

Assume that M, the scale up to which e, μ universality may hold, may be as high as 10^5 GeV, though we are far less certain of e, μ universality empirically. Now, even with $\sin^2\theta(M) = \frac{3}{8}$, we obtain M_e, M_μ, M_τ as low as 10^8 GeV for the case of three families. The "desert" has shrunk from 10^{13} GeV. If the number of families increased to four, the "desert" would shrink still further and stretch only between $M = 10^5$ and 10^6 GeV. Such is the sensitivity of logarithmic functions to (small) changes of inputs!

The conclusion we arrive at is that the stretch and the extent of the desert is crucially dependent on the assumption made. In particular, the simplest assumption of a breakdown of family universality can shrink the desert drastically. The same would happen if $N_q \neq N_\ell$: i.e. whenever the numbers of (left-handed) quark and leptonic doublets differed from each other; this could be the case if mirror fermions exist. Such theoretical extrapolations from present experience (no breakdown of universality, no mirrors) are aesthetically motivated. These however drastically affect the stretch of the desert simply because the renormalization group formulae which we use involve logarithms of masses.

I have so far been speaking about uniting the electroweak forces SU(2) and U(1) with two coupling parameters into one (non-Abelian) structure with one coupling parameter. One could carry out a similar analysis if we wish to unite SU(3) of colour together with $SU(2) \times U(1)$ into a (non-Abelian) symmetry with one coupling parameter - for example the symmetry SU(5) for each family. Here there are three evolution curves for the three couplings, and the demand that all <u>three</u> meet for the same M_U can apparently be met, with $M_U \sim 10^{14}$ GeV and $\sin^2\theta(m_W) \approx 0.21 - 0.23$ provided $\alpha_{colour}(m_W)$ is assumed to be of the order of $\approx \frac{1}{10}$.

Once again, if we did not entertain family universality holding right up to M_U (i.e. we assume the symmetry is $SU_c(5) \times SU_\mu(5) \times SU_\tau(5)$ which breaks at a lower energy to $SU(5)\big|_{e+\mu+\tau}$), the stretch of the desert can tumble down drastically (for example to 10^8 GeV or lower). Another possible symmetry which incorporates the three families is the symmetry $[SU(6)]^4$. This symmetry would predict new leptons in each family ($\sin^2\theta(M_U) = \frac{9}{28}$ instead of $\frac{3}{8}$) and also give $M_U \approx 10^8$ GeV (besides $\sin^2\theta(m_W) \approx 0.23$).

Clearly the determination of the energies up to which e, μ, τ universality holds becomes a parameter of crucial importance for determining the stretch of the "desert".

In the next section we consider the relaxing of the assumption about there being a desert of Higgs particles.

IV THE RICHNESS ASSOCIATED WITH HIGGS PARTICLES

Gauge theories have been surmised to resolve some of the outstanding puzzles of early cosmology - for example the problem of baryon-antibaryon asymmetry, and the problem of nucleation of galaxies. Likewise these theories may account for masses for neutrinos, and they may lead to neutron-antineutron oscillations. It is important to realize that this richness of physics comes associated with Higgs particles, a multitude of which must be introduced into the theory to bring about specific varieties of spontaneous symmetry breaking. The minimal Higgs structures associated with SU(3) x SU(2) x U(1) (one Higgs) or with the minimal SU(5) (two Higgs multiplets) are insufficient. Thus an important future experimental task is to test the phonomena predicted and to explore the energy regimes where such Higgs might be operative.

Consider for definiteness the grand unifying model SU(5). The minimal Higgs structure consists of a $\underset{\sim}{24}$ and a $\underset{\sim}{5}$. The $\underset{\sim}{24}$ breaks SU(5) into SU(3) x SU(2) x U(1) around $M_U \sim 10^{14}$ GeV, while the $\underset{\sim}{5}$ breaks SU(2) x U(1) to $U_{em}(1)$ around 10^2 GeV. Let us leave aside for the next section the hierarchy problem - i.e. the problem that if such a breaking is arranged through an appropriate choice of the Higgs parameters in a tree approximation, there is no "natural" way in which the radiative corrections can be made to respect and preserve this large ratio of around $10^{12} \simeq M_U/m_W$ for higher radiative corrections in a perturbative context, except possibly through an invocation of supersummetry (and its own complicated Higgs structure). In this section we shall simply wish to list in Table II those Higgs which have been postulated from time to time to explain away old puzzles or to predict new physics.

A similar proliferation of Higgs is necessary in SO(10) where the minimal set of $\underset{\sim}{45}$, $\underset{\sim}{126}$ plus $\underset{\sim}{10}$ needs supplementation with complex $\underset{\sim}{10}$ or $\underset{\sim}{120}$ to avoid undesirable family mass relations like $m_\mu/m_e \simeq m_s/m_d$ as well as a complexification of $\underset{\sim}{126}$, $\underset{\sim}{45}$ and $\underset{\sim}{10}$ to cope with the cosmological problems mentioned in the SU(5) context. The SO(10) model admits also of (V+A) currents and contains W_R as well as ν_R, with the possibility of an intermediate mass scale in the 10^6–10^7 GeV range. These and other grand unifying models [like SU(16)] which admit of possible $\Delta(B-L) \neq 0$ baryon decays into (1) three leptons (or three antileptons) or (2) into a lepton (as contrasted with the $\Delta(B-L) = 0$ decay baryon $\rightarrow$ antilepton, mediated in the minimal SU(5) by massive gauge particles) need Higgs in the mass ranges 10^5–10^6 GeV and 10^8–10^9 GeV, respectively[5]).

We now ask what are the allowed mass ranges for the Higgs I have mentioned? This is naturally what the accelerator-builder will want to know. In this context, Marciano at the Paris Conference in July 1982, presented the following constraints for an enlarged SU(5) model for those Higgs ($\underset{\sim}{5}$, $\underset{\sim}{10}$, $\underset{\sim}{15}$, $\underset{\sim}{45}$, $\underset{\sim}{50}$) which may couple to fermions.

Table II

SU(5) representation	Reason for introducing	
24 or 75	To break SU(5) → SU(3) x SU(2) x U(1) with associated energy scale ≈ 10^{14} GeV	minimal SU(5)
5	To break SU(2) x (U1) → $U_{em}(1)$ ≈ 10^2 GeV	minimal SU(5)
45	SU(5) used for all three families, predicts the successful relation $\frac{m_\tau}{m_\mu} \approx \frac{m_c}{m_s}$, but by the same token gives the disasterous $\frac{m_\mu}{m_e} \approx \frac{m_s}{m_d}$ (wrong by a factor of 20 or so). Need 45 to avoid this problem.	
15	To give ν_L a mass.	
More 5's or 45's	SU(5) (or any other grand unifying theory) gives an explanation of baryon-asymmetry. For minimal SU(5), however the predicted value is $\frac{n_B}{n_\gamma} \sim 10^{-15}$–$10^{-16}$ as against the empirical $\sim 10^{-11}$. To correct this, need extra 5's or 45"s.	
Complex 5	To motivate the axion mechanism and thereby avoid strong CP problem in SU(5).	
Complex 10	The axion mechanism, as a rule, brings with it the domain wall problem; such domains would make it hard to understand the isotropy of 3°K radiation. However in an inflationary universe scenario these problems may take on a different complexion.	
Further complex Higgs	To generate strings which produce density perturbations, which may nucleate galaxies.	
10, 15 or 50 or both	For neutron-antineutron oscillations and for $\Delta(B-L) = 0$ processes like $p+p \to e^+ + e^+$. (For $n \leftrightarrow \bar{n}$ with $\tau_{n-\bar{n}} \sim 10^7$ secs, need Higgs of mass ~ 10.5 GeV). Also need these multiplets for (B-L) violation.	

SU(5) SU(3) x SU(2) x U(1) content

$\underset{\sim}{5}$ = $(1,2,1) + (3,1,-2/3)$ $(p \to \mu^+ K^0, \bar{\nu}_\mu K^+ \ldots$ imply $m_2 > 10^{10}$ GeV)

ϕ_1 ϕ_2

$\underset{\sim}{10}$ $(1,1,2) + (3,1,-4/3) + (3,2,1/3)$

ϕ_3 ϕ_4 ϕ_5

$\underset{\sim}{15}$ $(1,3,2) + (3,3,1/3) + (6,1,-4/3)$

ϕ_6 ϕ_7 ϕ_8

(for $\underset{\sim}{10}$ and $\underset{\sim}{15}$:) B-L violation neutrino mass n-$\bar{n}$ oscillation

$\underset{\sim}{45}$ $(1,2,1) + (3,1,-2/3) + (3,3,-2/3) + (\bar{3},1,8/3) + (\bar{3},2,-7/3)$

ϕ_1 ϕ_{10} ϕ_{11} ϕ_{12} ϕ_{13}

$+ (\bar{6},1,-2/3) + (8,2,1)$

ϕ_{14} ϕ_{15}

$\underset{\sim}{50}$ $(1,1,-4) + (3,1,-2/3) + (\bar{3},2,-7/3) + (6,1,8/3)$

ϕ_{16} ϕ_{16} ϕ_{18} ϕ_{19}

$+ (\bar{6},3,-2/3) + (8,2,1)$

ϕ_{20} ϕ_{21}

One can show (from proton decay considerations) that m_2, m_{10}, m_{11}, m_{12}, $m_{17} > 10^{10}$ GeV. For the others (which mediate D^0-$\bar{D}^0$, B-$\bar{B}^0$, n-$\bar{n}$, H-$\bar{H}$

oscillations, neutrino mass and other rare processes) the deviations for the computed $\sin^2\theta(M_U)$ and τ_{proton} from the minimal SU(5) predictions provide the following constraints:

$$\sin^2\theta(m_W)=0.210-\frac{\alpha(m_W)}{36\pi}\,\ell n\left[\frac{m_w^{\ 2}\, m_5^4\, m_6^7\, m_7^4\, m_9^2\, m_{11}^{24}\, m_{20}^{33}}{m_2^2\, m_3\, m_4^3\, m_8^{11}\, m_{10}^2\, m_{12}^7\, m_{14}^9\, m_{15}^4\, m_{16}^4\, m_{17}^2\, m_{18}^4\, m_{19}^{19}\, m_{21}^4}\right]$$

$$\tau_P = 1 \times 10^{30} \text{ years } 1 \left(\frac{m_w\, m_3\, m_6^5\, m_9\, m_{11}^3\, m_{12}^4\, m_{13}^7\, m_{16}^4\, m_{18}^7\, m_{19}^4}{m_2\, m_5\, m_7\, m_8^4\, m_{10}\, m_{14}^6\, m_{15}^8\, m_{17}\, m_{20}^6\, m_{21}^8}\right)^{2/33} .$$

Clearly these are not too restrictive constraints.

Another rich source of new fermions, new Higgs and new physics, which has been speculated upon arises from the desire to remove the family degeneracy, inherent in family groups like SU(5) and SO(10).

One assumes that there exist trival unifying symmetries like SU(7) incorporating two of the known families (besides many new fermions), or SU(11) incorporating all the three families or SO(14) incorporating two of the known plus the mirror families or SO(10) x SO(10) or SU(5) x SU(5), etc. One then starts over again with new Higgs to mediate the breaking of these symmetries, and to push the unwanted fermions to higher unobserved masses. The variety of such symmetries, and their fermionic and suggested Higgs content, is so large that it would be pointless to list them here. Such symmetries would of course give rise to flavour changing neutral currents whose strength may be expected to be $\approx 1/M_F^2$, which is the characteristic mass scale at which these tribal symmetries break down to the simpler family symmetries like SU(5) or SO(10). Particularly relevant in this context would be the precise determination of the rate of rare decays like $K_L \to \mu^+ e^-$, etc., which can, of course, be undertaken at present accelerators, though their cross-sections would have "normal" rates beyond the "transition" energies of the relevant Higgs.

V RICHNESS ASSOCIATED WITH SUPERSYMMETRY

Besides the family problem, the most troublesome problem for grand unification ideas is the hierarchy problem - the "naturalness" of large numbers like $M_U/m_W \sim 10^{12}$ in a perturbative context. Supersymmetry has been suggested as a way out of this but even without this to commend it, supersymmetry - and its promise of richness of physics - must be taken very seriously.

Supersymmetry - the symmetry between fermions and bosons - is an incredible symmetry. It could have been discovered at any time subsequent to 1935 after the canons of quantum field theory had been established. However, even in 1971 when it was first conceived of in the USSR, its existence went unnoticed and its significance missed. The situation persisted till the symmetry was rediscovered in 1973. And even thereafter, though one recognized its elegance quite early, the freedom of supersymmetric Lagrangians from field theoretic infinities, and the remarkable positivity of supersymmetric Hamiltonians, the fact that there is no direct evidence for its existence at the low energies hitherto available has meant its being somewhat ignored hitherto.

The hierarchy problem arises because there is no mechanism in ordinary theories by which a spin-zero Higgs which starts life with a small mass ($\lesssim$ 100 GeV) can protect itself from acquiring mass of the order of 10^{14} GeV (through its interactions with other Higgs which are needed by the theory with large (10^{14} GeV) mass). Thus radiative corrections destroy any hierarchy with which we may start. There is however a protective mechanism for fermions - chirality. So why not tie all bosons with corresponding fermions through supersymmetry? Thus in a supersymmetric SU(5) we may "protect" a $\underset{\sim}{5}$ and a $\overline{\underset{\sim}{10}}$ of Higgs by placing them in the same supersymmetric multiplets as the $\underset{\sim}{5}$ and $\overline{\underset{\sim}{10}}$ of the fermions. Remember the doublet in the $\underset{\sim}{5}$-fold of Higgs was needed (around 100 GeV) to act as the familiar Higgs of SU(3) x SU(2) x U(1).

Unfortunately things are not all that simple. We do want, for example, the triplet contained inside the $\underset{\sim}{5}$-fold of Higgs to be as massive as 10^{14} GeV so as not to enhance proton decay rate. These conflicting phenomenological requirements (of a light doublet versus a heavy triplet inside the same $\underset{\sim}{5}$) can be met - as a rule by the standard device of doubling everything (if one $\underset{\sim}{5}$-fold does not work, take two) - but this needs a careful adjustment of parameters.

For normal renormalizable field theories, radiative corrections destroy such careful adjustments, but amazingly enough, not for supersymmetric theories, where they are stable. As a rule, after the adjustments and the doublings of the multiplets the final supersymmetric grand unifying theories which emerge are baroque affairs. Presumably, with experience, this will be set right.

But the major question which remains open is this: what is the scale of supersymmetry breaking m_S? It turns out that there are two theoretical choices. One is ≈ 1 TeV and the other is much higher around $m_S = (m_W m_{Planck})^{\frac{1}{2}} \approx 10^{10} - 10^{11}$ GeV. But even for the higher scale m_S one must emphasize that there is a promise of experimental signatures for lower than TeV energies, since the supersymmetric partners in such models can acquire masses differing by $(\frac{\alpha}{\pi})^n m_S$.

What are the signatures? All quarks and leptons have scalar partners "squarks" and "sleptons"; all gauge particles have fermion partners, gluinos $(\tilde{g})$, photinos $\tilde{\gamma}$, W-inos, Z-inos, etc. For supersymmetry breaking in the TeV range, the photinos may be expected below a few GeV and gluinos around 30 GeV. Present experimental analyses place only meagre limits on the $m(\tilde{g}) > 2$ GeV, m (sleptons) > 16 GeV.

So much for global supersymmetry. However, supersymmetry like all symmetries can be gauged, and the gauge particles turn out to be the graviton and its fermionic partner of spin $\frac{3}{2}$ - the gravitino. Clearly, if we wish to unite gravity with other forces, a gauging of supersymmetric grand unifying models is one way to accomplish this. As we shall see, in the next section, the ultimate expression of this line of thought is the self-gauged "N = 8 extended supergravity theory" which possesses a unique multiplet (with a unique self-coupling) consisting of one graviton, eight gravitinos, twenty-eight vector mesons, fifty-six spin $\frac{1}{2}$ two-component objects and seventy scalars. Can this multiplet and its self-interaction accommodate all known particles, their symmetries, and their interactions? The answer turn out to be NO; we shall examine this further in the next section on preons.

Let us for the moment be content with the humbler version of supergravity, where a supersymmetric grand unifying model like SU(5) interacts with the simplest version N = 1 of a supergravity multiplet, consisting of one graviton and one gravitino. What mass does the gravitino acquire on account of supersymmetry breaking? A simple way to compute this is to observe that in such a theory a cosmological constant arises which can be made to acquire its empirically determined value of "zero" by giving the gravitino a mass of $\approx m_S^2/m_{\text{Planck}}$. This could be as small as $\approx 10^{-16}$ GeV or as large as $\approx m_W$ (or larger) depending on at what energy (m_S) supersymmetry breaks. Since the helicity $\frac{1}{2}$ component of the gravitino has been shown by Fayet to have an effective interaction of strength E^2/m_S^2 for the lower value of gravitono mass (in the electron-volt range) its coupling could be large; on the other hand, with the higher scale for m_S, there is the exciting possibility that (the spin $\frac{3}{2}$) gravitono pair production (with gravitinos of mass $\gtrsim m_W$) may provide an important search-project for the $p\bar{p}$ collider or for the higher energy accelerators. One must emphasize that we are still very far from a standard and favoured supergravity or even a supersymmetric model. My purpose in mentioning these ideas is merely to emphasize the richness which these prospects promise.

VI THE NEXT LEVELS OF STRUCTURE; PREONS, PRE-PREONS

As repeatedly emphasized already, the most mysterious aspect of present day phenomenology is the existence of apparently "similar" recurrences of families of fifteen two-component quarks and leptons. Whenever such recurrences have shown themselves to occur in the past, we have eventually uncovered a new layer of elementarity. Will this happen again, before the onset of the energy scale $M_U \approx 10^{14}$ GeV. Do quarks and leptons possess radii much larger than $(M_U)^{-1}$? The present limits on leptonic radii are only in the (10 TeV - 100 TeV)$^{-1}$ range, the precise values depending on the definition adopted. Clearly this will be one of the most beckoning tasks for experimental searches in the near future, to determine quark and lepton structure.

The simplest "preonic" model is the one which associates light fundamental entities, one each with four colours (c = red, yellow, blue and B-L, and one each with four flavours (f = up-left, down-left, up-right and down-right). If the four chromons (c) are spin-zero and four flavons (f) are spin $\frac{1}{2}$, there may even exist a supersymmetric version of the preonic theory with four basic supersymmetric multiplets, where supersymmetry breaking is synonymous with the emergence of colour and flavour quantum number and "composite" symmetries $SU_c(4) \times SU_L(2) \times SU_R(2)$.

The family distinctions can be built into the theory in diverse ways; one of the simplest (in a non-supersymmetric version of preonic theory) is to postulate three familons in addition to chromons and flavons, one for each family. If familons are (analogoue) dyons carrying (analogue) electric and magnetic charges (e.g.) and chromons and flavons carry charges (-e,0) and (0,-g) (with $eg/4\pi = n/2$), the binding force could be an (analogue) magneto-electric U(1) force. Quarks, leptons, Higgs and SU(4) x SU(2) x SU(2) gauge particles would then be the uncharged composites of flavons, chromons and familons.

There are of course other versions of the preonic models. In one of these the unbroken SU(3) of colour and U(1) of electromagnetism are accorded a privileged status as truly fundamental forces with "elementary" gauge particles associated with them, while the $W^{\pm}$, Z^0 of the electro-weak force are composites. The "elementary" preonic fermions are assumed to be $(3,\bar{3})_{L,R}$ and $(3,3)_{L,R}$ of a $SU(3)_{\text{hyper-colour}} \times SU(3)_{\text{colour}} \times U(1) \times U(1)$ with the new strong hyper-colour gauge force binding the preons together. The family distinctions are brought in, through varying the numbers of preonic pairs in the composites.

Finally, there is the supergravity preonic model which treats the unique N = 8 supergravity multiplet as referring to preons rather than to quarks, leptons and Higgs, etc.[6]. I shall not describe this version of the theory in any further detail except to remark that the model may accommodate a chiral SU(5) grand unifying theory, though the question of whether the three families do indeed emerge as composites is not fully settled. However it is clear that if this preonic model is the correct one, the quark and leptonic radii are not likely to be larger than inverse Planck mass $\approx (10^{19}\text{ GeV})^{-1}$.

In the preonic context, an important question is to state criteria which guarantee that if preons are (chirally protected) massless fermions, the composites are massless as well. 't Hooft has attempted to formulate such criteria in terms of anomaly-matching of preonic multiplets with the expected multiplets of composite bound states. These criteria have proved difficult for realistic models[7] to satisfy and have led 't Hooft to suggest that a high degree of complexity in particle spectrum seems unavoidable. Stated differently, there may be an unending chain of "elementary" structures, quarks, preons, pre-preons,... associated with an unending chain of gauge groups SU(3), SU(4), SU(5),..., SU(N), where $N \to \infty$ on a linear energy scale. Presumably with this scenario the (accelerator) physicist will never be at a loss for new discoveries!

Contrast this with the view advocated by some of us that the preon, pre-preon,... chain may end "monotheistically" with one unique multiplet of one unique symmetry. I have mentioned N = 8 supergravity preonic theory in this context. There is a more unique supersymmetry, the N = 4 supersymmetry with the particle types 1 vector particle, four (Majorana) fermions and six (real) scalars, all in the adjoint representation of a non-Abelian symmetry group[8]) (like SU(2)). This theory has been shown to have no infinities whatsoever - even when the fermions and scalars are (N = 2 supersymmetrically) massive. And it may be the only theory in particle physics to exhibit this finiteness. Furthermore, Grisaru and Schnitzer have shown that a Reggeization of this theory could lead to a set of composites - none other than the N = 8 preons mentioned above, while Osborn[8]) has shown that the theory also possesses solitonic solutions which form a dual multiplet (of one vector, four fermionic and six bosonic triplets of SU(2)) which in its turn describes magnetic monopoles. (Since there is no renormalization of charge, there is no problem of whether it is the unrenormalized electric (e) and magnetic charges (g) or the renormalized ones which satisfy the Dirac condition $eg/4\pi = n/2$. In this dyonic form, is this the ultimate pre-preonic multiplet of which preons and then quarks and leptons are composed? An important experimental question will be: what are the mass parameters associated with this multiplet?[9]).

VII CONCLUDING REMARKS

High energy physics is an intoxicating subject - every generation has felt that it has nearly scaled the truth, and perhaps after the ideas it has espoused have been worked out, there will be a desert of basic principles. And every generation has been proved wrong in the past.

I have concentrated in this lecture on the physics riches which we can now perceive may be in store for future accelerator physics. These concern the physics associated with Higgs in grand unifying theories, physics associated with supersymmetry and with preonic ideas. I have not spoken of the dimly-perceived prospects - like those arising out of extra space-time dimensions and the Kaluza-Klein theories which live on them. Such ideas are intimately related to the prospects of supergravity theories, particularly the N = 8 theory whose most natural formulation is in terms of compactified eleven-dimensional space-time, and will presumably become relevant at much higher energies.

What physics is likely to be associated with the extra dimensions? Do they hold the secret of the charge concept?[10]) What is the topology of these compactified spaces; What is their likely relationship to the cosmology of the early Universe? What is the likely resolution of the unseasonably large cosmological constant such theories appear to support. As Nahm has stressed, are we likely to discover a new principle (perhaps with low energy experimentation) like the equivalence principle, in the context of an empirically vanishing cosmological constant. As I said these are dimly perceived questions at present. But before we grow too wildly speculative, we need experimental direction. And the amazing aspect of the interaction of theory and experiment is that even one well-conceived experiment can be the decisive pointer to give direction to our speculations. We are here to-day to ensure that experiment and theory do not get out of phase in this regard.

Footnotes

1) True enough, after 1979, with the work of Cremmer and Julia and the revival of Kaluza-Klein higher dimensional theories, there has been a slight remission downwards by more than an order of magnitude in energy scales as emphasized by Freund. We believe, for example, that already at a lower energy $\frac{e}{2\pi} m_{Planck}$, of the order of a mere 10^{15} TeV, space-time will have blossomed from four to eleven dimensions. Thus before Planck energies are reached, we may have a totally new regime to deal with.

2) An intermediate energy scale (not available within SU(5)) is the one associated with the breakdown of the left-right symmetry (V + A currents and predicted existence of W_R) characteristic of all grand-unifying theories, except SU(5). Such a breakdown may manifest itself anywhere between $\frac{1}{3}$ and 1000 TeV.

3) As was noted almost immediately after $SU_c(4)$ was postulated, one could consider breaking (B-L) (around 10^5 GeV) and mediate proton decays into three leptons. The breakdown of $SU_c(4)$ into $SU_c(3) \times U_{B-L}(1)$ could also give rise to monopoles of "light" mass ($\approx 10^6$ GeV) which future accelerators may produce.

4) $N_q = N_\ell$ can be theoretically motivated by the demand for axial anomaly cancellation between quarks and leptons. If however one admits the possibility of the existence of mirror fermions, coupling with (V + A) currents to $W^\pm$ and Z^0 (and one can give arguments that such fermions must appear below 300 GeV or so, if they exist at all) then there is no anomaly and no necessity for $N_q = N_\ell$.

5) It is important to remark that a rare process like proton $\rightarrow$ three leptons - rare below 10^5 GeV - would become "normal", beyond the transition energy >> 10^5 GeV.

6) If it is assumed that the N = 8 supergravity theory describes physical quarks and leptons and physical gauge particles, the maximal clasifying group cannot be larger than the vectorial $SU_{colour}(4) \times U_{L+R}(1)$. The $SU(4)_{colour}$ may break into $SU_c(3) \times U(1)_{B-L}$. For this picture $W^{\pm}$, Z^0 must in any case be treated as composites; so must the muon, the taon and the b-quark. It seems preferable therefore to treat all the objects in the multiplet on par as preons and to make all the known particles as composites of these.

7) One simple model (due to Albright, Schremp and Schremp) where a part of these criteria are met contains preons $(\underline{3},\underline{6},\underline{1})_L$ and $(\underline{3},\underline{1},\underline{6})_R$ of $[SU(3)_{hyper-colour} \times SU_L(6) \times SU_R(6) \times U_{L+R}(1)]$ symmetry with $(\underline{1},\underline{6},\underline{15})$ and $(\underline{1},\underline{15},\underline{16})$ of composite quarks and leptons. A notable prediction of this model (shared also by the simplest flavon-chromon model) is the lack of universality of taon-couplings with e,μ couplings (e.g. $e^+e^- \rightarrow \tau^+\tau^-$ should exhibit vanishing charge asymmetry).

8) The internal symmetry SU(p) may not be unique, except possibly for the case when we wish this spectrum to represent dyons when p = 2 is the simplest choice.

9) In this half century, in the science of biology, the analogue of our universal gauge principle was found in 1953 with the discovery of the double helix. However, this has not obscured from the biologist the fact that far from being the "end of molecular biology" this was only a beginning. "Something quite essential is missing in our basic understanding of life and we have not the slightest idea about the nature of lacunae in our knowledge", "The End of Molecular Biology" by A Sibatani, Trends in Biochemical Sciences, Vol. 4, No. 7 (Elsevier, 1979). I believe the same applies to particle physics with the unsolved problem of the nature of charge.

DISCUSSION

Willis. Suppose we have built an accelerator for 100 TeV in the centre of mass, and want to look for preons or Higgs particles, must we use the standard formula for point like cross sections?

Salam. No, if for example quarks have a form factor at 10^5 GeV, those form factors will have made their appearance. Whether that helps or hinders I do not know, but the point-like behaviour should have gone by that time. Events which are now rare ($e \to \mu$, $\mu \to e + \gamma$) will presumably become copious when the threshold for their becoming strong interactions has been passed.

Willis. "Copious" could still be on the scale of point like particles.

Salam. That's right.

Willis. If you specify only the energy to be reached, you have set us too easy a problem!

Salam. Considering again the quark, it has a certain size related to this basic energy of 100 TeV.

Willis. I believe that your answer to my earlier question is 'yes'. We want experiments with our 100 TeV accelerator to have at least one event per human lifetime for a point like cross-section.

Participant. That is very modest!

Hand. Has there been any theoretical progress in understanding the purely leptonic mass ratios? Does this not already indicate that there is a violation of family unversality going on somewhere? Could we not be fooled, in quoting $\mu \to e + \gamma$ as an example, by some other selection rule which prohibits it?

Salam. Higgs' are very powerful. You can 'fix up' μe masses by inventing Higgs mechanism just by putting in a new parameter. In that sense people have been 'fixing things up'; the hope is that someday one will be able to explain those parameters in term of some fundamental formula which will come. There is no theory of mass at the present time. The SU5 model, despite the very great success in predicting one mass ratio being equal to another, has also a great deficiency. If you look at the third multiplet it is out by a factor of twenty or thirty. The theory is inadequate at present; we need a dynamical theory that in calculable, and this doesn't exist yet.

Palmer. I am following up on Willis' questions a little. We know that there is very little scattering of electrons by neutrons at low energies, since overall the neutron is neutral. If you increase the energy and get within the neutron, interactions begin to appear because the charge distribution is not uniform. By analogy, above a certain energy in preon theory you will start seeing the preon 'charge' and the cross sections could suddenly become enormous.

Salam. Enormous relative to these energy factors...

Willis. You have chosen a bad example, which still has the $1/q^2$ in front of its form factor, but which also happens to have a vanishing form factor. We grant that you can put that form factor equal to unity, but I am worried about the kinematics.

Palmer. I wonder whether one can have a new scale factor which has the strength of this preon-preon interaction, which is orders of magnitude higher.

Salam. You are right, even if you have 1 instead of 10^{-2} or 10^{-4} his q^2 is overall and that is something which nobody can help.

Palmer. This is crucial for the luminosity of these machines, and has enormous implications for us machine builders.

Salam. Regretfully, yes. It is just quantum mechanics and we can't fight that!

THE PHYSICS OF PARTICLE ACCELERATION*

J D Lawson

Rutherford Appleton Laboratory, Chilton, Oxon. U.K.

1. Introduction

During the past five decades the energy reached by particle accelerators has increased roughly by a factor 25 per decade. The cost per MeV has decreased by about 16 per decade.(1) This remarkable progress has been sustained by the successive introduction of several new concepts, and of new or improved techniques. The logical outcome of this trend, if it continues with roughly the same level of funding as at present, is that there will be fewer and fewer ever larger accelerators, until eventually there is not enough money to build machines fast enough for interest in experimental high energy physics to be maintained.

We are now at the stage of 'few and large' accelerators; some countries, for example the UK, no longer have even one national laboratory with a machine capable of supporting high energy physics research. The question to be considered at this meeting is as follows: is there still a sufficient flow of new ideas and techniques to sustain the trend towards higher energies, and at the same time make costs fall sufficiently that the future of experimental particle physics is assured beyond the turn of the century?

Many methods exist by which particles may be accelerated. It is important that these should all be examined carefully to see whether some new departure from existing techniques may prove advantageous. At the same time, the basic constraints on existing systems must be well understood.(2).

---

*Minor modifications have been introduced in the paper as presented, to take account of later presentations at the meeting. Such material has been acknowledged by reference to the appropriate speaker.

Since the advantage of colliding beam techniques increases with energy, (the ratio of energy available by colliding similar beams to that of using a stationary target being $\sqrt{2\gamma}$), it follows that unless an enormous 'energy leap' is possible, the new machines must be of the colliding beam type. Cross sections for expected interesting events decrease, so that the demands on luminosity and beam quality are severe. High luminosity implies large mean currents, which in turn means that the requirements for power to operate the machine are great. For this reason the 'wallplug' efficiency, power station to beam, must be high.

2. Classification of Methods of Acceleration

To accelerate particles, an electric field is required. To focus them, magnetic fields can be used, and indeed, these turn out to be the most practical at high energies. New ideas could incorporate much higher electric fields than are now used, and/or more compact focusing systems. As the general understanding of particle dynamics, electromagnetism, and plasma physics becomes more complete, it naturally becomes more difficult to invent new accelerator concepts. It is helpful to try to categorize the different types of accelerator which have been proposed, to see whether any gaps can be found. Accelerating concepts may be classified in several different ways; the one proposed here seems simple and convenient, but is certainly not unique.

TABLE EXHIBITING CLASSIFICATION OF ACCELERATORS

	Accelerated Particles in Free Space		Accelerated Particles in a Medium
	No Free Charges in System	Free Charges in System	
Harmonic Accelerating Fields	1.1 LINAC SYNCHROTRON Inverse free electron laser	1.2 See section 3b in text	1.3 Inverse Cherenkov Beam-wave Laser beat-wave
Accelerating Fields not Harmonic	2.1 Betatron Induction linac	2.2 Ion drag Wake field	2.3 Ionisation front Electron ring

Only six categories are presented, but these can readily be subdivided. Vertically, in the table we distinguish between machines in which the accelerating field at a point varies harmonically, and those in which it does not. The latter category contains a number of lower energy machines, such as the betatron and all electrostatic accelerators, whereas the former contains all existing types of very high energy machine. Horizontally, we distinguish between systems in which the particles move in free space (as in all existing high energy machines) or in some medium such as a plasma or an intense beam of particles of a different type. The first of the horizontal categories is again divided according to whether the charges which produce the required accelerating and focusing fields are entirely bound within metals or dielectrics, or whether they are free, forming part of a plasma or particle beam.

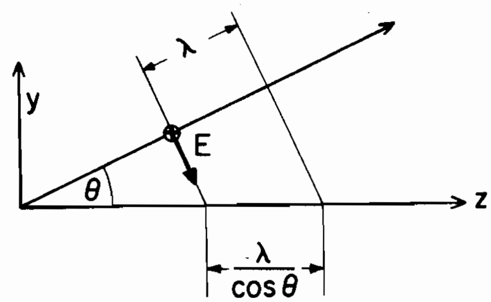
3. Discussion of the Various Categories

3a. Category (1.1)

The first category in the table is the most familiar. It can be examined by considering the very fundamental problem of the interaction of a plane wave and a particle, since all fields in regions of free space can be constructed from plane wave solutions, provided that evanescent as well as propagating waves are included.

We consider first systems in which a wave with component of electric field in the direction of propagation travels at the same velocity as the particle. It is well known that infinite plane waves propagating in free space cannot satisfy this condition. First, the wave always moves faster than the particle, and second, the electric field is perpendicular to the direction of motion. At extreme relativistic energies the wave and particle velocities are very little different, but if the wave is tilted at an angle θ to provide an accelerating field $E_0 \sin\theta$ along the orbit the phase velocity increases to $c \sec\theta$, so that the velocity difference for a significant value of θ is no longer small.

Evanescent waves, on the other hand, can provide the type of field required. Single semi-infinite evanescent phase waves are familiar in



Propagating Plane Wave, cos $\theta < 1$

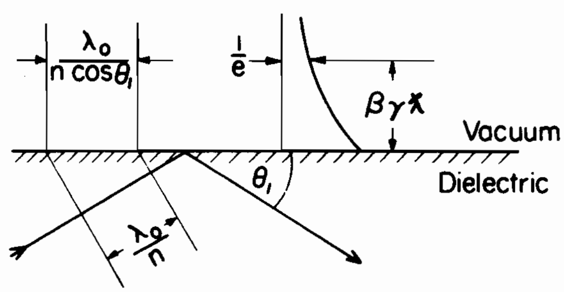
Field components:

$$E_y = -E_o \cos\theta \exp[-i(z\cos\theta + y\sin\theta)/\lambda\!\!\!^- + i\omega t]$$

$$E_z = E_o \sin\theta \exp[-i(z\cos\theta + y\sin\theta)/\lambda\!\!\!^- + i\omega t]$$

$$Z_o H_x = E_o \exp[-i(z\cos\theta + y\sin\theta)/\lambda\!\!\!^- + i\omega t]$$

$$v_z = c/\cos\theta > c$$



Evanescent Plane Wave, cos $\theta > 1$

(Write $\cos\theta = 1/\beta_w$, $\sin\theta = -i/\beta_w\gamma_w$)

Field components:

$$E_y = \frac{-E_o}{\beta_w} \exp\left(\frac{-y}{\beta_w\gamma_w\lambda\!\!\!^-_o}\right) \cos\left(\omega t - \frac{z}{\beta_w\lambda\!\!\!^-_o}\right)$$

$$E_z = \frac{-iE_o}{\beta_w\gamma_w} \exp\left(\frac{-y}{\beta_w\gamma_w\lambda\!\!\!^-_o}\right) \sin\left(\omega t - \frac{z}{\beta_w\lambda\!\!\!^-_o}\right)$$

$$Z_o H_x = E_o \exp\left(\frac{-y}{\beta_w\gamma_w\lambda\!\!\!^-_o}\right) \cos\left(\omega t - \frac{z}{\beta_w\lambda\!\!\!^-_o}\right)$$

Phase velocity $v_z = c/\cos\theta = \beta_w c < c$

Fig 1 Two forms of plane wave

optics as the fields which exist outside a block of dielectric in which total internal reflection is occurring. Both types of wave are illustrated in Fig. 1. The angle that the wave vector makes with the surface is complex, since the cosine of the 'angle of refraction' in free space is given as greater than unity by Snell's law. The phase velocity of the wave along the surface in $c/\cos\theta$, which, since $\cos\theta$ exceeds unity, is less than c. We can write the phase velocity

$$v = c/\cos\theta = \beta_w c \tag{1}$$

from which it follows that

$$\sin^2\theta = -1/\beta_w^2\gamma_w^2 \tag{2}$$

where

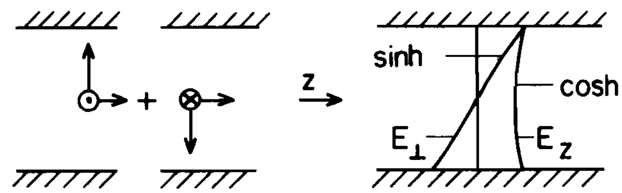
$$\gamma_w = (1-\beta_w^2)^{-\frac{1}{2}}. \tag{3}$$

Such a wave clearly is synchronous with a particle of energy $\gamma m_o c^2$. Henceforth the subscript w is omitted. It is readily shown that the fields decay to 1/e in a distance from the surface equal to $\beta\gamma\bar{\lambda}$, and that for a wave with the appropriate polarisation there is a component of electric field in the direction of propagation (parallel to the surface) given by

$$E_{\parallel} = E_{\perp}\left|\sin\theta\right| = E_{\perp}/\beta\gamma \tag{4}$$

where $E_{\perp}$ is the field component perpendicular to the surface. As γ increases, $E_{\parallel}/E_{\perp}$ decreases, becoming very small at relativistic energies. This seemingly fundamental disadvantage of using evanescent waves can be overcome in either of two ways. First, if a second surface is placed parallel to the first, then a wave can be propagated in which the components of $E_{\perp}$ and H cancel on the plane midway between them, but the $E_{\parallel}$ components add, as shown in Fig. 2a. Away from the central plane $E_{\perp}$ and H increase rapidly, but up to a distance $\pm\,\bar{\lambda}$ they are both less than $E_{\parallel}$. In conventional accelerators a manifold of such evanescent plane waves are combined to produce a system with axial symmetry.

Double-sided Solution



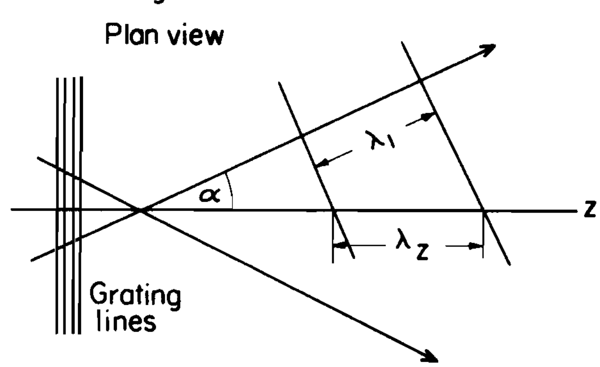
Field components:

$$E_y = \frac{-E_o}{\beta} \sinh\left(\frac{y}{\beta\gamma ƛ}\right) \cos\left(\omega t - \frac{z}{\beta ƛ}\right)$$

$$E_x = \frac{-iE_o}{\beta\gamma} \cosh\left(\frac{y}{\beta\gamma ƛ}\right) \sin\left(\omega t - \frac{z}{\beta ƛ}\right)$$

$$Z_o H_x = E_o \sinh\left(\frac{y}{\beta\gamma ƛ}\right) \cos\left(\omega t - \frac{z}{\beta ƛ}\right)$$

Single-sided Solution



$$\lambda_z = \lambda_1 \sec \alpha$$

$$\lambda_1 = \lambda_o \cos \theta$$

$$\lambda_z = \lambda_o \text{ when } \cos\theta\cos\alpha = 1$$

Fig 2 How to achieve finite E_z when γ is very large.

Instead of being supported by a dielectric, they can be carried by a corrugated surface exhibiting a reactive surface impedance, or as space-harmonics in an appropriate periodic structure. In practical systems other space-harmonics (including backward travelling waves) occur, but the essential physical features of the interaction are common to all types of linac and to synchrotrons. Interaction by means of evanescent waves implies that virtual photons are exchanged between the cavity wall and the accelerated charges. (It is well known that particles cannot absorb single real photons, since the conditions for energy and momentum conservation cannot simultaneously be satisfied).

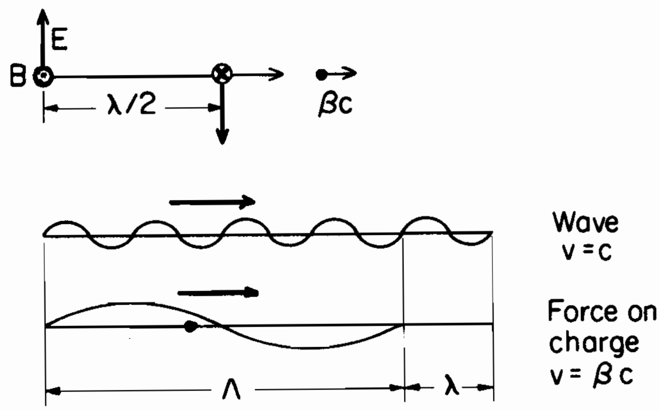
An alternative configuration, making use of a one-sided system, has been proposed by Palmer(3). This is illustrated in Fig. 2b. If we consider a space harmonic for which $\sin\theta$ is not small, so that $1/\cos\theta$ is not near to unity, then in a direction parallel to the surface but at an angle α to the wave normal, the phase velocity will be $v/\cos\theta$. If α is chosen such that

$$\cos\theta\cos\alpha = 1 \tag{5}$$

then the phase velocity of the wave along the z-direction will be c. For example, if $\alpha = 45°$, $\cos\theta = \sqrt{2}$ and in the z-direction $E_{\parallel}/E_{\perp} = |\sin\theta/\cos\theta| = 1/\sqrt{2}$. By using a grating rather than a dielectric surface, and generating a suitable manifold of wave-pairs travelling at $\pm\,\alpha$ to a line perpendicular to the grating lines, a one sided 'grating accelerator' suitable for optical wavelengths can be established.

Although different in form from existing accelerators, and operating at a different wavelength, the grating linac is essentially the same class of device. There is scope for looking in a very general way at accelerators of this broad class, to see whether the types of structure and frequencies of operation of present machines lie close to the optimum. (Just how this 'optimum' is to be characterised, of course, needs some thought). This may well show that there are more favourable operating regions when considered in terms of fundamental constraints, such as breakdown limits. Technological limits, such as availability of power sources without excessive development costs, may in fact turn out to be the more important.

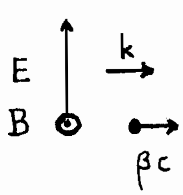
Charge and Plane Wave



$$\Lambda = \frac{\lambda}{1-\beta} \approx 2\gamma^2\lambda$$

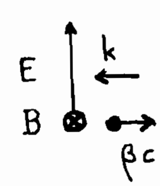
Modulate orbit 1) by wiggler of wavelength $\Lambda = 2\gamma^2\lambda$

or 2) by oppositely directed wave with $\Lambda = (1+\beta/1-\beta) \approx 4\gamma^2\lambda$



$F \propto 1-\beta$

Electric and magnetic forces almost cancel



$F \propto 1+\beta$

Electric and magnetic forces add giving stronger interaction

Fig 3. Illustrating the inverse free electron laser mechanism limits the transverse excursion. At high energies the loss by synchrotron radiation becomes important for electrons.

The transverse deflection can alternatively be provided by a second electromagnetic wave travelling in the opposite direction to the laser beam. The effects of the electric and magnetic fields of this second wave now add (rather than subtract) to produce a transverse deflection of wavelength

$$\Lambda = \lambda(1+\beta)/(1-\beta) \approx 4\gamma^2\lambda . \tag{9}$$

Yet another possibility is to use an evanescent wave travelling at an angle to the particle beam(4).

This type of acceleration makes use of laser light without the severe restrictions arising from proximity to conductors, and from the small transverse dimensions necessary in the grating accelerator described earlier. To pay for these advantages it has the problem that the accelerating field and particle velocity are almost perpendicular.

Despite these problems, impressive acceleration rates are indicated in a study by Pellegrini.(5) High intensity, however, makes severe demands on laser power and repetition rate.

3b. Category (1.2) in Table

As originally presented, there were no items in this category. It was pointed out at the meeting, however, by Tigner and Keefe(6) that if one regards a linear accelerator and its driving klystrons as a single system, then there are 'free charges' in the klystron. The klystron cavities and accelerator structure form an impedance transformer which transfers energy from high current low energy beams in the klystrons to a high energy low current beam in the accelerator. It was postulated that it might be possible to devise a system in which the beams were closer together and the 'transformer' structure simpler and more efficient.

The inverse free electron laser mechanism is also placed in category (1.1). Here the particle and wave velocites are different, and there are essentially two waves interacting simultaneously with the particles.(One of these may be a static 'wiggler' magnet, with finite k but zero frequency). This interaction will now be examined, again considering the basic components of plane wave and particle.

We consider first a particle and plane wave moving in the same direction, as shown in Fig. 3. The wave is characterised by wavelength λ, and the particle by normalized energy γ, where γ is large, so that

$$\beta \approx 1 - 1/2\gamma^2. \tag{6}$$

The particle experiences a transverse electric force, almost balanced by a magnetic Lorentz force in the opposite direction,

$$F = eE(1-\beta) \approx eE/2\gamma^2. \tag{7}$$

The direction of this force alternates, the particle experiencing one complete cycle in a distance

$$\Lambda = \frac{\lambda}{1-\beta} \approx 2\gamma^2\lambda . \tag{8}$$

If now a small transverse deflection of wavelength Λ is induced in the particle orbit, particles at the appropriate phase with respect to the plane wave experience a force which is outward when the particle moves outward, and inward when the particle moves inward. The particle thus gains energy at a rate $eE_{\perp}v_{\perp}$. Continuous acceleration is achieved, but this is a second order process since the electric field and orbit directions are almost orthogonal.

It may be asked why a large transverse deflection is not possible. First, if the deflection is not small, eq. (8) is no longer true, and for given λ, Λ becomes smaller. To combine this smaller Λ with large transverse velocities requires excessive fields in the deflecting magnets if λ is kept small so as to make use of the high fields obtainable from lasers. Further, high fields require a compact laser beam, so this also

3c. Category (1.3) in Table

Three examples of acceleration methods in this category may be cited, in which the medium is respectively a gas, dense particle beam, and plasma.

In a gas the phase velocity of electromagnetic radiation is slowed down. If n is the refractive index, continuous interaction can occur between electromagnetic waves travelling at an angle θ to a particle moving with velocity βc when

$$1/\beta n = \cos\theta \ . \qquad (10)$$

If the phase of the particle is such that energy passes from the wave to the particle this may be termed inverse Cherenkov acceleration. If laser light is used, limitations arise from gas breakdown and scattering of the accelerated particles. Scattering is excessive unless the gas density is low, in which case $\cos\theta$ is nearly unity, and the component of electric field in the direction of particle motion is small.(7)

The possibility of accelerating particles by the longitudinal electric fields associated with waves on electron beams has received extensive study over a number of years. Early experiments in the USSR are described by Fainberg.(8) A later suggestion has been to use a beam carrying a negative energy wave.(9) The idea is that as the particles are accelerated, energy is removed from the wave, which therefore grows. Unfortunately it has proved very difficult to realise this idea in practice. Both experimental experience and general analysis suggest that this concept is not promising for producing high accelerating gradients at particle velocities approaching that of light.

The latest idea in this general category is the beat-wave accelerator of Tajima and Dawson.(10) Two intense laser beams, with frequencies ω_1 and ω_2 are fired into a plasma of density such that $\omega_1 - \omega_2$ is equal to the plasma frequency. This causes strong non-linear bunching, which is manifest in the form a plasma wave which advances at almost the velocity of light. The suggestion is that extreme relativistic particles be accelerated in the longitudinal electric field of this wave. (This combination of light velocity and a longitudinal electric field is, as explained earlier, not possible in free space). Preliminary calculations suggest the possibility of generating enormous fields with the power densities obtainable from large modern lasers. An elegant simplified analysis of an accelerator system by Ruth and Chao indicates fields of 5 GeV/metre.(11) Whether fields of this order can be sustained and controlled over a sufficient distance to make a practical accelerator remains to be seen; further evaluation is awaited with interest.

3d Category (2.1) in Table

We now consider accelerators in which the accelerating field does not vary harmonically with time. In the first category of such devices the particles move in a vacuum. Several important types of accelerator are shown in the table; their limitations are well understood, and none of them shows promise as a very high energy machine. It is interesting to note that the first of these, the multigap accelerator, now proposed as the initial stage of a high current accelerator for heavy ion fusion(12), was first proposed as early as 1924.(13) At that time the technology to make it work was not available.

3e Category (2.2) in Table

An early example of this class of accelerator, in which free charges are used, though the particles to be accelerated are in a vacuum, was suggested in 1952.(14) An intense electron beam is focused through a sharp waist. Near to this waist is an intense radial electric field. If the whole beam is propelled sideways, then positive particles can be dragged along and accelerated. The scheme for moving the beam sideways originally proposed does not look technically very practical, and the interest in this method is now purely historical.

A new method of dragging relativistic particles along has recently been suggested by Voss and Weiland.(15) Relativistic electrons in the form of a ring are accelerated through a suitable shaped periodic structure, the axis of which coincides with the axis of the ring. For a ring containing of order 10^{11} particles, there is a trailing wake field at the centre of the structure which could have a field strength of order ten times that needed to accelerate the ring. Detailed evaluation, especially of the dynamics of the ring, remains to be done.

These two examples represent applications of the idea of dragging one group of particles along in the fields associated with a second group. Other examples may found in refs. 16 and 17.

3f Category (2.3) in Table

The final category comprises schemes in which a deep potential well in an electron beam moves forward with a velocity suitable for accelerating positive particles trapped within it. One of the best known and most thoroughly analysed schemes of this class is the elctron ring accelerator proposed originally by Veksler and others in 1967.(18) (The idea had been forshadowed in an unpublished note by Harvie in 1951.(19)) Although originally forseen as a high energy accelerator, greater understanding of ring stablity has revealed limitations to its performance, which suggest that it would be more appropriate as a heavy ion accelerator at non-relativistic energies. It is described in detail in ref. 17.

Another important example in this category is the ionisation front accelerator. This relies on the intense fields associated with a 'virtual cathode'. Unneutralized electron beams above a certain limiting current (of the order of 10-50KA, depending on energy and configuration) cannot be propagated down a hollow metal tube; this is because the space charge sets up retarding fields which are strong enough to reflect the electrons. If, however, there is gas present, this becomes ionised by fast electron collisions. The secondary electrons produced in this process are repelled to the conducting wall, and ions remain to neutralize the space-charge of the beam. By this means the virtual cathode region can advance along the pipe, carrying in its intense field ions, which are thereby accelerated. This velocity of advance can be varied by controlling the ionisation process, for example by a timed succession of laser pulses.(16,17)

This type of scheme, has not yet been developed sufficiently for practical application. In any case it would seem to be more appropriate to producing very short but intense bursts of ions at non-relativistic energies, than providing high field acceleration of relativistic particles.

Indeed, the authors of the DOE study 'Collective Accelerators'[16] conclude that "the acceleration of charged particles to very high energies is an unlikely application of collective accelerators". This study dealt with the schemes mentioned earlier in this review, (and many others), with the important exceptions of the beat-wave accelerator and the wake field accelerator discussed in sections 3c and 3e above.

4. Conclusions

Many methods of accelerating particles have been studied during the past thirty years. There is still the possibility of new ideas, recent examples being the wake field and beat-wave accelerator concepts still to be evaluated. The limits to conventional accelerators, in category (1.1) in the table, have by now be fairly carefully studied. There is still scope for a more general analysis to reveal more clearly the essential parameters, and to delineate the parameter space. In this more general analysis the grating linac and conventional accelerators such as SLAC would fall in two small well-defined areas, the region in between being more or less unexplored territory. Any optimisation, of course, must take into account factors such as development costs of appropriate power generators if these do not exist.

The categories in the table represent only one way of classifying the types of accelerator which exist or have been proposed. There may be more illuminating ways of making this classification if attention is restricted to the regime appropriate to ultra-high energies.

Finally, a clear understanding of the constraints on reaching ultra-high energies should be obtained. It is important to be clear as to what is required in the way of luminosity, (and hence beam quality and mean current) as well as energy. The aim is not merely to stay on the Livingston curve, but to do better. The 'answer' if there is one, hardly seems yet in sight.

1. REFERENCES

1. Report of the Subpanel on Accelerator Research and Development of the High Energy Physics Advisory Panel, (Chairman M. Tigner), DOE report ER-0067 UC-34, Washington, June 1980.

2. Proc. Second ICFA Workshop on Possibilities and Limitations and Detectors, Les Diablerets, 4-10 Oct 1979. Also G. A. Voss, this meeting.

3. R.B. Palmer, this meeting and Particle Accelerators, 11, 81 (1980).

4. R. Pantell and T. I. Smith, Workshop on the Laser Acceleration of Particles, Los Alamos, Feb 1982.

5. C. Pellegrini, this meeting and Los Alamos Workshop, (as ref. 4).

6. M. Tigner and D. Keefe, presented paper and panel discussion respectively, this meeting.

7. R. Pantell, Los Alamos Workshop (as ref. 4) and Pellegrini, this meeting.

8. Ya. B. Fainberg, Soviet Physics Uspekhi, 10, 750 (1980).

9. M. L. Sloan and W. E. Drummond, Phys. Rev. Letts. 31, 1234 (1973).

10. T. Tajima and J. M. Dawson, Phys. Rev. Letts. 43, 267 (1979).

11. R. D. Ruth and A. W. Chao, Lawrence Berkeley Laboratory report LBL-14332 (1982).

12. A. Faltens and D. Keefe, Proc. Heavy Ion Fusion Workshop, Oct-Nov 1979, Lawrence Berkeley Lab Report LBL-10301, p.157 (1980).

13. G. Ising, Arkiv för Matematik och Fysik 18, 1 (1924).

14. H. Alfvén and O. Wernholm, Arkiv för Fysik 5, 175 (1952).

15. G. A. Voss, and T. Weiland, 'Particle Acceleration by Wake Fields', DESY internal report M-82-10, Hamburg, and this meeting.

16. 'Collective Accelerators, a study carried out for the US. DOE', Chairman F. T. Cole, DOE report FN-355 0102.000 (1981).

17. C. L. Olson and U. Schumacher, 'Collective Ion Acceleration' Springer Tracts in Modern Physics, Vol. 84, (1979).

18. V. I. Veksler et al., Proc. 6th Int. Conf. on Accelerators, Cambridge, Mass, 1967. Cambridge Electron Accelerator Lab. Report CEAL-2000, p. 289 (1968).

19. R. B. R. S-Harvie, 'A Proposed Proton Linear Accelerator', Harwell memo AERE G/M 87, 1951.

DISCUSSION

Winterberg. In your analysis of what is feasible you do not include non-linear fluid mechanics and the possibility of collisionless shock waves and similar phenomena.

Lawson. This would be included in the category in my table of non-harmonic accelerating fields with particles moving in a medium. The schemes actually listed are examples and not intended to be exhaustive.

Zotter. The diagram you showed for the energies reached by existing machines is in terms of equivalent rather than real energies. Is not this rather misleading?

Lawson. The curve is an extrapolation from Livingston's book published about twenty years ago. It was presented at the 1980 Accelerator Conference at Geneva. Now that we are in the highly relativistic regime it is the centre of mass energy which is significant for physics.