Topics in Grand Unified Theories:

i) The Naturalness Problem

ii) Monopoles and Fermion Number Violation

Thesis by

Balachandran Sathiapalan

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California.

1983

(Submitted May 24, 1983)

Acknowledgements

I thank my advisor, John Schwarz, for his constant encouragement, guidance and support.

I thank Theodore Tomaras for a very enlightening collaboration on which much of the material in the second part of this dissertation is based, for many useful discussions in general, and also for his forbearance in putting up with an occasional digression into history and linguistics.

I am greatly indebted to all my fellow graduate students, especially Neil Marcus, Olivier Martin, Ariel Caticha and Mohit Randeria for their unfailing support and friendship and for the physics I learnt from them through countless hours of discussions.

क्षुरस्य धारा निशिता दुरत्यया ।
दुर्गम् पथस्तत् कवयो वदन्ति ॥

(कठोपनिषद्)

The razor's edge is sharp and difficult to tread.

So is the path (to Knowledge) hard - the Seers say.

(Kátha-Upanishad)

Abstract

This dissertation consists of two parts. The first part contains a discussion of the 'fine-tuning' and 'naturalness' problems in grand unified theories. It is argued that, while it is impossible to solve these problems in conventional theories which contain scalars, supersymmetric theories that require no fine tuning can be constructed. In these theories the problem reduces to that of obtaining a light Higgs doublet at the tree level, without any unnatural adjustment of parameters. A realistic supersymmetric grand unified theory that has this feature is constructed. It is based on the gauge group SO(10). Supersymmetry is explicitly broken through terms of dimension two.

The second part is an analysis of the interaction of fermions with a non-Abelian ('t Hooft-Polyakov) monopole. Monopoles are invariably present in grand unified theories, and recent studies with massless isospin half fermions have shown that monopoles catalyse fermion number violation. We show that this phenomenon can be described in simple terms using the language of instanton physics. This description also permits a straightforward extension of previous results to arbitrary fermion representations. The importance of half-integer winding numbers is stressed. An explicit calculation is done in the case of isovector fermions.

Table of Contents

INTRODUCTION

One of the principal goals of physics since the beginning of the century has been the unification of all particle interactions. Today we are aware of four kinds of interactions that, at least at low energies, behave in qualitatively different ways. They are the gravitational, weak, electromagnetic and strong interactions. The first major step towards unification was taken when it was realized that the weak interactions could be described by a gauge theory very similar to electromagnetism and it was noticed by Weinberg and Salam (independently) that both interactions could be described by an $SU(2)\times U(1)$ gauge theory. Having a non simple gauge group and consequently two independent and unrelated coupling constants meant that it was not a true unification in the literal sense of the word. However it was a conceptual unification in that both interactions were now parts of one bigger gauge theory. Belief in gauge theories was further strengthened by a series of developments: The Weinberg-Salam theory was shown to be "renormalizable"- which meant that one could now make detailed and unambiguous calculations of scattering amplitudes and compare them with experiment. The strong interactions were also shown to be quite adequately described by a gauge theory based on the group SU(3). While calculations were difficult to do because of technical complications - infrared divergences - there was no problem of principle. The discovery of asymptotic freedom opened up the high-energy regions to calculations and even made some successful predictions such as the violation of scaling.

Having discovered a complete theory (at least for low energies) of the three basic interactions and being struck by the fact that they were all described by gauge theories based on compact semi-simple Lie groups, it was time for physicists to be more ambitious. It was shown that one could achieve a true unification of the three interactions by embedding $SU(3)\times SU(2)\times U(1)$

in the simple Lie group SU(5). Besidesthe aesthetic appeal of unification, it also automatically solved the longstanding problem of explaining the quantization of electric charge.

As more detailed calculations with the minimal SU(5) theory were done, it became clear that it could not be taken to be a complete theory, since some of its predictions, particularly fermion masses, were not quite right. The special successes of the SU(5) theory were that it predicted correctly $\sin^2\theta_W$ and $\frac{m_b}{m_\tau}$. Apart from these specific quantitative predictions, the hypothesis of grand unification also suggested solutions to the puzzle of the observed baryon - antibaryon asymmetry observed in the universe. Thus the idea of grand unification has been widely accepted though the difficulties with the SU(5) model prompted innumerable other models based on different gauge groups to be constructed.

To make further progress one needs to develop some criteria to enable one to select from a plethora of possible theories the few that are worthy of detailed analysis. We already have some abstract principles, namely, we believe in local renormalizable quantum field theory, at least up to energies comparable to the Planck mass, and we can insist on theories based on a gauge principle, since all known theories (including gravity) are such. A new additional requirement that can be imposed is that of naturalness. We require a theory to be more than just renormalizable, we require it to be natural. i.e., one should have natural explanations of any dimensionless number that cannot be said to be of O(1) (by any stretch of the imagination). Grand unified theories, as we explain later, abound in such unexplained small numbers. Imposing this requirement, in fact, would directly rule out all conventional grand unified theories. The Technicolour and Extended-technicolour theories were motivated by just such considerations. These have not been very successful phenomenologically, however. More attractive in this respect are supersymmetric theories. Supersymmetry is a symmetry that relates bosons with fermions and in the process effects many cancellations of divergences in quantum calculations. It affords a way to make grand unified theories more natural.

Another independent requirement one would want to impose on a theory is that it describe quantum gravity as well. This is a non trivial requirement, as evinced by the fact that gravity is nonrenormalizable and is also extremely singular at the quantum level as soon as it is coupled to ordinary matter. Remarkably enough, supersymmetry may come to the rescue here also. To describe gravity all we have to do is to require that supersymmetry be a local (gauge) symmetry. (This is also in keeping with our proclivity towards gauge principles.) Since the commutator of two supersymmetries generates a space-time translation, we automatically get a theory of gravity when it is made local. Supergravity theories (there are eight of them known as N=1 to N=8 supergravity) are all much less singular in their quantum properties. There is even the faint hope that N=8 supergravity might be completely finite to all orders, though it is suspected to be divergent at either three loops or seven loops. In any case, it is significant that these two superficially unrelated requirements, namely, that the theory describe gravity and that it be natural, seem to point towards the same physical principle - that of supersymmetry.We think there is strong motivation, then, for trying to construct supersymmetric grand unified theories that share the successes of ordinary grand unified theories and furthermore satisfy the requirement of naturalness. In the first part of this thesis, we describe an attempt to construct a realistic supersymmetric grand unified theory that satisfies certain naturalness criteria.

The ultimate test of a physical theory lies in comparing its predictions with what is observed in the real world. One of the most significant unverified predictions of grand unified theories is the existence of extended objects (solitons) that look like magnetic monopoles at large distances. They are, in fact, a serious problem in the old standard cosmology, since it has been argued that monopoles should have been produced in significant numbers in the early universe whereas none have been observed experimentally. In the inflationary universe scenario this problem does not arise since the number density of monopoles is sufficiently diluted during the inflationary phase. Still, it is important to be able to set precise experimental limits on the number density of monopoles in order to be able to compare with predictions of different cosmological models.

Recent studies on the properties of monopoles indicate that much more stringent limits can be placed on the number density of monopoles than was hitherto thought possible. This is due to the phenomenon of monopole catalysed nucleon decay. In the second part of this thesis we investigate this phenomenon in some detail. We consider an SU(2) gauge theory broken to U(1). We study the interaction of fermions with the monopoles in this theory. The monopoles that occur in grand unified theories are essentially the same as these, since, at least locally, they can be described by the SU(2) subgroup of the full gauge group. It is expected that these results will be applicable to the grand unified monopoles also - this is a subject of ongoing research.

Supersymmetric Grand Unified Theories

1. The Hierarchy, Naturalness and Fine-Tuning Problems:

In the real world there are various observed mass scales, which on the face of it seem to be completely unrelated to each other. In order of decreasing magnitude : the Planck mass ($M_p \approx 10^{19} GeV.$), the weak interaction scale ($M_W \approx 10^2 GeV.$, the strong interaction scale ($\Lambda_{QCD} \approx 10^2 MeV.$), the electron mass $m_e \approx 1 Mev$, the inverse radius of the universe $\approx 10^{-32} ev$ - to mention only the most striking ones. In addition to these scales grand unified theories introduce yet another unobserved scale M_X - the scale at which the strong, weak and electromagnetic coupling constants become equal, which is on the order of $10^{16} GeV.$. What is striking is - and this is essentially the statement of the hierarchy problem - the large ratios of some of these numbers to the others. There are extremely small dimensionless numbers like $\frac{M_W}{M_p} \approx 10^{-17} or \frac{M_W}{M_X} \approx 10^{-14}$ in our theories and no explanation or understanding of why they are so small. In the absence of a sensible (finite or renormalizable) quantum theory of gravity, it is perhaps premature to worry about the ratio $\frac{M_W}{M_p}$. However, the ratio $\frac{M_W}{M_X}$ in grand-unified theories still remains to be explained. The same holds for most of the other small numbers that occur.*

In addition to having two widely separated unexplained mass scales , grand- unified theories have other unpleasant features. In order to get the right low energy phenomenology, one has to 'fine-tune' parameters in the Lagrangian to an incredible degree of accuracy. This has been called the 'naturalness' or 'fine-tuning' problem. In its usual form, this involves choosing

---

* The ratio $\frac{\Lambda_{QCD}}{M_X}$ is usually 'explained' as the exponential of a not so small number.

two large mass parameters $M_A, M_B \approx M_X$ such that $M_A - M_B \approx M_W$. This implies a fine-tuning of the relative values of M_A and M_B to an accuracy of fourteen decimal places. It is plausible that a satisfactory physical theory should not be dependent on such delicate adjustments but rather have predictions that are not excessively sensitive to small changes in the bare parameters of the Lagrangian. In ordinary renormalizable field theories with scalars this problem also manifests itself at the quantum level, for example in corrections to physical quantities such as masses of the scalar particles. The corrections δm^2 are of $O(\Lambda^2)$ where Λ is the cutoff in the theory. This entails a subtraction to keep the physical mass at the experimentally observed value, i.e.

$$m^2{}_{phys} = m^2{}_{phys} + m^2{}_{c.t.} - \alpha\Lambda^2.$$

Here m_{phys} is the physical mass of the scalar particle and is taken to be the renormalized parameter in the Lagrangian. $m^2{}_{phys} + m^2{}_{c.t.}$ is the bare $(mass)^2$ (c.t. stands for counterterm) and is chosen to cancel the quantum corrections that are generated, which include the divergent piece $\alpha\Lambda^2$. If $m^2{}_{phys} \approx 100 GeV.^2$ and $\Lambda \approx M_p$ then we have a fine tuning of 34 decimal places! This has to be repeated at each order in perturbation theory. Thus, whether or not we have to tune things at the tree level (as in grand-unified theories), in all field theories with scalars we have to tune parameters at the quantum (loop) level. It should be emphasised that there is no inconsistency in any of these operations. However, one would like to restrict a physical theory to be more than just consistent, one requires it to be natural. We can phrase this requirement in the following way: a change $\delta\lambda$ in the bare parameters λ of the theory such that $\frac{\delta\lambda}{\lambda} \ll 1$ should lead to a change δP in the value of a physical quantity P , satisfying $\frac{\delta P}{P} \ll 1$. This requirement is imposed before any subtraction or any other form of renormalization is done , so a physical cutoff is kept in the theory. Stated in this form this requirement essentially forbids any form of fine tuning both at the classical and quantum level.

In the above description the problem of scales was divided into two separate problems. One was the difficult issue of how the scales were generated in the first place. The other was the question of how, given two widely separated scales, they are to be kept apart without any unnatural adjustments of parameters, either at the tree level or at the loop level. It is clear that any theory that claims to solve the first one (the hierarchy problem) must necessarily solve the second one also.

There have been some attempts to solve the hierarchy problem. Many of them are based on having no fundamental scalars in the theory but only composite ones , made out of fermions. These are the technicolour schemes [1]. This automatically solves the problem of quadratically divergent corrections to the masses of scalars (simply because there are no fundamental scalars*), and hence the fine tuning problem. The solution to the hierarchy problem then consists of generating a scale $\Lambda_{technicolour}$ in a manner analogous to the way Λ_{QCD} is generated . The full gauge group thus is of the form $G_{technicolour}$ X SU(3) X SU(2) X U(1). At this scale the technicolour interactions cause the techniquarks to confine and also break chiral symmetry, and the resulting technipions become the Higgs scalars necessary for the Weinberg- Salam mechanism. While extremely elegant in its conception this scheme suffers from a serious defect - there is no mechanism for generating quark and lepton masses. To get around this various modifications have been proposed like the extended technicolour schemes [2] but none of these have really been successful. The only other serious attempt to solve the hierarchy problem is the inverted hierarchy scheme [3]. However, we shall not describe it here. In this thesis we shall not have anything new to say about the hierarchy problem. We shall only discuss an attempt to solve the fine tuning problem . In the next section we describe field theories that have global supersymmetry and

* Fermion mass corrections can only be logarithmically divergent and a number like log $M_p / M_W \approx 20$ which is not unacceptable from the point of view of the fine tuning problem.

introduce N=1 superfield notation. In sec. 1.3 we discuss a supersymmetric grand unified theory that does require a fine tuning. We also motivate explicit supersymmetry breaking.

In sec. 1.4 a realistic SO(10) theory that does not require any fine tuning, is described and we conclude in sec 1.5 with some thoughts on linking up supersymmetric grand unified theories with higher N supersymmetry and supergravity theories.

2. Supersymmetry

Supersymmetry is a symmetry that relates particles of different spin with each other. There exists a huge mass of literature [4] on field theories with global or local supersymmetry, so we shall not discuss these theories in detail here. However we shall mention a few salient features that make these theories so interesting even in the absence of any experimental evidence for supersymmetry. First, the unification of particles of different spins into one supermultiplet is aesthetically very attractive. Second, the supersymmetry algebra can include both internal symmetry and space time symmetry generators. The commutator of two supersymmetry transformations gives a spacetime translation.* Thus when supersymmetry is made local we naturally obtain a theory of gravity. Since the algebra contains internal symmetry generators also, there is for the first time , the possibility of a real unification of gravity with the other three interactions. Finally, and perhaps most importantly, theories with super symmetry, local or global are invariably less singular in their quantum properties than corresponding non supersymmetric theories. As an extreme example N=4 supersymmetric Yang-Mills theory has been proven to be finite to all orders of perturbation theory. The N=8 supergravity theory is known to be one-loop finite even though it contains a rich spectrum of particles besides the graviton. Ordinary gravity on the other hand is extremely singular as soon as you include any 'matter' particles (particles of spin <2). The N=1 supersymmetric theories also have much better quantum properties than ordinary field theories. They obey certain 'no-renormalization' [5] properties. These have the consequence that, apart from overall wave function renormalizations, only supersymmetric expressions which can be written in the form $\int d^4\theta(\ldots)$ are generated by graphs in any order of perturbation theory. This implies , then, that masses and Yukawa couplings are not renormalized. This property makes supersymmetry very

*The supersymmetry algebra is described in Appendix A useful when one tries to solve the fine tuning problem. In non-supersymmetric grand-unified theories with scalars there is no hope of solving either the hierarchy or fine-tuning problems.

We proceed to describe the construction and properties of renormalizable field theories that are supersymmetric. Only theories with one supersymmetry are described here.

In Appendix A the abstract supersymmetry algebra with N supersymmetry generators (N < 4 for reasons explained there), has been written down and the field content of the supermultiplets that form irreducible representations of this algebra is also given. We reproduce here for convenience the relevant part for N=1 supersymmetry. The algebra is

$$\{Q_\alpha, Q_\beta\} = \{\bar{Q}_{\dot{\alpha}}, \bar{Q}_{\dot{\beta}}\} = 0$$

$$\{Q_\alpha, \bar{Q}_{\dot{\beta}}\} = 2\sigma^\mu_{\alpha\dot{\beta}} P_\mu$$

$$[Q_\alpha, P_\mu] = [\bar{Q}_{\dot{\alpha}}, P_\mu] = 0$$

$$[Q_\alpha, M_{\mu\nu}] = i(\sigma_{\mu\nu} Q)_\alpha$$

and the irreducible representations are (0 , 1/2) - scalar multiplet which consists of one complex scalar and one Majorana (or Weyl) spinor. (1/2 , 1)- vector multiplet with one Majorana (or Weyl) spinor and one real vector. In order to write down Lagrangians that have invariance under supersymmetry we have to represent the abstract algebra on fields. The first such representation found was :

$$\delta A = i\bar{\alpha}\psi$$

$$\delta B = i\bar{\alpha}\gamma_5\psi$$

$$\delta\psi = \partial_\mu(A - \gamma_5 B)\gamma^\mu\alpha + (F + \gamma_5 B)\alpha$$

$$\delta F = i\bar{\alpha}\gamma^\mu\partial_\mu\alpha$$

$$\delta G = i\overline{\alpha}\gamma_5\gamma^\mu\partial_\mu\psi \quad .$$

A and F are scalars , B and G are pseudoscalars and ψ is a Majorana spinor. One can verify that $(\delta_1\delta_2 - \delta_2\delta_1)A = -2i\overline{\alpha}_1\gamma^\mu\alpha_2\partial_\mu A$. Thus the commutator of two supersymmetry transformations gives a space-time translation as required by the abstract algebra. Furthermore, using the equations of motion

$$(\Box - m^2)A = (\Box - m^2)B = (\partial\!\!\!/ + m)\psi = 0 \quad ,$$

one can verify that the combinations (mA + F) and (mB + G) are invariant under supersymmetry. Thus one can take F = -mA and G = -mB without spoiling the supersymmetry. Thus the independent degrees of freedom are A, B and ψ. This constitutes the (0, 1/2) representation mentioned earlier. F and G are auxiliary fields and can be eliminated using their algebraic equations of motion. It is possible to describe all the fields (A, B, ψ , F, G) as the components of one superfield - known as the chiral superfield, as follows:

$$\Phi(x,\theta) = \varphi(x) + \theta^\alpha\psi_\alpha + \theta^2 z + i\theta^\alpha\overline{\theta}^{\dot\beta}\partial_{\alpha\dot\beta}\varphi + i/2\theta^2\overline{\theta}^{\dot\beta}\partial_{\alpha\dot\beta}\psi^\alpha(x) - 1/4\theta^2\overline{\theta}^2\Box\varphi$$

where $\varphi = A - iB$, $z = \frac{F + iG}{\sqrt{2}}$, θ^α is a two component Weyl spinor. $\overline{\theta}_{\dot\alpha}$ is its complex conjugate with the opposite chirality. It is a Grassmann (anticommuting) number i.e. it satisfies $\{\theta_\alpha, \theta_\beta\} = 0$, $\int d\theta = 0$ and $\int d\theta\theta = 1$. x^μ, θ and $\overline{\theta}$ constitute the coordinates of a superspace. If θ is assumed to have negative chirality then ψ describes a left handed spinor.

To write down a Lagrangian in terms of Φ it is necessary, first, to write down the kinetic terms. In components the supersymmetric kinetic term is :

$$\frac{-1}{2}(\partial_\mu A)^2 - \frac{1}{2}(\partial_\mu B)^2 - \frac{1}{2}i\overline{\psi}\partial\!\!\!/\psi + \frac{1}{2}F^2 + \frac{1}{2}G^2.$$

In terms of superfields it is $\int d^4\theta\overline{\Phi}\Phi$. A mass term is easily written down in terms of chiral superfields:

$$\frac{1}{2}\int d^2\theta m \Phi^2 + h.c. = \sqrt{2}m(FA + GB) - \frac{m}{2}i\overline{\psi}\psi \ .$$

A cubic interaction term would be

$$\frac{1}{3}\int d^2\theta\Phi^3 + h.c. = \sqrt{2}(FA^2 - FB^2 + 2GAB) - i\bar{\psi}\psi A + i\bar{\psi}\gamma_5\psi B$$

A word about the dimensionality of the superfields : Φ has the dimensions of mass since it starts off as $\varphi = \frac{1}{2}(A - iB)$ and A and B are canonical scalar fields. Furthermore ψ has a canonical dimension of 3/2 which implies that θ has dimension -1/2. From $\int d\theta\theta = 1$ we conclude that $d\theta$ has dimension +1/2. Thus $\int d^2\theta\Phi^3$ is a dimension four term, as it should be. The others can be checked similarly. We can now write down a general Lagrangian in the form

$$\int d^4\theta\bar{\Phi}\Phi + \int d^2\theta\, W(\Phi) + h.c.$$

The Lagrangian is manifestly supersymmetric, being written in terms of superfields and integrated over all θ. $W(\Phi)$ is called the superpotential for reasons that will become clear and has the generic form $m\Phi^2 + g\Phi^3$. To get the component level Lagrangian we have to eliminate the auxiliary fields F and G using their equations of motion. If, for e.g., $W(\Phi) = m\Phi^2 + \frac{4}{3}g\Phi^3$ then:

$$F + mA + g(A^2 - B^2) = 0$$

$$G + mB + 2gAB = 0$$

Substituting these expressions for F and G into the mass, kinetic and interaction terms the full scalar potential is obtained. A very convenient shorthand way of writing this potential is:

$$V(A,B) = F^2 + G^2$$

where it is understood that for F and G one is to substitute their equations of motion. This tells us two things right away : i) $V(A,B) \geq 0$ in supersymmetric theories. ii) If F=G=0 at the minimum, i.e. if supersymmetry is unbroken, then $V(A,B)_{\min} = 0$. Also, if we can find a solution to the equations F=G=0, we have automatically minimised the potential and also verified that supersymmetry is unbroken at the minimum.

Let us denote F-iG by z and F+iG by z^*. Then we require $z = z^* = 0$. The equation $z = 0$ implies (using the equation of motion of z):

$$z^* = \int \frac{\partial W}{\partial z^*} d^2\theta = 0.$$

where $W = m\Phi^2 + \frac{4}{3} g \Phi^3$. But

$$\int d^2\theta \frac{\partial W[\Phi]}{\partial z} = \int d^2\theta \frac{\partial W[\Phi]}{\partial \Phi} \cdot \frac{\partial \Phi}{\partial z}$$

Since $\int d^2\theta \frac{\partial \Phi}{\partial z} = 1$, we get $\frac{\partial W}{\partial \Phi}|_{\theta=0} = 0$. It is easier to say this in words: write down the superpotential and replace the superfields by their scalar components and then minimise this 'potential'. All this assumes that the equations F=G=0 have a solution. If they do not, we cannot use this short-cut. We have to minimise the full potential $V(A,B)=F^2(A,B) + G^2(A,B)$. Furthermore in that case supersymmetry is spontaneously broken.

We now turn to the vector superfield: $V(x,\theta,\bar{\theta})$ can be expanded as (in the Wess-Zumino gauge)

$$V^{\alpha\dot{\beta}}\theta_\alpha\bar{\theta}_{\dot{\beta}} + i\theta^2\bar{\theta}^{\dot{\alpha}}\lambda_{\dot{\alpha}} - i\bar{\theta}^2\theta^\alpha\lambda_\alpha + \frac{1}{2}\theta^2\bar{\theta}^2 D$$

Internal symmetry indices have been suppressed. We do not bother to write down the supersymmetry transformation laws of the components fields. The gauge invariant supersymmetric kinetic term is

$$\frac{-1}{4}F_{\mu\nu}^2 - \frac{1}{2}i\bar{\lambda}\partial\lambda + \frac{1}{2}D^2$$

which shows that D is an auxiliary field. We can now write down a gauge invariant interaction between the chiral superfields and the vector superfield:

$$\int d^4\theta \bar{\Phi} e^{2eV} \Phi$$

On expanding into components, in addition to the usual covariantising of derivatives we get an additional term $\varphi^i D^a (t^a)_i^j \varphi_j^*$ where t^a are the Hermitian generators of the group. Thus the equation of motion for D becomes:

$$D^a = \varphi^i (t^a)_i^j \varphi_j^*$$

The potential of the scalar fields V(A,B) becomes now

$$V(A,B) = F^2(A,B) + G^2(A,B) + D^2(A,B)$$

To minimise this we have to require F=G=D=0. The extra condition D (A,B)=0 is usually not very difficult to satisfy. For instance we can write down the implications of requiring D=0 for some representations of SO(10) that will be useful in a later section.

10 : $A^i (t^a)_i^j B_j - B^i (t^a)_i^j A_j = 0$

for all values of α. Either A=0 or B=0 or A B.

45,54 : $[A,B]^{[ij]} = 0$

16,$\overline{16}$: $\chi_1^i (t^a)_i^j \chi_{1j}^* - \chi_2^i (t^{*a})_i^j \chi_{2j}^* = 0$

If we assume that $<\chi_1^T>=<\chi_2^T> = (0,0,\ldots ,\ldots ,1)$ as usually required for SO(10), then each term is individually zero unless t^a is diagonal. Then however $t^a = t^{a*}$ and so a necessary and sufficient condition for the vanishing of the D term is $<\chi_1> = <\chi_2>$.

Having described the N=1 supersymmetric Lagrangian and the tree level properties of the superpotential we move on to the quantum properties of these theories.[5] The quantum calculations can be done in a manifestly supersymmetric way using the technique of supergraphs. A result due to Grisaru, Rocek and Siegel is that to any order in perturbation theory the effective action can be written as an expression local in θ :

$$\Gamma(\varphi,\overline{\varphi},V) = \int d^4x d^4\theta \, G(\varphi,\overline{\varphi},V,\ldots;x_i,\theta).$$

Power counting leads to the following result for the degree of divergence of a graph: d = 2 - E - M (-1 if the graph has only chiral or only anti-chiral external lines) where E is the number of external chiral (or anti chiral) lines and M is the number of mass insertions in the internal lines. This formula is easy to explain : $d^4\theta$ takes up two powers of mass, each external chiral line takes up one power of mass and so do the mass insertions. Finally, if all the external lines are either chiral or anti-chiral then one needs either a D^2 or a $\overline{D}^2$ to get a non zero $d^4\theta$ integration.

The most important point is that, as mentioned before, neither the masses nor the Yukawa coupling constants get renormalized except through wave function renormalizations. This is because they are all of the form $\int d^2\theta\, W(\Phi)$ and not $\int d^4\theta$. Thus we are allowed to leave out mass terms or Yukawa terms, even if there is no global or local symmetry forbidding them, without compromising renormalizability. We will have occasion to comment on this later.

From the phenomenological point of view the N=1 theories with supersymmetry broken softly can be made consistent with all the low-energy experiments. At higher energies they are expected to show significant differences from ordinary grand-unified theories, not least because they predict a host of new particles that are not necessarily superheavy. The higher N supersymmetric theories have a basic problem in that they are all left-right symmetric (for the spin 1/2 particles) contrary to what is observed in the real world. This is because there are two supersymmetry generators and in turn they relate a helicity +1/2 to a helicity 0 particle and the helicity zero particle to a helicity -1/2 particle all in the same representation of the gauge group (because the supersymmetry generator is a gauge singlet). The problem is less severe in the gauged supergravity theories where the supersymmetry generators themselves carry internal symmetry indices. In that case the left-handed and right-handed fermions, related by supersymmetry, could have different internal quantum numbers such as electric charge or colour.

We restrict ourselves in this thesis to N=1 supersymmetry except for a brief discussion of the N=2 case in the Appendices. We shall describe the supersymmetrised version [7] of the SU(5) grand-unified theory [8] and use that opportunity to illustrate the fine tuning problem. We shall then describe in detail an SO(10) model which does not have this problem.

3. A Supersymmetric Model Illustrating the Fine-Tuning Problem.

Let us look at the following SU(5) model, which illustrates the fine tuning problem:

$$W = \lambda_1[\frac{1}{3}Tr\Sigma^3 + \frac{m}{2}Tr\Sigma^2] + \lambda_2 H'_x[\Sigma^x_y + 3m'\delta^x_y]H^y + Matter \quad coupling.$$

W is the superpotential, i.e. the Lagrangian is $\int d^4x d^2\theta\, W + h.c.$ + kinetic terms. Σ is a chiral superfield transforming as a 24 under SU(5). H and H' are also chiral superfields that transform as a 5 and $\overline{5}$ respectively. Σ contains in it the adjoint 24 needed to break SU(5) down to SU(3)XSU(2)XU(1). H and H' are responsible for breaking SU(2)XU(1) down to U(1). Unlike the case of ordinary SU(5) the super symmetric theory does not allow the complex conjugate of the 5 in place of a $\overline{5}$, because complex conjugation changes the chirality of the fermions in the superfield. The chirality of the superfield is important, because the coupling $\int d^2\theta \varphi_1\varphi_2\varphi_2$ contains Yukawa couplings between the fermions of each superfield. Thus we need a separate left-handed $\overline{5}$. To minimise the potential we have to solve the equations F=G=0. These are:

$$\frac{\partial W}{\partial \Sigma^y_x} = 0 \quad => \quad \lambda_1(\Sigma^2 - \frac{1}{5}Tr\Sigma^2)^y_x + \lambda_1 m\Sigma^y_x + \lambda_2 H'_x H^y = 0 \tag{1}$$

$$\frac{\partial W}{\partial H'_x} = 0 \quad => \quad \lambda_2(\Sigma^x_y + 3m'\delta^x_y)H^y = 0$$

$$\frac{\partial W}{\partial H_y} = 0 \quad => \quad \lambda_2 H'_x(\Sigma^x_y + 3m'\delta^x_y) = 0$$

We are using some shorthand notation : in the equations, Σ, H, and H' stand for the (non auxiliary) A+iB components of the superfields Σ, H and H' respectively. There are three possible solutions to these equations corresponding to three vacua each with a different symmetry and all of them degenerate in energy. One of them has $\langle H\rangle = \langle H'\rangle = \langle \Sigma \rangle = 0$ and SU(5) is unbroken . Another has SU(5) broken to SU(4)XU(1). The third one, which is the one we shall study, has SU(5) broken to SU(3)XSU(2)XU(1).

$$\langle\Sigma\rangle = \text{Diag.}(3,3,-2,-2,-2)$$

Consider the term $\lambda_2 H' (\Sigma + 3m'\delta)H$. On substituting for the scalar part of Σ, the vacuum expectation value (vev) shown above, it becomes clear that the colour triplet Higgs gets a $(mass)^2 \propto (2m+3m')^2$ whereas the doublet which is the Weinberg-Salam Higgs gets a $(mass)^2 \propto (3m-3m')^2$. Since the doublet has to get a vev at lower energies it has to be kept light. The colour triplet on the other hand, has to be made superheavy in order to prevent proton decay from occurring with too high a rate. Thus we must impose m - m' $\approx O(M_W)$. m is of order $10^{15} GeV$. since it characterizes the SU(5) breaking scale and if we choose m' also of the same order the colour triplets become superheavy. We have a situation, then, where m, m' $\approx O(10^{15} GeV.)$, but m-m' $\approx O(10^2 GeV.)$. Let us apply our naturalness criterion to this case : we require for $\frac{\delta\lambda}{\lambda} \ll 1$ where λ are the parameters of the theory $\frac{\delta P}{P} \ll 1$ where P is any physical quantity. In the above situation we can vary m and m' independently since no symmetry relates them, so we let $\frac{\delta m}{m} = 10^{-3}$ (say) and let $\delta m' = 0$. Then we find since $\delta m = 10^{-3} \times 10^{15} = 10^{12} GeV.$ the Higgs mass becomes $O(10^{12} GeV.)$. Thus $\frac{\delta P}{P}$ is certainly not much less than 1 (It can vary between 1 and 10^9 depending on whether one takes the initial P or the final P in the denominator). Thus it looks like a case of fine tuning. However, before we come to such a conclusion, we must be a bit more careful. In supersymmetric theories the quantum corrections are very mild - only wave function renormalization for the Yukawa and mass terms. In the above example it is easy to see that both m and m' are renormalized by the same factor - the wave function renormalization of the Σ field. In the perturbative trivial vacuum where $<\Sigma> = 0$ and $< H > = < H' > = 0$, m and m' have a physical significance. They are the masses of the Σ field and H field respectively. However, it is conceivable that in the new vacuum that has only SU(3)XSU(2)XU(1) symmetry one might be able to define new parameters that have independent physical significance and in terms of which there is no fine tuning. Let us then try to define m-m' $= O(M_W)$ as the physical parameter. Expanding around the new vacuum, we have some fields of mass $O(M_W)$, others of mass $O(m)$, and there appears to be no fine tuning required, since we have defined these to be the parameters of the theory. However, this soon leads (not unexpectedly) to trouble. The Higgs triplet now has a mass that is some complicated function of m and M_{Ψ}. The underlying SU(5) symmetry requires that it have that precise value for its mass. But if we insist on viewing the theory as one whose parameters are m and M_{Ψ} and ignore the original structure, then the Higgs triplet mass has no particular reason to have that precise value. Thus the fine tuning is still present, showing itself in a different place, and this redefinition of parameters has not improved the situation. The moral of this rather long-winded discussion is that while at first sight there appears to be a certain amount of arbitrariness in our interpretation of the fine-tuning problem because of ambiguities in definitions of what the bare parameters of the theory are, in fact, the problem is well defined.

The basic problem, then, in supersymmetric grand-unified theories is to obtain at the tree level a "naturally" light Higgs doublet. Supersymmetry ensures that the quantum corrections do not reintroduce fine tuning.

A realistic theory cannot be supersymmetric. All the particles in a supermultiplet have the same mass because $P_\mu P^\mu$ is a Casimir operator of the supersymmetry algebra. This degeneracy is certainly not observed in the real world, and therefore supersymmetry has to be broken. The breaking can either be explicit or spontaneous. There are two reasons why explicit breaking is preferred. First, global symmetries are unnatural and so one prefers to associate the supermultiplet structure of particles with an underlying local supersymmetry (supergravity). There are also independent reasons that make supergravity attractive. As mentioned earlier, they hold out a promise of providing a sensible quantum theory of gravity and at the same time unifying it with the other three interactions. At low energies, the effective theory with $\kappa = 0$ derived from this supergravity theory, will have a global supersymmetry. Furthermore, if the local supersymmetry is spontaneously broken at some scale, then in the low energy ($k = 0$) theory there are terms that explicitly break this global supersymmetry. This provides a rationale for introducing explicit supersymmetry breaking. Second, spontaneously broken global supersymmetry would require the existence of a massless spin 1/2 particle sometimes called the 'goldstino'. It is not observed experimentally. Realistic models that have global supersymmetry spontaneously broken are extremely contrived and invariably require the addition of a large number of extra fields that have the sole function of breaking supersymmetry spontaneously and hiding the resultant goldstino from the eyes of the experimentalist. If the theory has underlying local supersymmetry, then the goldstino is eaten up by the gravitino which becomes massive via a super-Higgs effect, avoiding experimental conflict. The explicit supersymmetry breaking terms should be 'soft', i.e., they should not reintroduce quadratic divergences into the theory. The allowed terms for both N=1 and N=2 theories have been classified [9]. The result for the N=1 can be summarised as follows: One can add any dimension-two term (i.e. mass term for the scalars), but the only dimension-three terms allowed are mass terms for the gauginos (superpartners of the gauge bosons) and a particular cubic scalar coupling.* We now have all the information we need to construct a specific model.

*See Appendix B for details.

4. A Supersymmetric SO(10) Model with No Fine Tuning.

We now describe in detail a realistic model that has the feature that it requires no fine tuning of its parameters to obtain phenomenologically acceptable predictions. To be precise for any change $\delta\lambda$: $\frac{\delta\lambda}{\lambda} \ll 1$ of its parameters the predictions 'P' satisfy $\frac{\delta P}{P} \ll 1$. We emphasize that no attempt is made to understand the existence of widely separated scales (the hierarchy problem). The scales are put in by hand, i.e., we start with a Lagrangian that contains both large and small mass parameters.

The gauge group is SO(10). The superfield content is the following: i)the gauge vector multiplet V (45). ii)chiral superfields M_a (16), one for each family, 'a' being a family index, iii)chiral superfields $\Phi(54)$,Σ andΣ_1 (*both* 45),$\chi_1(16)$ and $\chi_2(\overline{16})$, iv) three chiral superfields H, H' and M (all 10's). The 45's are written as 10X10 antisymmetric matrices Σ^{ij} , Σ_1^{ij}, where i,j are vector indices and take values from 0 to 9. The 45 can also be written in the spinor representation of SO(10) as a 16x16 matrix,i.e., $\Sigma_a^b = (\sigma^{ij})_a^b \Sigma^{ij}$, where a,b are spinor indices that run from 1 to 16. σ^{ij} are the SO(10) generators represented by 16x16 matrices. They are the Clebsch Gordan coefficients that pick out the 45 from 16×$\overline{16}$. We use the conventions of ref. [13], reproduced in Appendix C. The 54 is a traceless symmetric 10x10 matrix.

Before proceeding further, we would like to motivate this choice of representations. The gauge multiplet V and matter M_a need no motivation. The 54, 45, 16, and $\overline{16}$ are all needed to break SO(10)-> SU(3)XSU(2)XU(1). To see this let us write down the most general superpotential consistent with SO(10) invariance*:

$$W = \frac{M_1}{2} Tr\,\Phi^2 + \frac{\lambda_1}{3} Tr\,\Phi^3 + \frac{M_2}{2} Tr\,\Sigma^2 + \lambda_2 Tr\,\Sigma\Phi\Sigma + g\,\chi_1\Sigma\chi_2 + m\,\chi_1\chi_2. \qquad (2)$$

*Note that there are no Σ^3 terms because Σ is antisymmetric in its indices. Also neither 16×16 nor 16×$\overline{16}$ contains a 54.

Matrix notation is being used here. Thus, for instance, $\chi_1 \Sigma \chi_2$ stands for $\chi_1^a(\Sigma)_a^b \chi_{2b}$. The equations for the minimum are:

$$\frac{\partial W}{\partial \Phi^{ij}} = 0 \;=>\; [M_1\Phi + \lambda_1(\Phi^2 - \frac{1}{10}Tr\,\Phi^2) + \lambda_2(\Sigma^2 - \frac{1}{10}Tr\,\Sigma^2)]^{ij} = 0 \tag{3.1}$$

$$\frac{\partial W}{\partial \Sigma^{ij}} = 0 \;=>\; M_2\Sigma^{ij} + \lambda_2(\Sigma\Phi + \Phi\Sigma)^{ij} + g\chi_1^a(\sigma^{ij})_a^b\chi_{2b} = 0 \tag{3.2}$$

$$\frac{\partial W}{\partial \chi_1^a} = 0 \;=>\; m\chi_{2a} + g(\Sigma)_a^b\chi_{2b} = 0 \tag{3.3}$$

We try the following ansatz for the vev that gives us the required SU(3)XSU(2)X U(1) symmetry:

$$\langle\Phi\rangle = v\begin{bmatrix} 3 &&&&&&&&& \\ & 3 &&&&&& 0 && \\ && 3 &&&&&&& \\ &&& 3 &&&&&& \\ &&&& -2 &&&&& \\ &&&&& -2 &&&& \\ && 0 &&&& -2 &&& \\ &&&&&&& -2 && \\ &&&&&&&& -2 & \\ &&&&&&&&& -2 \end{bmatrix} \tag{4}$$

$$\langle\Sigma\rangle = \sigma\begin{bmatrix} &&& a &&&&&& \\ && a &&&& 0 &&& \\ & -a &&&&&&&& \\ -a &&&&&&&&& \\ &&&&& b &&&& \\ &&&& -b &&&&& \\ &&&&&&&&& b \\ && 0 &&&&&& b & \\ &&&&&&& -b && \\ &&&&&& -b &&& \end{bmatrix}$$

$= a(\Sigma^{03} + \Sigma^{12}) + b(\Sigma^{69} + \Sigma^{78} + \Sigma^{45})$

$$\chi_1 = x\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \chi_2$$

Using these one can check that all the equations are satisfied provided the following scalar equations are satisfied :

$$M_1 v - \lambda_1 v^2 + \frac{\lambda_2}{5}(b^2 - a^2)\sigma = 0 \tag{5.1}$$

$$a\sigma(M_2 + 6\lambda_2 v) = gx^2 \tag{5.2}$$

$$b\sigma(M_2 - 4\lambda_2 v) = gx^2$$

$$m + g(3b + 2a)\sigma = 0 \tag{5.3}$$

It is easy to convince oneself that none of the representations (45, 54, 16, $\overline{16}$) are redundant if one is to have solutions where the final symmetry is broken down to SU(3)XSU(2)XU(1). For instance, without the 54 one could have a solution where b=a, but this would have a residual SU(5) symmetry. The other point to note is that if $m \approx M_1 \approx M_2 \approx 10^{16} GeV.$, then $v, \sigma, x \approx 10^{16} GeV.$, so SO(10) -->SU(3)XSU(2)XU(1)at the superheavy scale M_X directly, without any intermediate scales. Intermediate scales, as in SO(10) -> $SU(3)XSU(2)_L XSU(2)_R XU(1) \dashrightarrow SU(3)XSU(2)_L XU(1)$, usually result in an unacceptably large value of $\sin^2\theta_W$[11], and are best avoided.

As mentioned before, the colour triplets contained in the 10 of Higgs, mediate proton decay and therefore have to be made superheavy. At the same time the doublets have to be kept light, because at low energies they have to develop a vev and break $SU(2)_L XU(1)$ down to U(1). A possible way of doing this was first suggested by Dimopoulos and Wilczek [10]. The idea is as follows: If the 45 of SO(10) is given a vev of the form:

$$\begin{bmatrix} 0 & & & 0 & & \\ & & & & 1 & \\ & & & -1 & & \\ 0 & & & & & 1 \\ & & & & -1 & \\ & & & & & \ddots \end{bmatrix}$$

then one can have the coupling 10 x 45 x 10 and the first four components of the 10 (the complex Higgs doublet) would be light and the last six (=3+ $\overline{3}$ of SU(3)) would get a mass of the order of the vev of the 45. Thus the doublets and the triplets are automatically split. This would solve the fine tuning problem. A possible way to implement this in the minimal model described above (with 45, 54, 16, $\overline{16}$) would be to let <16>=< $\overline{16}$ >=0, i.e. x=0 in equation (5.2) and also let 'a' (one of the parameters in equation (4)) = 0. This is clearly one possible solution. However the residual symmetry is now SO(4) XSU(3)XU(1). The SO(4) can be written as $SU(2)_L XSU(2)_R$, where $SU(2)_L$ can be identified with the $SU(2)_L$ of the Weinberg-Salam model. One can, now, at a much lower scale, give a vev to the 16 (by adding some explicit supersymmetry breaking terms) and break $SU(2)_R$. But this also modifies the vev of the 45 and makes the parameter 'a' of eq. (4) non zero. This means that the Weinberg-Salam doublet gets a mass of the same order as the scale of $SU(2)_R$ breaking. This scale therefore cannot be much higher than $10^3 GeV.$. An intermediate scale as low as $10^3 GeV.$ is phenomenologically unacceptable because of problems with $\sin^2\theta_W$ mentioned earlier. The conclusion then is that in the minimal model with only one 45 it is not possible to solve the fine-tuning problem.

We can make use of the Dimopoulos-Wilczek scheme by expanding our minimal set of fields to include another 45 called Σ_1. In that case one of the 45's can be used to give a mass to the triplet Higgs while the other 45 can be used to produce the symmetry breaking pattern SO(10) -> SU(3)XSU(2)XU(1).

The superpotential in (2) is modified by the addition of some terms involving the new 45. It reads now

$$W = [\frac{M_1}{2} Tr\Phi^2 + \frac{\lambda_1}{3} Tr\Phi^3 + \frac{M_2}{2} Tr\Sigma^2 + \lambda_2 Tr\Sigma\Phi\Sigma + m\chi_1\chi_2 + g\chi_1\Sigma\chi_2]$$
$$+ (\frac{M_3}{2} Tr\Sigma_1^2 + \lambda_3 Tr\Sigma_1\Phi\Sigma_1 + fH\Sigma_1 H') + \mu H^2 + \mu' H'^2 + f_{ab} M_a \Gamma M_b H$$

Here M_1, M_2, M_3 and $m \approx O(10^{16} GeV.)$ and $\mu, \mu' \approx O(10^2 GeV.)$. The terms in the square brackets are the same as in the minimal model. The new terms are those involving the field Σ_1 (the second 45) in the curved brackets. We need two 10's (H and H') because Σ_1^{ij} is antisymmetric in its two indices and a term like $H\Sigma_1 H$ would vanish identically. This term will be responsible for splitting the colour triplets from the doublets by giving the triplets a large mass, while keeping the doublets light. The last term is the Yukawa coupling of the Higgs to matter fields, and 'Γ' stands for the Clebsch-Gordan coefficients for the 10 contained in 16X16 (see Appendix C).

It should be noted at this point that this superpotential is not the most general that can be written down consistent with supersymmetry and SO(10) invariance. For example the term $\chi_1\Sigma_1\chi_2$ has not been included. In general, in supersymmetric theories one is allowed to do this because of the no-renormalization theorems of supersymmetry [5]. However, as earlier pointed out one would like more than mere consistency, one wants naturalness and it is not enough to invoke the no renormalization theorems. The question of whether or not setting a parameter equal to zero is a violation of naturalness requires some consideration. We had, earlier, defined a theory to be natural if it satisfies $\frac{\delta P}{P} \ll 1$ for $\frac{\delta\lambda}{\lambda} \ll 1$ where P is a physical property (e.g. the mass of the W) and λ is a bare parameter. Choose λ to be the coefficient of a term that has not been included in the Lagrangian, i.e. $\lambda = 0$. Changing λ from 0 to $\delta\lambda > 0$ is not a small change since $\frac{\delta\lambda}{\lambda}$ is not a small number however small $\delta\lambda$ might be. This is a reflection of the fact that the theory is qualitatively modified when λ is changed from zero. So if a bare parameter is zero, then in testing the naturalness of a theory, this parameter should not be varied at all. Only the non zero parameters should be varied. Thus setting bare parameters to zero in supersymmetric theories does not by itself make the theory unnatural. We think this is sufficient justification for omitting some terms from the Lagrangian. However for the sceptical reader this is justified separately on the grounds that the Lagrangian that has been written down is the most general consistent with supersymmetry, SO(10) and the following discrete symmetries:

$$\Sigma_1 \to -\Sigma_1, \quad H' \to -H' \tag{6a}$$

$$\chi_1 \to -\chi_1, \quad \chi_2 \to -\chi_2 \tag{6b}$$

$$H \to -H, H' \to -H', M_a \to e^{i\frac{\pi}{2}} M_a \tag{6c}$$

The term $\chi_1\Sigma_1\chi_2$ for example is ruled out by (a). The complete list of terms otherwise allowed but ruled out by these symmetries is given below: (The letter in paranthesis refers to the particular symmetry that rules out the term)

$$Tr\,\Sigma\Sigma_1(a)$$

$$Tr\,\Sigma\Phi\Sigma_1(a)$$

$$M_a\chi_2(b)$$

$$\chi_1\Sigma_1\chi_2(a), \quad M_a\Sigma_1\chi_2(a), \quad M_a\Sigma\chi_2(b)$$

$$HH'(a)$$

$$H\Sigma H'(a)$$

$$M_a\Gamma M_b H'(a), \quad \chi_1\Gamma\chi_1 H'(a), \quad \chi_1\Gamma M_a H'(a), \quad \chi_2\Gamma\chi_2 H'(a),$$

$$\chi_1\Gamma\chi_1 H(c), \quad \chi_1\Gamma M_a H(b,c), \quad \chi_2\Gamma\chi_2 H(c)$$

The analysis of the minimum of the potential proceeds as follows: Setting the auxiliary fields to zero gives : (matter fields have been put equal to zero in equations 7.1 to 7.7)

$$D_i^j = 0 \tag{7}$$

$$F_\Phi = 0 \ : \ M_1\Phi + \lambda_1(\Phi^2 - \frac{1}{10}Tr\,\Phi^2) + \lambda_2(\Sigma^2 - \frac{1}{10}Tr\,\Sigma^2) + \tag{7.1}$$

$$+ \lambda_3(\Sigma_1^2 - \frac{1}{10}Tr\,\Sigma_1^2) = 0$$

$$F_\Sigma = 0 \ : \ M_2\Sigma^{ij} + \lambda_2(\Sigma\Phi + \Phi\Sigma)^{ij} + g\,\chi_1^a(\sigma^{ij})_a^b\chi_{2b} = 0 \tag{7.2}$$

$$F_{\Sigma_1} = 0 \;:\; M_3\Sigma_1^{ij} + \lambda_3(\Sigma_1\Phi + \Phi\Sigma_1)^{ij} + \frac{f}{2}(H^iH'^j - H'^iH^j) = 0 \tag{7.3}$$

$$F_{\chi_1} = 0 \;:\; m\chi_{2a} + g\,\Sigma_a^b\chi_{2b} = 0 \tag{7.4}$$

$$F_{\chi_2} = 0 \;:\; m\chi_1^a + g\chi_1^b\Sigma_b^a = 0 \tag{7.5}$$

$$F_H = 0 \;:\; f\,\Sigma_1H' + 2\mu H = 0 \tag{7.6}$$

$$F_{H'} = 0 \;:\; fH\Sigma_1 + 2\mu' H' = 0 \tag{7.7}$$

Setting D = 0 forces $\chi_1 = \chi_2$ and $[A_\Sigma, B_\Sigma] = 0 = [A_\Phi, B_\Phi]$ where A and B are the real and imaginary parts of the scalar components of the superfield. The parametrisation (4) along with the following .

$$\begin{bmatrix} 0 & & & \\ & \begin{smallmatrix} & 1 \\ -1 & \end{smallmatrix} & & \\ & & \ddots & \\ & & & \begin{smallmatrix} & & & 1 \\ & & 1 & \\ & -1 & & \\ -1 & & & \end{smallmatrix} \end{bmatrix} \qquad \begin{aligned} H^T &= (0, 0 \ldots 0) \\ H'^T &= (0, 0 \cdots 0) \end{aligned}$$

gives us the following algebraic equations:

$$M_1v - \lambda_1v^2 + \frac{\lambda_2}{5}(b^2 - a^2)\sigma^2 + \frac{\lambda_3}{5}\sigma_1^2 = 0 \tag{8.1}$$

$$a\,\sigma(M_2 + 6\lambda_2v) = gx^2 \tag{8.2}$$

$$b\,\sigma(M_2 - 4\lambda_2v) = gx^2$$

$$m + g(3b + 2a)\sigma = 0 \tag{8.3}$$

$$M_3 = 4\lambda_3v \tag{8.4}$$

These can be solved as follows:

$$v = \frac{M_3}{4\lambda_3}$$

Using (8.2) and (8.4),

$$\frac{a}{b} = \frac{M_2 - 4v\lambda_2}{M_2 + 6v\lambda_2}.$$

Using (8.3),

$$b\sigma = \frac{-m}{g(\frac{2a}{b}) + 3}.$$
$$x^2 = b\frac{\sigma}{M_2 - 4v\lambda_2}$$

This follows from (8.2). The quantities on the R.H.S. are all known. Finally, we can substitute into (8.1) the expressions for v, b, σ and a/b to get an expression for σ_1. Given that M_1, M_2, M_3 and m are all of O(10^{16} GeV.) it is clear that x, σ, σ_1 and v are all also of the same order of magnitude. $<\Phi>$,$<\chi_1>$ and $<\Sigma>$ have little groups SO(4) X SO(6), SU(5) and SU(3)XSU(2)XU(1) respectively. The residual symmetry of the vacuum is therefore SU(3)XSU(2)XU(1) being the little group of the combined system (Φ, χ_1, χ_2, Σ)* This symmetry breaking occurs at 10^{16}. Note that $<\Sigma_1>$ has a little group SO(4)XU(3). Furthermore since the couplings of Σ_1 are not the most general SO(10) invariant ones, one can show that the parametrization chosen for $<\Sigma_1>$ is not unique. One can see this as follows: Equations (7.1) and (7.3) (which are the only ones containing Σ_1) have an SO(4)XSO(6) covariance. Thus any transformation generated by the elements of $\frac{SO(6)}{U(3)}$ changes the parametrization of $<\Sigma_1>$ but will leave it a solution of the equations because the form of the equations is unchanged. Thus we have a continuum of degenerate vacuum states corresponding to the different solutions for $<\Sigma_1>$. At the tree level, then, we have massless particles, pseudo-Goldstone bosons, parametrizing the coset space $\frac{SO(6)}{U(3)}$ scalars corresponding to 3+ $\overline{3}$ of SU(3). The expression pseudo Goldstone boson signifies the fact that the extra vacuum degeneracy is only present at the tree level. The full Lagrangian, in particular the gauge interaction, does not have the extra symmetry that corresponds to this vacuum degeneracy, and so quantum corrections remove the degeneracy and make the pseudo Goldstone bosons massive. Actually, in supersymmetric theories the preceding

*There are other solutions to the equations corresponding to the completely symmetric vacuum and to one with symmetry SU(4)XU(1). This ambiguity is resolved when supersymmetry is explicitly broken.

statement is not quite true. Unbroken supersymmetry ensures that the vacuum that had zero energy at the tree level continues to have zero energy even in the full quantum theory and the pseudo Goldstone particles remain exactly massless. Their fermionic partners are also exactly massless. Only when supersymmetry is broken do quantum corrections give masses to these particles. The scale of supersymmetry breaking has a lower bound coming from the requirement that these coloured strongly interacting particles get sufficiently large masses.

Before discussing supersymmetry breaking let us briefly review our position. The SO(10) symmetry is broken at a superheavy scale $M_X \approx 10^{16} GeV$. directly to SU(3)XSU(2)XU(1) and the $\frac{SO(10)}{SU(3)XSU(2)XU(1)}$ gauge bosons become superheavy via the usual Higgs mechanism. SU(2)XU(1) is still unbroken. The only light (massless) scalars in the theory are the Weinberg-Salam doublets and the $3+\overline{3}$ pseudo Goldstone bosons. The rest of the scalars, including the colour-triplet partners of the Weinberg-Salam doublets and the colour triplets of H' are superheavy . This splitting is achieved without any fine tuning of parameters. The price paid for this is the addition of a 45 to the minimal set of superfields. The superpartners of the ordinary particles are all degenerate in mass, because supersymmetry has not been broken. Thus, in particular, we have a host of unwanted light particles : the scalar partners of the quarks and leptons, the fermionic partners of the gauge bosons (gauginos), Higgs doublets, and the pseudo Goldstone bosons. The different vacua (the symmetric vacuum where all fields have zero vevs, the one with SU(4)XU(1)symmetry and the one with SU(3)XSU(2)XU(1)) symmetry are all degenerate in energy. This follows from the fact that supersymmetry is unbroken in each of the vacua, and that unbroken global supersymmetry automatically implies that the ground state has zero energy (see Appendix A). The present situation is thus not phenomenologically satisfactory. However all these problems can be solved by the simple expedient of breaking supersymmetry.

We have argued in sec.1.3 that soft explicit breaking of global supersymmetry is preferable to spontaneous breaking for several reasons. We therefore now modify our supersymmetric Lagrangian by adding non-supersymmetric terms. The kinds of soft supersymmetry breaking terms that are allowed to be added are listed in Appendix B (we remind the reader that "soft" here means "does not give rise to quadratic divergences"). Only dimension-two scalar mass terms are considered for reasons that will become clear soon. Thus,for instance, mass terms for the gauginos are not added since they have dimension three. Once a dimension-two term is added renormalizability requires all possible SO(10) invariant dimension two terms. Symanzik's theorem implies that terms of higher dimensionality are not required.

We need a term $-\mu_1^2 H^{2*}$ to give H^T a vev $(h_0,0,0,h_3,0,..0)$ which breaks SU(2)XU(1) -> U(1). This fixes μ_1 to be O(10^2) GeV. We also add positive mass terms for the scalar partners of the quarks and leptons, i.e., $\mu_2^2 M_a M_a$. (Note : In superfield notation a mass term of the form $\mu^2(A^2+B^2)$ is given by $\int d^4\theta\, U\overline{\varphi}\varphi$ where U = $\mu^2\theta^2\overline{\theta}^2$ and a mass term $\mu^2(A^2-B^2)$ is given by $\int d^2\theta\chi\varphi\varphi + h.c.$ where $\chi = \mu^2\theta^2$. See Appendix B for further details). Furthermore to pick the right SU(3)XSU(2)XU(1) vacuum a term $\mu_3^2 Tr\Phi^2$ is added. With the addition of these three terms, we have achieved the following: We have SU(2)XU(1) -> U(1) at the right scale $O(10^2)$GeV. This automatically makes the $\frac{SU(2)}{U(1)}$ gauge fields massive and to zeroth order in the supersymmetry breaking parameters μ_i^2, the corresponding gauginos also get masses of O(10^2) GeV. as required by low energy phenomenology. In this process the Higgs fermions in H also become heavy by pairing up with the gauge fermions to make a massive Dirac fermion. All the unwanted scalars except for the pseudo Goldstone bosons have masses of $O(10^2)GeV$. It is also clear that adding SO(10) invariant mass terms $\mu^2 Tr\Sigma_1^2, \mu^2 Tr\Sigma^2, \mu^2 Tr\Sigma_1\Sigma$ cannot make the pseudo Goldstone

*As before, we use the same symbol for the superfield and its scalar component. It is always clear from the context which is being referred to.

bosons heavy at the tree level, because the terms $\mu^2 Tr\Sigma^2$ and $\mu^2 Tr\Sigma_1^2$ being SO(10) invariant have the same value at the different minima and, therefore, do not distinguish between them. The term $\mu^2 Tr\Sigma\Sigma_1$ distinguishes between them, but due to the fact that the Σ fields are superheavy this mixing results only in masses of $O(\frac{\mu^2}{M})$ for the pseudo Goldstone bosons. These particles acquire a significant mass only when radiative corrections are included. The same is true for the gluinos (superpartners of the gluons) and the photino. The tree-level mass spectrum is shown in fig. 1. It agrees with low energy observations except for the afore-mentioned gauginos and pseudo Goldstone bosons.

Consider, now, the effects of the other dimension two non-supersymmetric terms which are required by renormalizability. In particular we have to ensure that they do not change h_0 and h_3, which set the weak breaking scale, by large amounts. The only supersymmetry breaking terms that directly affect h_0 and $h_3$are $\mu^2 H'^2$ and $\mu^2 HH'$. These are of the same order of magnitude as $\mu^2 H^2$ and do not cause h_0 and h_3 to shift dangerously. The other dimension two terms such as $\mu^2 Tr\Phi^2$,$\mu^2 Tr\Sigma^2$,$\mu^2\chi_1\chi_2$ etc. affect h_0 and h_3 only indirectly through changes in the vevs of Φ,Σ,Σ_1,χ_1,χ_2, etc. However, only those fields that are not superheavy can have their vevs changed significantly. This can be seen as follows: If ψ is a field with mass M and vev ψ_0, i.e. $\psi = \psi_0 + \tilde{\psi}$, then the potential near the minimum is approximated by $M^2\tilde{\psi}^2$. If we add a term $-\mu^2\psi^2$ the equation for its minimum becomes

$$2M^2\tilde{\psi} - 2\mu^2(\psi_0 + \tilde{\psi}) = 0.$$

Therefore $\tilde{\psi} \approx \frac{\mu^2}{M}$. Furthermore the superheavy fields ψ can only occur quadratically in an equation for h. (This follows from group theory considerations). They occur in equations of the form $\mu^2 h + (\psi_1^2 - \psi_2^2)h' = 0$ where $\psi_{1,2}$ are superheavy, and h, h' are 10's and μ is of $O(10^2)GeV$.. Before adding the perturbing term $\psi_1 = \psi_2$ at the minimum. Then $\delta h \approx O(\frac{\psi_1\delta\psi_1 h'}{\mu^2}) \approx O(\mu)$,which is harmless. Apart from the pseudo Goldstone bosons all the particles contained in Σ, Σ_1, Φ and $\chi_{1,2}$ are superheavy and do no harm. The pseudo Goldstone bosons in turn are contained in Σ_1^{ij} with i,j :4-9 , and they do not affect the equation for h_0 and h_3. Thus, we conclude that the addition of all possible dimension two terms has no significant effect on the vev of H and consequently does not induce any fine tuning. Note that this argument would not have worked if we had included dimension three terms. These would result in changes $\delta\psi \approx O(\mu)$. If the previous analysis is repeated with dimension-three terms, we get $\delta h \approx \frac{\psi\delta\psi h'}{\mu^2} \approx O(M)$. This would mean that the value of h at its minimum changes by an enormous amount and one would have to perform an extreme fine tuning to recover a value of $10^2 GeV.$ for the scale of the weak interactions. We conclude that we cannot add any dimension three terms, even if they do not introduce quadratic divergences.

This concludes the tree level analysis of the spectrum. At the loop level fig. 2a shows the pseudo Goldstone boson mass generation. Fig. 2b is the corresponding diagram for the fermionic superpartners of the pseudo Goldstone bosons. The pseudo Goldstone boson mass term generated is $\int d^4\theta\, U\overline{\Sigma}_1\Sigma_1$ where $U = \mu^2\theta^2\overline{\theta}^2$. The mass term for the fermions is $\frac{1}{m^3}\int d^4\theta \chi\overline{\chi}(D_\alpha\Sigma_1)(D^\alpha\Sigma_1)$. In both these expressions Σ_1 stands for the $3 + \overline{3}$ part of the 45.

The graph in fig. 2b is estimated to be $\frac{\alpha_s^2}{16\pi^2}\frac{\mu^4}{m^3}$. The two powers of μ^2 come from the two explicit insertions. The gluino couples strongly, hence, α_s^2, ($\alpha_s = \frac{g^2_s}{4\pi}$, where g_s is the QCD coupling constant. $\alpha_s \approx .1$) and the 16 π^2 in the denominator is because it is a two loop graph. m is the mass of the particles running around the loop and has to be $O(\mu)$ for this approximation to be valid. One has to choose $\mu \approx m \approx O(10^6 GeV.)$ for the mass generated to be $O(10^2) GeV.$. From this we get the constraint that supersymmetry has to be broken at a scale $O(10^6 GeV.)$, which is a few orders higher than the weak interaction scale. This could induce, either at the tree level, or via radiative corrections, a certain amount of fine tuning to the tune of three decimal places. However, in any case, in all such softly broken supersymmetric theories there is always another source of a three decimal place fine tuning which is needed for the super-GIM [7] mechanism to work* which we shall describe shortly. The graph in fig. 2b also forces us to add an extra representation of fields for the following reason: In the vacuum polarization loop of the gluino superfield the particles running around are colour triplets 'M'. The graph requires an explicit supersymmetry breaking mass term of the form $\int d^2\theta \chi M^2$. This means that M has to be a real representation of SO(10) (the mass term $\int d^4\theta\, U\overline{M}M$ would not require M to be in a real representation; however, this term would not serve our purpose). Furthermore, it cannot be superheavy. The ordinary quarks cannot be used because they belong to a 16 which is not real. We do have real 10's of Higgs but the colour triplets in them must be heavy (to prevent fast proton decay). So we have to introduce a multiplet , say a 10, having a mass of $O(10^6) GeV.$. At this point we have to check that the theory is asymptotically free : each superfield family ('superfamily') is equivalent to 3/2 ordinary families. The 3+$\overline{3}$ pseudo Goldstone super multiplet and the 3+ $\overline{3}$ from M can be combined into one superfamily. That makes four superfamilies (= 6 families) coming from the matter Higgs fields. The gluino being in the adjoint (and also Majorana) counts as 3/2 families. So we have a total of seven and a half families. Asymptotic freedom allows 8.25 families. The theory is therefore asymptotically free but cannot accommodate any more superfamilies.

The gluino mass comes from the diagram in fig. 3. For the supersymmetry breaking scale of $O(10^6 GeV.)$ the gluino mass is estimated to be $O(10^4 GeV.)$ which makes it completely unobservable.

The final superpotential and the explicit supersymmetry breaking terms are summarised below:

$$L = \int d^2\theta\, W + h.c. + \Delta L_{s.b.} + \textit{Kinetic} \text{ and } \textit{gauge terms}.$$

*Some models in which the supersymmetry breaking terms automatically have a family independent structure are exceptions. [12].

$$W = [\frac{M_1}{2} Tr\Phi^2 + \frac{\lambda_1}{3} Tr\Phi^3 + \frac{M_2}{2} Tr\Sigma^2 + \lambda_2 Tr\Sigma\Phi\Sigma + m\chi_1\chi_2 + g\chi_1\Sigma\chi_2]$$

$$+ (\frac{M_3}{2} Tr\Sigma^2 + \lambda_3 Tr\Sigma_1\Phi\Sigma_1 + fH\Sigma_1 H') +$$

$$+ \mu H^2 + \mu' H'^2 + \mu M^2 + f_{ab} M_a \Gamma M b H + g M_a \chi_2 S$$

$$\Delta L_{s.b.} = \int d^2\theta [\chi_1 M^2 + \chi_2 H^2 + \chi_3 h'^2 + \chi_4 \Sigma_1^2 + .].+.h.c.$$

$$\int d^4\theta [U_1 H\overline{H} + U_2 M_a M_{\overline{a}} + U_3 \Sigma_1 \Sigma_{\overline{1}} + ...]$$

where $\chi_i \equiv \mu_i \theta^2$, $U_i \equiv \mu_i^2 \theta^2 \overline{\theta}^2$ The three dots indicate that all other SO(10) invariant combinations should be included. The masses of some of these fields are constrained (as mentioned above). $\chi_1$and$\chi_4 \approx 10^6 GeV.$, $U_1 \approx -10^2 GeV.$, (responsible for the vev of H), $U_2 \approx 10^2 GeV.$, and should also be independent of the family in the first approximation i.e. $\frac{\Delta m^2}{m^2} \leq 10^{-3}$. This comes from requiring absence of flavour changing neutral currents, which we describe below. Apart from this there are no special requirements on the values of the other mass parameters.

To the superpotential the term μM^2 has been added for reasons already explained. The list of discrete symmetries can also be extended to include a fourth one: M-> -M. This prevents unnecessary couplings for M.

Consider the neutrinos. In this model the neutrino gets a Dirac mass term along with the other leptons and quarks. This is unsatisfactory phenomenologically. This problem can be solved by giving the right-handed neutrino a large mass. This causes it to decouple from the left-handed neutrino leaving the latter almost massless. For this purpose a singlet S has been added. When χ_2 gets a vev the right-handed neutrino becomes superheavy and decouples from the left-handed neutrino. Having added the singlet field S one has to modify the discrete symmetry (b) to $\chi_1 \to -\chi_1, \chi_2 \to -\chi_2, S \to -S$ and (c) to $H \to -H, H' \to -H', M_a \to e^{i\frac{\pi}{2}} M_a, S \to e^{-i\frac{\pi}{2}}$ to ensure the invariance of the added term $M_a \chi_2 S$.

We now turn to the $K_0 - \bar{K}_0$ mixing problem. In the standard model the G.I.M. mechanism takes care of this (fig. 4a). In supersymmetric theories there is a corresponding diagram where the $W^{\mp}$ bosons are replaced by the corresponding fermions, and the u, c quarks by their scalar superpartners (fig. 4b). In order for the cancellation to occur, the masses of these scalars have to satisfy the relation $\frac{\Delta m^2}{m^2} \leq 10^{-3}$. This has to be put in by hand and is a source of fine tuning that cannot be avoided in such models. It should be pointed out that in models where the softly broken supersymmetric theory is derived as the low energy limit of coupled matter-supergravity systems, this mass relation is automatically implemented because of the flavour independence of the gravity interaction.

Finally, we comment on proton decay modes in supersymmetric models [14]. Unlike conventional grand-unified theories where the baryon number violation takes place via the superheavy vector bosons that result in effective dimension-6 operators, in supersymmetric grand-unified theories it is possible to have dimension-5 operators. These are shown in fig. 5a. The operator is, therefore, suppressed by only one power of the superheavy mass. Completing it to a four fermion operator introduces only an additional suppression of $\frac{1}{\tilde{M}_W}$,where $\tilde{W}$ is the gaugino (fig. 5b). Fig. 5a has a coefficient $\frac{g_H^2}{m_H}$ where g_H is a Yukawa coupling $\approx \frac{m_{quark}}{<v>}$ and m_H is the mass of the Higgs triplet. Fig. 5b contributes a factor $\frac{1}{16\pi^2} \frac{g^2}{M_W}$ where g is a gauge coupling. Thus there is an overall factor $\frac{1}{16\pi^2} \frac{g_H^2}{\tilde{m}_W} \frac{g^2}{m_H}$. In fig. 5a the (q q sq sq) operator has to involve one set (u, c, t) of quarks at one vertex and the other set (d, s, b) at the other vertex. This is seen by explicitly writing out all the dimension five operators that can be generated by the graphs of fig. 5 [14]. So $g_H^2 \approx m_u \frac{m_s}{M_W^2}$. Thus the overall factor becomes $\frac{m_u m_s}{16\pi^2 M_W^2} \cdot \frac{g^2}{\tilde{M}_W M_H}$. This tells us two things : first, although dimension five operators are involved, because the Yukawa couplings are small the proton lifetime much longer than naive expectations based on power counting, and is compatible with experimental lower limits. Second, because of the factor m_s strange decays are preferred to non strange decays in contrast with conventional grand-unified theories. Thus typically N -> $\bar{\nu}_\tau$ + *strange* and $N \to \bar{\nu}_\mu$ + *strange* are the dominant modes [14] rather than N -> e^+ + pion as in conventional grand-unified theories. Lifetimes are of the same order as in non-supersymmetric theories - 10^{31} years.

5. Conclusions

We have described a possible solution to the fine tuning problem that has plagued grand-unified theories from the beginning. The essential ingredient for the solution was supersymmetry. There are independent and strong motivations for positing that supersymmetry might be an exact symmetry of nature at some high energy scale. These come from the promise held out by some of the supergravity theories, like N=8 supergravity (or perhaps the superstring theory [15]) of both providing a sensible quantum theory of gravity, and unifying gravity with the other three interactions. The gauged N=8 supergravity [16] seems to hold particular promise. We would like, therefore, to conclude this chapter with a short analysis of possible directions that could be pursued in trying to relate supergravity to the real world.

We assume for the purposes of this discussion that the N=8 supergravity theory (or the string theory, which has N=8 supergravity as its low energy limit) is the ultimate theory of nature. To establish a connection with N=8 supergravity, the globally supersymmetric grand-unified theories have to be extended in two obvious directions. First, gravity has to be included, and second, the other seven supersymmetries have to be accounted for. These seven supersymmetries could, in principle, remain good symmetries up to energies well below the Planck mass. However, a closer inspection reveals that this is unlikely, simply because the fermion representation content of higher N supersymmetric theories is vector like. For every left-handed particle, there is a right-handed particle with exactly the same quantum numbers. The observed particle spectrum does not show this left-right symmetry. In fact it seems to be impossible to write down an N=2 globally supersymmetric version of the Weinberg-Salam theory, which would reduce at some energy scale to an effective N=1 supersymmetric theory, and at yet lower scales to the standard Weinberg-Salam model. Thus, it seems highly likely that the seven remaining supersymmetries are broken at around the Planck mass. If that is the case, it would seem more fruitful to study the matter-supergravity systems starting with N=1 and then proceed with the question of the remaining supersymmetries in the context of supergravity theories. In these theories there is the possibility that the particle representations can be flavour chiral, because the supersymmetry charges themselves can carry internal quantum numbers. The N=1 matter supergravity system has received some attention. A plausible scenario that has been found to be phenomenologically viable is the following: The remaining supersymmetry is broken at some reasonably high scale 'M', by the super-Higgs effect in some sector of the theory. The effective low energy theory is a globally supersymmetric theory with soft explicit supersymmetry breaking terms. The mass of the gravitino $m_{\frac{3}{2}}$ is the relevant scale of global supersymmetry breaking. This scheme also has the advantage that [12] one finds, for instance, a common flavour independent mass term for all scalar particles. This is useful for the super GIM mechanism. The supergravity effects also remove the vacuum degeneracy problem. To make further progress, however, we need a more thorough understanding of the N=8 supergravity theory.

Appendix A : Super Poincare Algebra and Representations

$$\{Q^i_\alpha, Q^j_\beta\} = \{\bar{Q}_{\dot{\alpha}i}, \bar{Q}_{\dot{\beta}j}\} = 0 \tag{A1}$$

$$\{Q^i_\alpha, \bar{Q}_{\dot{\beta}j}\} = 2\sigma^\mu_{\alpha\dot{\beta}}\delta^i_j P_\mu$$

$$[Q^i_\alpha, P_\mu] = [\bar{Q}_{\dot{\alpha}i}, P_\mu] = 0$$

$$[Q^i_\alpha, M_{\mu\nu}] = i(\sigma_{\mu\nu}Q^i)_\alpha$$

$$[Q^i_\alpha, B_a] = iS^i_{a\ j}$$

$$[B_a, B_b] = if_{abc}B_c$$

Q's are the supersymmetry generators. We have written them in two component notation $(Q^i_\alpha)^* = \bar{Q}_{\dot{\alpha}i}$. 'i' is the internal symmetry index and $\bar{Q}_{\dot{\alpha}i}$ transforms as the complex conjugate representation of Q^i_α. B_a is the internal symmetry generator represented by the Hermitian matrices $S^i_{a\ j}$ on the Q's.

From the relation $\{Q^i_\alpha, \bar{Q}_{\dot{\beta}j}\} = \sigma^\mu_{\alpha\dot{\beta}}P_\mu$ we can derive the following :

$$\{Q_1, Q_{\dot{1}}\} + \{Q_2, Q_{\dot{2}}\} = 2P_0$$

If $Q^i_\alpha|0\rangle = 0$ i.e. the vacuum is supersymmetric then $\langle 0|\{Q^i_\alpha, \bar{Q}_{\dot{\alpha}i}\}|0\rangle = 0 \Rightarrow \langle 0|P_0|0\rangle = 0$ This implies that the vacuum energy is zero. This is an operator identity and is exact provided the supersymmetry algebra is valid. It is easy to see that $P_\mu P^\mu$ is a Casimir invariant for the above algebra whereas $W_\mu W^\mu$ (W is the Pauli-Lubanski vector) is not. Thus a single miultiplet contains particles of different spins. We can now study the representations. Take the massless case : (assume only one supersymmetry)

$$\{Q^i_\alpha, \bar{Q}_{\dot{\beta}j}\} = 2\sigma^\mu_{\alpha\dot{\beta}}P_\mu \quad \{Q^i_\alpha, Q^j_\beta\}=0 \tag{A2}$$

Choosing a frame where $P_\mu = (1,0,0,1)P$

$$\{Q^i_\alpha, \bar{Q}_{\dot{\beta}j}\} = 2P(\sigma^0 + \sigma^3)_{\alpha\dot{\beta}} \tag{A3}$$

In the representation

$$\sigma^0 = -\sigma_0 = -I \qquad \sigma^3 = \sigma_3 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$$

We get

$$\{Q_1, \overline{Q}_{\dot{1}}\} = 0 \tag{A4}$$

$$\{Q_2, \overline{Q}_{\dot{2}}\} = 2P$$

Thus applying $Q_1 or \overline{Q}_{\dot{1}}$ on a state gives a state of zero norm and we only use Q_2 or $\overline{Q}_{\dot{2}}$. If we start with a state $|n>$ of helicity n that satisfies $\overline{Q}_{\dot{2}}\ |n> = 0$ we can construct a state $Q_2\ |\ n>$. It can be shown using the commutation relations that thes states have helicity n + 1/2 . Thus $Q_2|n> = |n+\frac{1}{2}>$ and $Q_2Q_2|n> = 0$ because of the anticommuting nature of the Q's. Thus the multiplets have the form (n, n +1/2). (0, 1/2) is the scalar multiplet, and (1/2 , 1) is the vector multiplet. It is understood that the CPT conjugates are also added to preserve CPT invariance.

We can modify the above construction easily to get represention of the higher N superalgebras. Let us derive the N=2 representation which will be needed in Appendix B. We adjoin to the Q's an internal symmetry index i that takes on the values 1 and 2. Thus A4 is modified to

$$\{Q_1^i, \overline{Q}_{\dot{1}_j}\} = 0 \qquad \{Q_2^i, \overline{Q}_{\dot{2}_j}\} = 2P\delta_j^i \tag{A5}$$

Starting with $|n>$ satisfying $\overline{Q}_{\dot{2}_i}\ |n> = 0$ we get $Q_2^1\ |n>$ and $Q_2^2\ |n>$ as two other states with helicity n + 1/2 and helicity n + 1 respectively. Thus the multiplet contains helicities (0, 1/2 , 1) or (-1/2 , 0, +1/2). The first two are the vector and scalar hypermultiplet [17]. Once again it is understood that the CPT conjugate of each particle is also included.

Proceeding in a similar manner we can construct representations for all the higher N theories. Globally supersymmetric theories can have up to four supersymmetries so that spin ≤ 1 and supergravity theories can have up to eight without exceeding a spin of two. The table below gives the representations of superalgebras with $N \leq 4$.

Appendix B : Soft Supersymmetry Breaking

The soft breaking terms for N=1 supersymmetry are the following [9]:

$$\frac{1}{2}\int d^4\theta\, U\overline{\varphi}\varphi = \frac{\mu^2}{2}(A^2 + B^2) \qquad U = \mu^2\theta^2\overline{\theta}^2 \tag{a}$$

$$\frac{1}{2}\int d^2\theta \chi\varphi^2 + h.c. = \mu^2(A^2 - B^2) \qquad \chi = \mu^2\theta^2 \tag{b}$$

$$\int d^2\theta \eta W^\alpha W_\alpha + h.c. = \mu\lambda\overline{\lambda} \qquad \eta = \mu\theta^2 \tag{c}$$

$$\int d^2\theta \eta\varphi^3 + h.c. = \frac{\mu}{\sqrt{2}}(A^3 - 3AB^{2)} \qquad \eta = \mu\theta^2 \tag{d}$$

N=2 Supersymmetry

In terms of N=1 superfields the Lagrangian is the following:

$$L = \int d^4\theta[W^\alpha W_\alpha + \overline{N}e^{2eV}N + \overline{S}e^{2eV}S + \overline{T}e^{-2eV}T] + \int d^2\theta(4e\overline{T}SN + m\overline{T}S) + h.c.$$

The notation is the following : V is a vector superfield. N, S and $\overline{T}$ are scalar superfields. (V, N) form an N=2 vector hypermultiplet and (T, S) form an N=2 scalar hypermultiplet. (For further details see ref. [17]). This Lagrangian has an internal SU(2) symmetry that rotates the two supersymmetry generators into each other. More useful for us will be the following approximate 'R' symmetry: $N \to e^{ix}N, S \to e^{iy}S, T \to e^{i(x+y)}T$ for *arbitrary* x, y. It is only an approximate symmetry because the mass term violates it.

All the soft breaking terms allowed by N=1 supersymmetry are allowed here also. In addition the following are allowed:

$$\int d^4\theta\, UD^\alpha ND_\alpha N = \mu\overline{\psi}\psi \qquad U = \mu\theta^2\overline{\theta}^2 \tag{1}$$

In the N=1 case gauge fermion mass terms were allowed. This mass term is the N=2 extension of that. This expression has R- number 2x . From power counting we see that quadratic divergences have to be of the following forms: (Note that U has dimension -1)

(i) $\int d^4\theta\, UN$. This term is allowed if N is a singlet i.e. if the gauge group has a U(1) factor. It has R- number x. The difference of x has to be due to the mass term. Thus at least one power of m should multiply this expression, which means the divergence can at most be logarithmic.

(ii) $\int d^4\theta\, UT, \int d^4\theta\, US$. These terms have the wrong R-number and cannot occur. Moreover they would have to be singlets for this expression to be allowed. In N=2 supersymmetry singlets decouple from everything else and are uninteresting.

(iii) $\int d^4\theta\, UU...$ terms with higher powers of U. These would require $D^2$and$\overline{D}^2$ to survive the θ integration. Since these have dimensions of mass they reduce the degree of divergence.

Thus the term $\int d^4\theta\, UD^\alpha N D_\alpha N$ is soft.

(2) $\int d^4\theta\, UD^\alpha T D_\alpha T = m\bar{\chi}\chi$ is allowed by an analysis identical to the one above. Thus fermion masses are allowed.

(3) $\int d^4\theta\, U(T + \overline{T})^3 = mA^3$. Has R-number of 3(x+y). or (x+y). Violates internal SU(2). The possible quadratic divergences are:

(i) $\int d^4\theta\, UN$. Has R-number x and cannot be generated.

(ii) $\int d^4\theta\, UT$. This can be there if T is a singlet. (figure 6)

(iii) $\int d^4\theta\, US$. Has R-number y and would require an m insertion. Hence only logarithmically divergent.

(4) $\int d^4\theta\, U(T+\overline{T})(S+\overline{S})(N+\overline{N})$. Has R-number 0, 2x, 2y, 2x+2y.

(i) $\int d^4\theta\, UN$. Has R-number x. So at least one m insertion is needed. Only logarithmic divergence.

(ii) $\int d^4\theta\, UT$. Has R- number x+y. Not generated.

(iii) $\int d^4\theta\, US$ Has R- number y. Not generated.

(5) $\int d^4\theta\, U(T^3+\overline{T}^3) = \int d^2\theta\eta(T^3+\overline{T}^3)$ where $\eta = \mu\theta^2$. This term was allowed in N=1 supersymmetry also so it is certainly allowed here.

(6) $\int d^4\theta\, UT\overline{T}(T+\overline{T}) = \mu A(A^2+B^2)$. Has R-number $\pm(x+y)$.

(i) $\int d^4\theta\, UN$. Has R-number x. Hence not generated.

(ii) $\int d^4\theta\, US$. Has R-number y. Requires one 'm' insertion. So it is not dangerous.

(iii) $\int d^4\theta\, UT$. Is generated if T is a singlet (figure 6).

(7) $\int d^4\theta(N^3+\overline{N}^3)$. Allowed since it is allowed in N=1 supersymmetry.

(8) $\int d^4\theta N\overline{N}(N+\overline{N})$. Generates quadratic divergences $d^4\theta\, UN$ if N is a singlet.

(9) $d^4\theta(N+\overline{N})^3$. Has R-number $\pm x, \pm 3x$.

(i) $\int d^4\theta\, UN$. Generated if N is a singlet. (figure 7)

(ii) $\int d^4\theta\, UT$. Has R-number x+y. Hence not generated.

Thus we conclude that if there are no singlets all dimension three terms are allowed. If N is a singlet (U(1) factor in the gauge group) then terms (8) and (9) are not allowed. If T is a singlet, terms (3) and (6) are not allowed. This is the complete list of soft supersymmetry breaking terms for N=2 supersymmetry.

Appendix C

The SO(10) generators in the spinor (16) representation are reproduced for convenience. If we let $\vec{\sigma},\vec{\tau},\vec{\eta},\vec{\rho}$ stand for the following:

$$\vec{\sigma} \equiv \begin{bmatrix} \vec{\sigma} & & & & & & & \\ & \vec{\sigma} & & & & & 0 & \\ & & \vec{\sigma} & & & & & \\ & & & \vec{\sigma} & & & & \\ & & & & \vec{\sigma} & & & \\ & 0 & & & & \vec{\sigma} & & \\ & & & & & & \vec{\sigma} & \\ & & & & & & & \vec{\sigma} \end{bmatrix} \qquad \vec{\tau} \equiv \begin{bmatrix} \vec{\tau}\times 1_\sigma & & & \\ & \vec{\tau}\times 1_\sigma & & \\ & & \vec{\tau}\times 1_\sigma & \\ & & & \vec{\tau}\times 1_\sigma \end{bmatrix}$$

$$\vec{\eta} \equiv \begin{bmatrix} \vec{\eta}\times 1_\tau \times 1_\sigma & 0 \\ 0 & \vec{\eta}\times 1_\tau \times 1_\sigma \end{bmatrix} \qquad \vec{\rho} \equiv \begin{bmatrix} \vec{\rho} \times 1_\eta \times 1_\tau \times 1_\sigma \end{bmatrix}$$

i.e. $\vec{\sigma}$ acts on the smallest 2 X 2 space, $\vec{\tau}$ on the next smallest and so on, then the generators are:[13]

$$\mu_{ij} = \varepsilon_{ijk}\eta_k \qquad \mu_{i+3\,j+3} = \varepsilon_{ijk}\sigma_k \qquad \mu_{i+6\,j+6} = \varepsilon_{ijk}\tau_k$$

$$\mu_{i0} = \eta_i\rho_3 \qquad \mu_{i+3\,0} = \sigma_i\rho_1 \qquad \mu_{i+6\,0} = \tau_1\rho_2$$

$$\mu_{i+3\,j+6} = \sigma_i\tau_j\rho_3 \qquad \mu_{i+6\,j} = \tau_i\eta_j\rho_1 \qquad \mu_{i\,j+3} = \eta_i\sigma_j\rho_2$$

where i,j:1-3.

The 'σ' matrices in this representation are μ_{0j}. This is the analog of the σ matrices of SO(4). Thus the 10 can be written as $H^0 + i\mu_{0j}H^j$ in terms of 16X16 matrices. (Analogous to writing the 4 of SO(4) as $a^0 + a^i\sigma^i$).

References

[1] L. Susskind, P.R.D18 (1978) 4500

[2] S. Dimopoulos, L. Susskind, Nucl. Phys. B155 237

[3] E. Witten, P.L. 105B (1981) 267

[4] P. Fayet, S. Ferrara, Phys. Rep. 32C (1977) 249 ; P. Van Nieuwenhuizen, Phys. Rep. 68(1981) 189

[5] M. Grisaru, M. Rocek, W. Siegel Nucl. Phys. B159 (1979) 429

[6] J. Wess, B. Zumino Phys. Lett. 32 (1974) 52

[7] S. Dimopoulos, H. Georgi Nucl. Phys. B193 150 (1981)

[8] H. Georgi, S. Glashow Phys. Rev. Lett. 32 (1974) 438

[9] M. Grisaru, G. Girardello Nucl. Phys. B 194 (1982) 65

[10] S. Dimopoulos, F. Wilczek Phys. Rev. D24 (1981) 1681

[11] R. N. Mohapatra, MPI-PAE/Pth 8/81, Feb 1981.

[12] A. H. Chamseddine, P. Nath, R. Arnowitt Phys. Rev. Lett D24 (1981) 1681

[13] H. Georgi, D. Nanopoulos Nucl. Phys. B155 (1979) 52

[14] J. Ellis et al. Nucl. Phys. B 202 (1982) 43

[15] J. Schwarz Phys. Rep. 89 No.3 (1982)

[16] B. De Wit, H. Nicolai, Phys. Lett. 108B 285 1982

[17] P. Fayet Nucl. Phys. B113 (1976) 135

Figure Captions

[1] Tree level particle spectrum. The twiddle indicates a superpartner. e.g. $\tilde{g}$ stands for the gluino. 'pgb' stands for pseudo-Goldstone boson.

[2a] The finite part of this graph gives a mass to the pseudo goldstone bosons A and B which are a 3+$\bar{3}$. $\int d^4\theta\, U\Sigma_1\bar{\Sigma}_1 \approx \mu^2(A^2+B^2)$.

[2b] A graph which gives a mass to the fermionic partners of the pseudo goldstone bosons. $\frac{1}{m^3}\int d^4\theta\bar{\chi}\chi(D^\alpha\Sigma_1)(D_\alpha\Sigma_1) \approx \frac{\mu^4}{m_3}\psi^\alpha\psi_\alpha$. Here Σ_1 and ψ stand for the 3+$\bar{3}$ piece of the 45.

[3] Graph which gives a mass to the gluino. $\frac{1}{m}\int d^4\theta\chi Tr(\bar{D}^2D^\alpha V)(D_\alpha V) \approx \frac{\mu^2}{m}Tr\,\lambda^\alpha\lambda_\alpha$.

[4a] ΔS = 2 box diagram - GIM mechanism.

[4b] ΔS = 2 box diagram -Super GIM mechanism. Twiddles indicate superpartners.

[5a] ($ff\tilde{f}\tilde{f}$) dimension five operator is generated by exchange of the Higgs superfields. Arrows indicate the flow of chirality of a superfield. Note that the corresponding diagram with a gauge superfield exchange is suppressed by chirality conservation.

[5b] Exchange of a $\tilde{g}$ or $\tilde{W}$ converts the scalar matter to ordinary fermionic matter. This is only suppressed by $\frac{1}{\tilde{M}_W}$.

[6] The vertex $\int d^4\theta\, UTT\bar{T}$ results in a quadratic divergence $\int d^4\theta\, UT$ if T is a singlet.

[7] Exactly the same situation as in fig.6 with T replaced by N

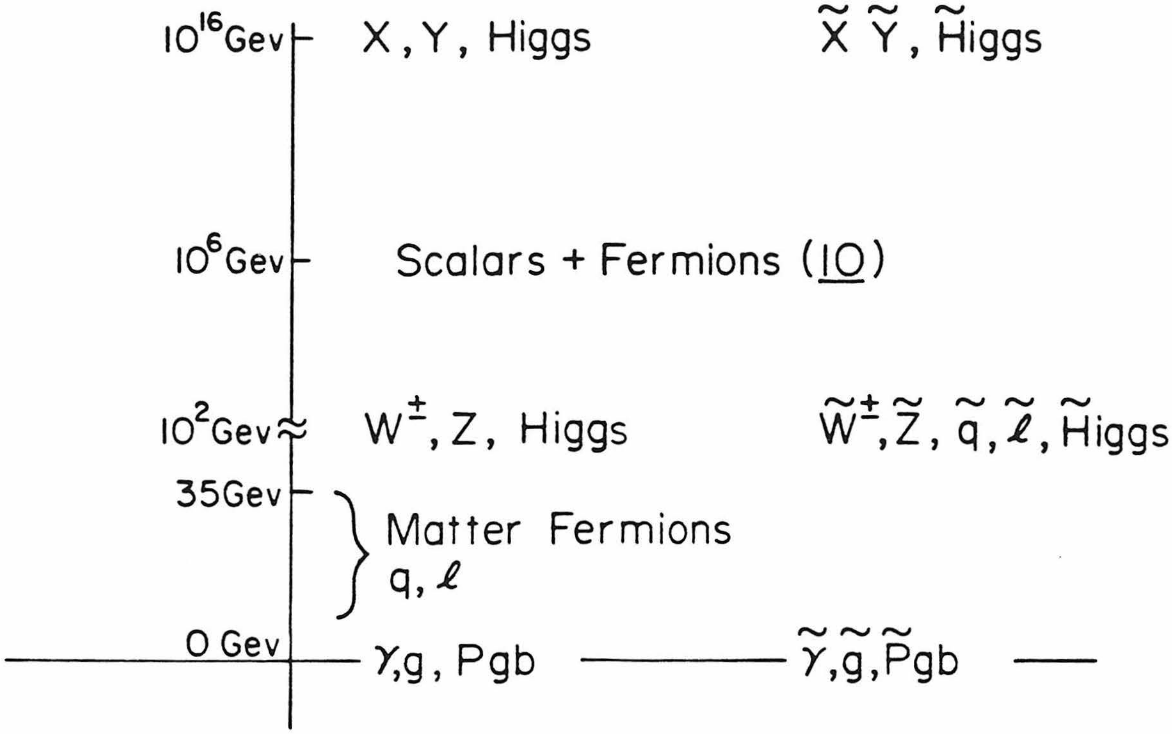
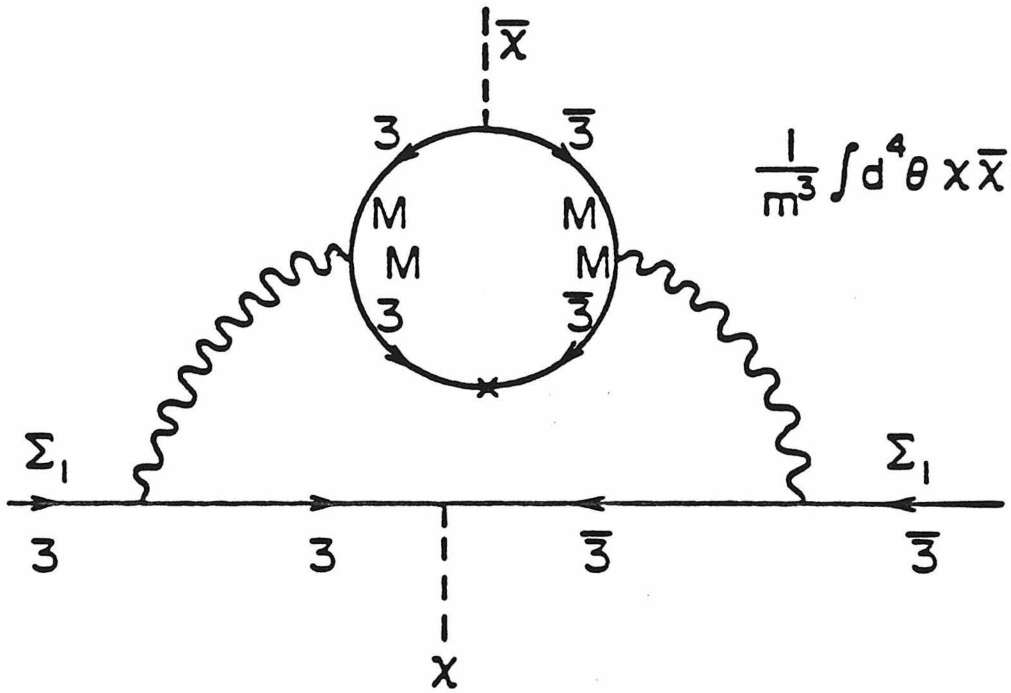
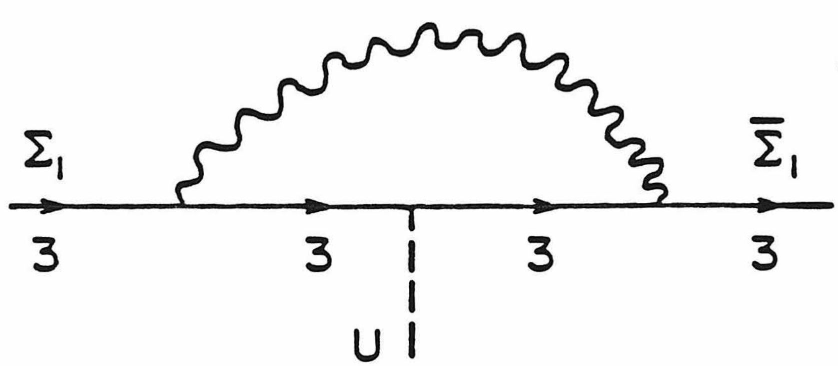

10^{16} Gev
X, Y, Higgs
$\tilde{X}$ $\tilde{Y}$, $\widetilde{Higgs}$
10^{6} Gev
Scalars + Fermions ($\underline{10}$)
10^{2} Gev
$W^{\pm}$, Z, Higgs
$\tilde{W}^{\pm}$, $\tilde{Z}$, $\tilde{q}$, $\tilde{\ell}$, $\widetilde{Higgs}$
35Gev
Matter Fermions
q, ℓ
0 Gev
γ, g, Pgb
$\tilde{\gamma}$, $\tilde{g}$, $\widetilde{Pgb}$

Fig. 1.



$$\frac{1}{m^3}\int d^4\theta\, X\bar{X}(D^\alpha\Sigma_1)(D_\alpha\Sigma_1)$$

Fig. 2b



$$\int d^4\theta\, U\Sigma_1\bar{\Sigma}_1$$

Fig. 2a

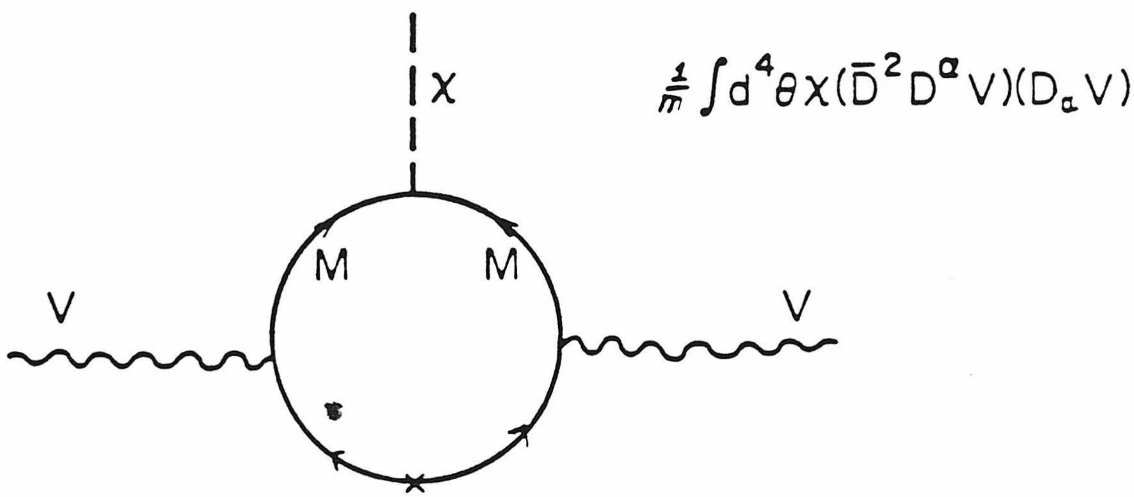
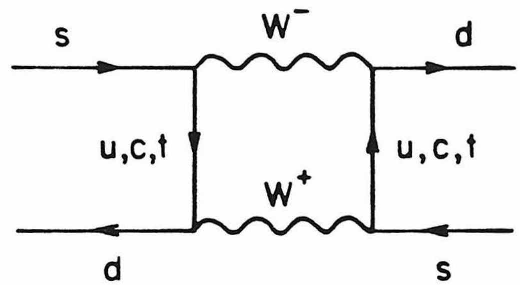
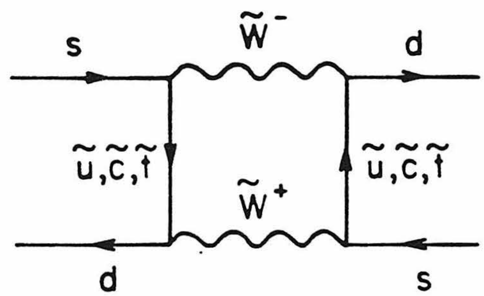


Fig. 3



GIM

Fig. 4a.



Super - GIM

Fig. 4b.

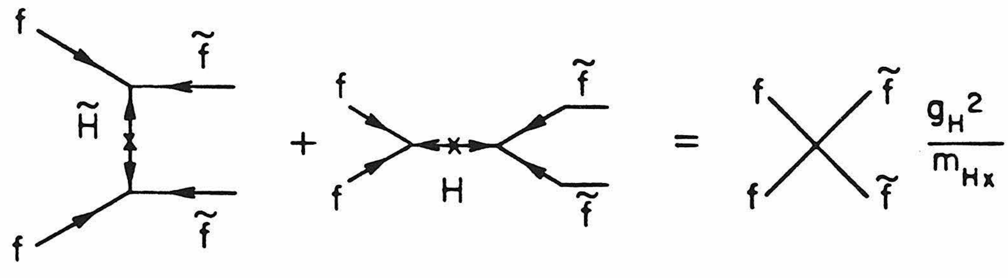

f
$\tilde{f}$
$\tilde{H}$
+
f
H
=
$\frac{g_H^2}{m_{Hx}}$

Fig. 5a.

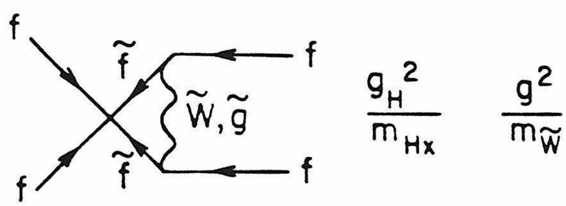

f
$\tilde{f}$
$\tilde{W}, \tilde{g}$
$\frac{g_H^2}{m_{Hx}}$
$\frac{g^2}{m_{\tilde{W}}}$

Fig. 5b.

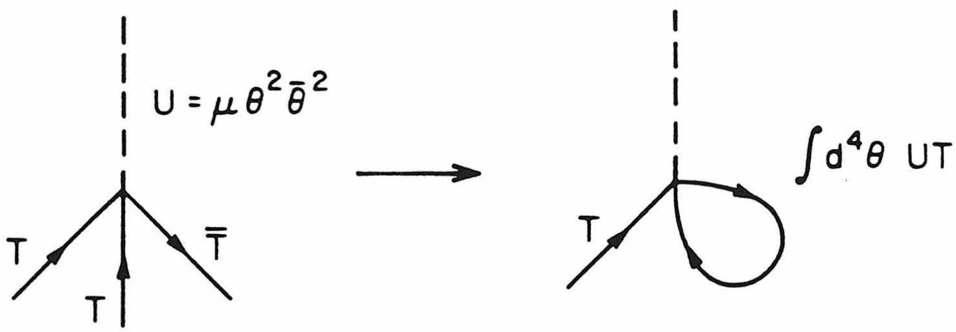

$U = \mu\theta^2\bar{\theta}^2$
T
T
$\bar{T}$
T
$\int d^4\theta$ UT

Fig. 6

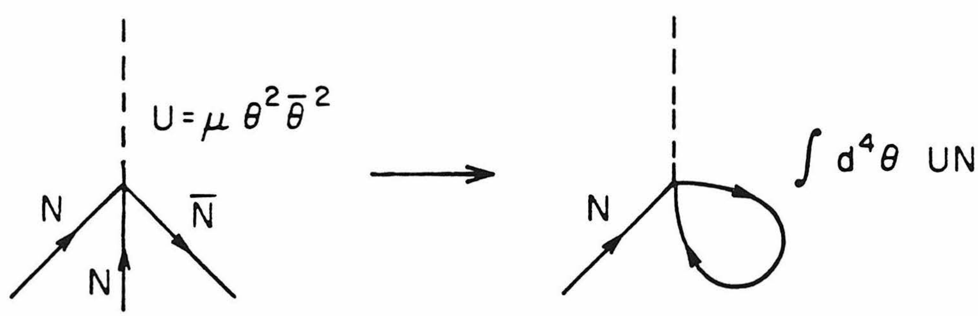

$U = \mu\theta^2\bar{\theta}^2$
N
N
$\bar{N}$
N
$\int d^4\theta$ UN

Fig. 7.

Monopoles and Fermion Number Violation

1. Grand Unified Monopoles and Baryon Number Violation

Recently, some interesting properties of the fermion-monopole system have been discovered [1,2]. It was found that monopoles can catalyse processes that violate fermion number and chirality, processes that are otherwise strongly suppressed. The calculations were done in a spontaneously broken SU(2) theory where the surviving symmetry is U(1) (Georgi-Glashow model), and with fermions in the isospinor (2) representation. It was found that in the presence of these (massless) fermions, the monopole is surrounded by a chirality and fermion number violating condensate of operators multilinear in the fermion fields. This implies that the ground state in the monopole sector breaks fermion number and chirality. This would manifest itself in chirality violating scattering amplitudes. Even more remarkably, the condensate formation is not suppressed by any powers of the coupling constant or inverse powers of the large mass parameter in the theory. Thus, in the presence of a monopole, these chirality and fermion- number violating processes would have large cross sections. This has phenomenological consequences in grand unified theories which are known to possess classical solutions that describe magnetic monopoles [3].

The monopoles of grand unified theories have a finite but very large mass and are singularity free, unlike the Dirac monopole. It has been argued [4] that they should have been produced copiously in the early universe. There has even been a possible experimental indication of their existence [5] although no subsequent experiment has seen any evidence.

The spherically-symmetric monopole solutions of the SU(5) theory have been classified [6]. If one were to extend the fermion number and chirality violation results of the SU(2) theory to the SU(5) case, one would be led to the remarkable conclusion that baryon number is strongly violated by these monopoles and the cross sections would have the magnitude typical of strong interaction cross sections. There have been attempts to pin down the various observational consequences this would have, especially in astrophysics [7] and also in the proton decay experiments. Typically, these can be converted to bounds on the number density of monopoles, which in turn gives some useful information about the early universe. However, one needs a much more detailed and quantitative understanding of the various reaction rates and selection rules of processes that could be catalysed by monopoles. Before anything definite can be said, the calculations of Rubakov and Callan which were done with SU(2) monopoles and isospinor fermions must be extended to different fermion representations, to different kinds of monopoles with different magnetic charges, and one must also include the effects of non-zero fermion masses. In this dissertation, we take a first step towards generalizing these calculations by analysing arbitrary fermion representations. This extension is quite straightforward when one realizes that the underlying physics is essentially the same as that responsible for $U_A(1)$ violation in QCD [8]. This chapter is organized as follows. First, we review briefly some classical and quantum aspects of the 't Hooft-Polyakov monopole in theories without any fermions. In section 1.3 we introduce massless fermions into the theory and give a qualitative picture of condensate formation around monopoles. This is followed up in section 1.4 with an explicit calculation. In section 1.5 these results are interpreted and a discussion of the general features of a fermion-monopole system is given. In section 1.6 we generalize to the case where fermions in arbitrary representations of the gauge group are present. In section 1.7 the relevance of these calculations to the real world is illustrated with two examples.

2. The 't Hooft-Polyakov Monopole.

Consider an SU(2) gauge theory with a triplet of scalars, which gets a vev and breaks the gauge group down to U(1). There is a continuum of degenerate minima for the scalar potential corresponding to the coset space SU(2)/U(1). The following configuration of the scalar fields has a conserved topological winding number associated with it and is therefore stable against decay (fig. 8) : take a point 'O' where the Higgs has zero vev, and a surface S_2 around it. On the surface S_2 the Higgs vev takes values in SU(2)/U(1) in such a way that it defines a mapping from $S_2 \to \frac{SU(2)}{U(1)}$ corresponding to the element 1 of $\pi_2(\frac{SU(2)}{U(1)})$*. This element is the winding number of the map $S_2 \to \frac{SU(2)}{U(1)}$, i.e., it counts the number of distinct points of S_2 that are mapped onto each element of $\frac{SU(2)}{U(1)}$. If we fix the axes of SU(2) isospin in the figure so that τ^1 is along the x-axis, τ^2 along the y-axis and so on, a configuration with winding number one, is one in which the Higgs field points radially in isospin space. (The trivial configuration where the Higgs fields point in the same direction everywhere has winding number = 0.) To change the winding number of a configuration, the Higgs fields everywhere have to be rotated by a finite angle in isospin space. A finite kinetic energy has to be expended per unit volume if this rotation is done in a finite period of time. If we assume an infinite universe, this process requires infinite action and does not take place. These configurations are therefore stable. At this point, as described, these configurations seem to have infinite energy because the derivatives $\partial_i\varphi \approx \frac{1}{r}\partial_\theta\varphi \approx \frac{1}{r}$. This means that energy (E)

$$E \approx \int d^3x\,(\partial_i\varphi)^2 \approx \int r^2 dr \frac{1}{r^2} \approx \int dr \to \infty$$

However, the theory has a local gauge invariance and this is illusory. One can

---

* $\pi_2(\frac{SU(2)}{U(1)})$*. is the second homotopy group of $\frac{SU(2)}{U(1)}$. The elements of this group are the different classes of maps from S_2 to $\frac{SU(2)}{U(1)}$. Two maps correspond to the same element of π_2 if they can be continuously deformed into each other. $\pi_2(\frac{SU(2)}{U(1)}) \approx Z$ (group of integers).

choose a gauge (coordinate system) in such a way that the Higgs fields appear to point in the same direction everywhere. This is equivalent to having the connections W_μ satisfy $D_\mu\varphi = 0$ (although $\partial_\mu\varphi \neq 0$). From the topology of the configuration it is clear that $D_\mu\varphi$ cannot be zero all the way down to the origin without encountering a singularity. In fact, on solving the equations of motion, one finds $D_\mu\varphi \approx e^{-Mr}$,where M is the scale of SU(2) breaking rather than $\frac{1}{r}$ as in the theory without gauge invariance. The solution to the equations of motion in the radial (spherically symmetric) gauge is the following:

$$\varphi^a = \frac{c\hat{r}}{gr}H(r)$$

$$W_i^a = \frac{\varepsilon^{aij}r_j}{gr}F(r) \qquad (1)$$

$$W_0^a = 0$$

where a = 1,2,3 is the SU(2) isospin index and i = 1,2,3 labels Cartesian spatial directions. H(r) and F(r) are as shown in fig. 9. They satisfy H(0)=F(0)=0 and H(∞)=F(∞)= 1, approaching these values asymptotically. The scale of variation is characterised by r_M, which can be called the core radius. It is inversely proportional to the Higgs vev and the mass of the heavy bosons. Outside the core, the non vanishing component of the field strength $\vec{B}$ points in the unbroken U(1) direction $\hat{r}.\vec{T}$. This can be identified as the magnetic field of a monopole. The quantity $\int_S \vec{B}.d\vec{S}$ is a constant conserved quantity independent of the above surface S, and can be called the magnetic charge 'm'. For the configuration (1), $\frac{mg}{4\pi} = -1$, where that 'g' is the electric charge of the $W^\pm$ boson. If isospin half fermions are introduced with charge q = $\pm\frac{g}{2}$, then $\frac{qm}{4\pi} = \frac{1}{2}$, the smallest value allowed by the Dirac quantization condition.

The solution (1) is symmetric under a combined isospin ($\vec{T}$) and space ($\vec{L}$) rotation. It satisfies

$$[L_i + T_i, W_j] = i\varepsilon^{ijk} W_k$$

$$[L_i + T_i, \varphi] = 0,$$

i.e. it is a scalar under $\vec{L}+\vec{S}+\vec{T}$ ($\vec{S}$ is spin). Thus in the monopole background the quantity $\vec{J}=\vec{L}+\vec{S}+\vec{T}$ is conserved. It also has the commutation relations appropriate to angular momentum and hence it can be called angular momentum.

Thus far the discussion is purely classical. Quantization of the theory requires treating the various collective coordinates carefully [9]. Collective coordinates describe zero frequency motions of solitons. They reflect the underlying symmetries of the theory that are broken by the soliton solution. Thus the 't Hooft Polyakov monopole has four collective coordinates : three corresponding to translations and one corresponding to isospin rotations in the unbroken U(1) direction. (There is no collective coordinate corresponding to global spatial rotations since this can be compensated by global isospin rotations.*) We will not be interested in translations of the monopole since they are irrelevant to the phenomenon we are interested in. Let us assume that the monopole is extremely heavy and work in the rest frame of the monopole with the origin of the coordinate system at the centre of the monopole. The fourth collective coordinate is the angle of global isospin rotation around the unbroken U(1) axis. Its conjugate momentum is electric charge. Thus a charged monopole, or dyon, can be pictured as a 't Hooft-Polyakov monopole rotating in internal space. Since this coordinate is an angle, its conjugate momentum, the electric charge, is quantized. Thus, in the quantum theory, the existence of dyons with quantized electric charge is a direct consequence of the existence of monopoles. To see how time dependent rotations in isospin space can give rise to electric charge, consider the following configuration in the temporal gauge:

$$A_0 = 0 \tag{2}$$

* There might be some confusion at this point : Is it possible then that one can do a similar (reverse) thing for the aforementioned U(1) i.e. can that be undone by a global space rotation ? The answer is no because the unbroken U(1) rotations are not, strictly speaking, global rotations. The generator is $\boldsymbol{r}\,.\vec{T}$ which depends on position.

$$A_i = h^{-1} A_i^{cl} h + h^{-1} \partial_i h$$

where $h = \exp(2i\omega\,\hat{r}.\vec{T})$, $\omega(0,t) = 0$ It should be emphasized that, since $\omega(r,t)$ depends on time, this configuration is not a gauge transformation of the monopole configuration. Let us now perform a gauge transformation with h^{-1} (which does not affect the electric field) to make the effect more transparent:

$$A_0 = h\partial_t h^{-1} = -2\partial_t \omega\,\hat{r}.\vec{T}$$

$$A_i = A_i^{cl}.$$

The electric field is $E_r = -2\partial_r \partial_t \omega\,\hat{r}\,\vec{T}$. This can be described as being due to an electric charge. Though we will not be interested in dyon solutions of the classical equations of motion, virtual dyonic excitations (due to quantum fluctuations) of the form (2) play a crucial role in the phenomenon we shall investigate.

This was a brief survey of some of the classical and quantum properties of monopoles. We now introduce fermions into the problem.

3. Monopoles and Fermions

In this section we discuss the effect of introducing fermions into the monopole sector and the behaviour of fermions when the gauge fields have non trivial topology. Section 1.4 presents some explicit and detailed calculations with isovector fermions and section 1.5 is a discussion of the physical interpretation of the results.

Before discussing the interactions of fermions with non-Abelian monopoles, let us first look at the (Abelian) Dirac monopole fermion system. This has been studied in detail by Kazama, Yang and Goldhaber [11] and by Goldhaber [12]. One of the interesting results is the fact that there is a non-vanishing helicity flip amplitude for a spin half particle scattering from a monopole. In the presence of a monopole there is an extra contribution to the angular momentum of a charged particle given by $\frac{qgm}{4\pi}\hat{r}$ where qg is the charge of the particle and m is the magnetic charge of the monopole. This angular momentum is stored in the electromagnetic field as can be inferred from the fact that there is a non-vanishing $\vec{E}\times\vec{B}$ field circulating around the axis joining the charged particle and the monopole. If q is a multiple of $\frac{1}{2}$, then $g\frac{m}{4\pi}$ = 1 from the Dirac quantization condition. Thus the electromagnetic contribution to J is q, and the lowest angular momentum state for an electron, which has spin $\frac{1}{2}$, has J = $|q|$ - 1/2. It was found that this state can only undergo scattering if accompanied by a helicity flip.* This is a bit surprising because normally in a static magnetic field, helicity is conserved. However, as was shown in ref [12] this is not true in the presence of a monopole. The monopole field is singular at the origin. If a particle is allowed to go through the singularity, unusual processes can take place because neither the Hamiltonian nor the helicity operator is self adjoint. In fact, the J = $|q|$ - 1/2 state is precisely the one whose wave function does not vanish at the origin because a

* Note that for a massless particle, helicity, which is the component of spin along the momentum, and chirality, Q_5, which is the eigenvalue of γ^5, are the same.

particle in this state does not see a repulsive 'centrifugal barrier'. If a particle goes through the monopole its helicity has to flip. See fig. 10. All we need is conservation of angular momentum. As the particle goes through the centre, the sign of $\hat{r}$ changes, and the term $\frac{qm}{4\pi}\hat{r}$ changes sign. Conservation of angular momentum then requires that the particle's helicity also changes sign. One can formalize this result by the following procedure (only an outline is given since Abelian monopoles are not the subject of this dissertation. Details can be found in refs. [11,12]). First regularize the singularity by adding an extra magnetic moment k to the fermion. On solving the wave equation, one finds that even the J = |q| - 1/2 wave function vanishes at the origin making the Hamiltonian well defined. Calculate scattering amplitudes and then take the limit $k \to 0$. This procedure has been shown to reproduce the result that the state with J = |q| - 1/2 contributes only to helicity flip amplitude.

The fact that chirality can flip in the presence of a monopole should not come as a surprise. Consider a state with a non Abelian monopole. (Note that the Abelian monopole can be considered to be the limiting case of a non Abelian monopole as $M_X \to \infty$ where M_X is the scale of SU(2) breaking). Witten has argued [14] that a rotation of 2π generated by the electric charge operator, Q, changes the winding number (defined to be $\frac{1}{64\pi^2}\int d^4x F_{\mu\nu}\tilde{F}^{\mu\nu}$) of the configuration by one unit. From the anomaly equation † it follows that this changes the chirality of the state. Thus the operator Q is a translation operator in the Q_5 representation, in the presence of a monopole,

i.e. we can represent Q by $m\frac{\partial}{\partial Q_5}$. This implies $[Q,Q_5] \approx m$, where m is the

† $\partial_\mu j^\mu{}_5 = \frac{g^2}{64\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}$, where J_5^μ is the axial current $\bar{\psi}\gamma^\mu\gamma^5\psi$ and $Q_5 = \int d^3x J_5^0$

• Another, equally heuristic, way of saying the same thing is that Q_5 changes the vacuum angle θ and therefore the charge of a state with a monopole since the monopole has charge $e\frac{\theta}{2\pi}$. Thus Q_5 is a translation operator in the Q representation. $Q_5 \approx m\frac{\partial}{\partial Q} \Rightarrow [Q_5, Q] \approx m$

magnetic charge of a monopole.* This shows that a state with a monopole can either have definite charge, or definite chirality, but not both. Charge, unlike chirality, costs energy since it couples to gauge fields. The ground state is thus an uncharged monopole which does not, therefore, have definite chirality.

Another dramatic consequence of having massless fermions is the decay of dyons. It was shown by Blaer, Christ and Tang [13] that dyons are unstable and decay into fermion pairs giving out charge and chirality. An easy way of seeing this is the following : the anomaly equation tells us that

$$\int d^3x \; E.B = \frac{dQ_5}{dt}$$

The L.H.S. is non zero for a dyon since it has $\vec{E} \approx E_r$ and $\vec{B} \approx B_r$. The R.H.S. has to be zero in any stationary state. This shows that a dyon is not a stationary state. Yet another aspect of the fermion monopole system was noticed by Callan [2] who showed that the charge, $e\frac{\theta}{2\pi}$ (due to the vacuum angle θ [14]), of a monopole gets spread out over an infinite volume in the presence of massless fermions. Therefore inside any arbitrarily large but finite volume containing the monopole, the net charge is zero. It was shown that the collective coordinate, conjugate to charge, which in the absence of fermions has the relatively simple dynamics of a rigid rotor, does not have this simple behaviour in the presence of fermions. One finds on solving the equations of motion that there are no solutions that correspond to a dyon.

Finally it was shown by Rubakov [1] and later by Callan [2] that the (uncharged) monopole is surrounded by a chirality and fermion number violating condensate. Rubakov likened this to the fermion condensate caused by instantons [8], which violates the $U(1)_A$ symmetry of QCD. In the presence of monopoles such effects are enhanced. In this and subsequent sections, it is this phenomenon of condensate formation that we shall discuss in detail. We shall see that many of the concepts that are useful in conventional instanton physics have their counterparts here and can be used to obtain a simple description of this phenomenon. Such analogies will also enable us to extend the results of Rubakov and Callan to higher fermion representations.

It is useful to recapitulate some well known facts from instanton physics [15]. Consider the theory of SU(2) gauge fields coupled to massless fermions. Quantum fluctuations of the gauge fields can give rise to configurations of non zero winding number.* It can be shown by using the anomaly equations that configurations with non-zero winding number take the trivial vacuum at Euclidean time $t = -\infty$ to a state with fermions at $t = +\infty$. Thus the amplitude $\langle 1|0\rangle=0$, where, by $|n\rangle$, we denote the vacuum with winding number n and no fermions. An amplitude of the form $\langle 1|\psi_1\psi_2(x)|0\rangle$† can be non zero, and as a function of 'x', can be interpreted as the amplitude that during the process $|0\rangle \to |1\rangle$ the fermions ψ_1 and ψ_2 appear at the point x (to be annihilated by the operator $\psi_1\psi_2(x)$). It can also be interpreted as a condensate of the composite field $\psi_1\psi_2$. This condensate violates chirality. All these facts follow from the anomaly equation $\partial_\mu j^\mu{}_5 \propto F_{\mu\nu}\tilde{F}^{\mu\nu}$. Integrating over space-time, we get $\Delta Q_5 \propto \nu$, where ν is the winding number of the configuration, and gives the number of chirality-violating ~~fermion states that~~ are generated. The functional integral formalism provides another way of looking at the same process. In a given background gauge field configuration, the index theorem (of which the anomaly equation is the local version) determines the number of normalizable zero modes of the Dirac operator. This determines those amplitudes of the form $\langle 1|\psi_1\psi_2(x)...|0\rangle$ that can be non zero. (The normalizability of a mode means that one fermion of that type is 'produced' between Euclidean time $t=-\infty$ and $t = +\infty$.) Thus we have different and consistent ways of looking at the same phenomenon. In doing calculations one typically uses a saddle point approximation around classical solutions of the (Euclidean) equations of motion, that have non zero winding number. These are the well known

---

*Winding number is defined to be $\frac{1}{32\pi^2}\int d^4x\, Tr F_{\mu\nu}\tilde{F}^{\mu\nu}$.

† The number of fermion fields to be inserted depends on the fermion content of the theory. We will come to this later.

instanton solutions. They have integer winding number, ν, and finite action $\frac{8\pi^2}{g^2}\nu$. This means that processes involving instantons are suppressed by the factor $\exp(-\frac{8\pi^2}{g^2}\nu)$.

One can use essentially the same picture to describe the formation of condensates around monopoles (unless otherwise stated, we restrict ourselves to the case of a 't Hooft Polyakov monopole in an SU(2) -> U(1) gauge theory) with the following modifications : a) In the presence of monopoles, configurations with non zero winding number exist, which have arbitrarily small action, hence there is no suppression factor. b) The winding number does not have to be an integer, it can be any real number. c) The back action of the fermions on the gauge field is important and we do not treat the fermions as propagating in a fixed background.

Let us look at each of these modifications a bit more closely : a) Witten showed [14] that in the monopole sector one can connect vacua $|M;n>$ with different winding number by means of histories of the form given in equation (2)

$$A_0 = 0 \tag{2}$$

$$A_i = h^{-1}A_i^{cl}h + h^{-1}\partial_i h$$

with $h = \exp(2i\omega(r,t)\hat{r}.\vec{T})$. $\hat{r}.\vec{T}$ is the unbroken U(1) generator. To avoid a singularity at the origin we require $\omega(0,t) = 0$. When ω is time independent, this is an allowed gauge transformation because it does not rotate the Higgs fields. Consider ω as shown in fig. 11. The history 'h' that has $\omega(r,-\infty) = 0$ and $\omega(r,+\infty)$ as shown in the figure, has winding number k. This can be seen as follows (asuume that we have a Dirac monopole, for simplicity):

$$\nu \approx \int F_{\mu\nu}\tilde{F}^{\mu\nu}d^4x$$

$$= \int d^3x \int dt\, \varepsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma}$$

$$= \int d^3x \int dt\, \varepsilon_{\theta\varphi r t}F^{\theta\varphi}F^{rt}$$

$$= \int d^3x \int dt B.\partial_r \partial_t \omega = \int d^3x B.\, \partial_r(\omega(r,+\infty) - \omega(r,-\infty))$$

with ν = winding number. Integrating by parts, and using $\partial_i B_i = g\,\delta^3(r)$ as well as the boundary condition on ω

$$\nu \quad = -\int d^3x\, \delta^3(r)[\omega(r,+\infty) - \omega(r,-\infty)] + \int d^2S.B[\omega(\infty,+\infty) - \omega(\infty,-\infty)]$$

and we get winding number m k (we have used $\int B.dS = m$, magnetic charge). It is clear from this that the presence of the monopole is crucial. Furthermore, as T-> ∞ and ω becomes time independent, h reduces to an allowed gauge transformation. This makes one suspect that there is no potential barrier for this process and that the cost in action comes entirely from the time derivatives in the kinetic term. This can also be shown by a simple scaling argument. We require

$$\int_{-T}^{+T} dt \int d^3x\ E.B \approx 1.$$

In the monopole sector there is a constant B field and so E_r scales like 1/T. Furthermore, it can also be shown that the E field is purely radial: $E_i = F_{0i} = \partial_t A_i - \partial_i A_0 - [A_0, A_i]$. $A_i = A_i^{cl} = \varepsilon_{aij} \frac{\tau^a r_j}{2ir}$ and $A_0 = i\partial_t \omega \hat{r}.\vec{\tau}$ This gives $\partial_i A_o = i\partial_i \partial_r \omega \hat{r}.\vec{\tau} + i\partial_t \omega \frac{P_{ij}}{r}\tau^j$ where P_{ij} is the projection operator $(r_i r_j - \delta_{ij})$.

$$[A_i, A_0] = i\partial_t \omega[\varepsilon_{aij}\tau^a \frac{r_j}{2ir}, \hat{r}.\tau] = -i\frac{(P_{ij}\tau^j)}{r}\partial_t \omega$$

This gives $F_{0i} = -i\partial_i \partial_t \omega \hat{r}.\vec{\tau}$. Since $\omega(r,t)$ depends only on 'r' it follows that the E field is purely radial. If we impose $\partial_r \partial_t \omega \approx O(\frac{1}{T^2})$, the action for this history is finite and $\approx |\vec{E}|^2 T \approx \frac{1}{T^2}T \approx \frac{1}{T}$. Thus as T->$\infty$, the action is zero, as conjectured. (We have subtracted out the monopole energy). This is to be contrasted with the instanton case, where E, B $\approx \frac{1}{\sqrt{T}}$ and the action goes as $(E^2 + B^2)T \approx \frac{T}{T} \approx$ constant. Thus, unlike the pure SU(2) gauge theory, where, for configurations of nonzero winding number, there is a lower bound on the action [15] of $8\frac{\pi^2}{g^2}\nu$ (ν = winding number , g = coupling constant) which is realized by the instanton solution in the monopole sector, one has configurations of nonzero winding number with arbitrarily small action. In the presence of such configurations, massless fermions have zero modes, as dictated by the index theorem [16], and this results in condensate formation just as in the trivial vacuum sector.

b) In the monopole sector one has finite action field configurations with non integer winding number $\int d^4x\, TrF_{\mu\nu}\tilde{F}^{\mu\nu}$. It was shown in the previous paragraph that a history of the type shown in fig. 11 has winding number k. There is no reason for k to be an integer. This is not true in a pure gauge theory where all finite action configurations have integer winding number. To see this, let us study the instanton solutions a bit more carefully :

$$A_\mu = \frac{\sigma^{\mu\nu}x^\nu}{(x^2+1)} \quad : \quad A^4 = i\sigma^i \frac{\tau_i}{(x^2+1)} \qquad A_i = \sigma^{ij}\frac{\tau_j}{(x^2+1)} \tag{3}$$

where $\sigma^{ij} = \frac{1}{2}\varepsilon^{ijk}\sigma^k$, $\sigma^{i4} = \frac{1}{2}\sigma^i$ σ^i being the Pauli matrices and $x = \sqrt{x^\mu x_\mu}$. To compare it with the configurations of the form in eq. 2 it is useful to transform the instanton solution to the temporal gauge $A^4 = 0$. A gauge transformation 'h' (not to be confused with the h of eq. 2) that accomplishes this is determined. h satisfies

$$h^{-1}\partial_4 h = A_4$$

Assuming $h = \exp(i\omega\hat{r}.\vec{\sigma})$, we get $\partial_t\omega = \dfrac{1}{\vec{x}^2 + x_4^2 + 1}$

$$\omega = \frac{1}{\sqrt{\vec{x}^2+1}}\tan^{-1}\frac{x_4}{\sqrt{\vec{x}^2+1}} + const.$$

It is convenient to rewrite the instanton solution (3) as

$$A_\mu = \frac{x^2}{1+x^2}g^{-1}\partial_\mu g \quad ; \quad g = \frac{x_4 - i\vec{x}.\vec{\sigma}}{\sqrt{x^2}} \qquad A_4 = \frac{x^2}{1+x^2}g^{-1}\partial_4 g$$

In the temporal gauge

$$A_4^T = 0$$

$$A_i^T = h\partial_i h^{-1} + hA_i h^{-1} \quad = \quad \frac{x^2}{1+x^2}\, hg^{-1}\partial_i(hg^{-1})^{-1} + \frac{1}{1+x^2}\, h\partial_i h^{-1} \tag{4}$$

Here the superscript T denotes 'temporal gauge'. (Note that as $\vec{x}$ or $x_4 \to \infty$, $A_i^T = hg^{-1}\partial_i(hg^{-1})^{-1}$ which is pure gauge. This is expected, since in the vacuum sector all configurations have to reduce to pure gauge at infinity if the action is to be finite). Furthermore, using

$$g^{-1} = \exp\,(i\,(\frac{\pi}{2} - \tan^{-1}\frac{x_4}{\sqrt{\vec{x}^2}})\vec{x}.\vec{\sigma}\,) \tag{5}$$

we get

$$hg^{-1} = \exp(i\,[\,\frac{\pi}{2} + \frac{\vec{x}}{\sqrt{1+\vec{x}^2}}\tan^{-1}\frac{x_4}{\sqrt{1+\vec{x}^2}} - \tan^{-1}\frac{x_4}{\sqrt{\vec{x}^2}}\,]\,\vec{x}.\vec{\sigma})$$

It is also clear that as $x \to \infty$, $h\,g^{-1}$ becomes time independent, as required for a pure gauge transformation in the temporal gauge.

$h\,g^{-1}$ is plotted in figure 12 along with the dyonic excitations of eq. 2 for comparison. For large x_4 and large $\vec{x}$, hg^{-1} contains all the information about the instanton configuration. This is also sufficient to determine the winding number. Also at $\vec{x} = 0$, the piece $h\partial_i h^{-1}$ is zero so we can ignore it. At $\vec{x} = 0$, $x_4 \approx 0$, $hg^{-1}\partial_i(hg^{-1})^{-1}$ is singular because it changes discontinuously. However it is multiplied by $\frac{x^2}{1+x^2}$ which smooths this. The important point to note in the first column of fig. 12 is that the winding number arises due to the change in the value of ω from 0 to $-\pi$ at the origin and not at spatial infinity where its value is constant. Since the group element is well defined at the origin only if it is a constant (± 1), it follows that ω has to change in steps of π. In complete contrast, as the figures in the second column indicate, the winding number in the monopole case arises due to changes in ω at infinity, and thus is unconstrained. When fermions are added, the situation changes since the winding number also counts fermion states via the anomaly equation. This quantizes the winding number, but still allows half integral values in certain cases.

c) We have already seen in the case of a dyon that fermions have a very significant effect on the dynamics of the gauge field. Thus, the calculation will be organized as follows. The fermions are first integrated out. This modifies the kinetic term for the gauge field. After that, the different gauge field configurations are integrated over. Thus, no fixed background gauge field configuration is put in by hand. Once the integration is done it can be checked, a posteriori, that the dominant configurations have winding numbers consistent with the index theorem.

4. Isovector Fermions

In this section we analyse the monopole fermion system with the fermion in the adjoint representation. The monopole is the usual 't Hooft Polyakov one

$$A_0^{cl} = 0$$

$$A_i^{cl} = -igA_i^a T^a = \varepsilon_{aij} T^a r_j \frac{F(r)}{ir}$$

$$\varphi^{cl} = c\hat{r}.\vec{T}\,H(r).$$

F(r) and H(r) are the functions shown in fig. 9. $<\varphi> = c$, far away from the monopole and the mass M_X of the heavy bosons $\approx$ gc. The core radius r_M is $O(\frac{1}{M_X})$. To simplify calculations, we take the limit $M_X \to \infty$ or $r_M \to 0$. If the gauge bosons are much heavier than the fermions, this is a good approximation. The effects of the monopole persist even in this limit, which means that they are not suppressed by factors of $1/M_X$, as one might suppose naively. The effects of the monopole manifest itself in certain boundary conditions that the fermion fields have to satisfy at the centre. Even after the limit $r_M \to 0$ is taken, these boundary conditions convey information about the core to the outside. Therefore we shall first write down the Lagrangian keeping a finite core radius, obtain the boundary conditions, and then let the core radius r_M go to zero.

The excitations of the gauge fields being considered are of the form (2). The background, as well as these excitations, are invariant under $\vec{J} = \vec{L} + \vec{S} + \vec{T}$ (the total angular momentum). It is convenient to employ a Fourier decomposition of the fermion fields. The operators $\vec{J}.\vec{J}$, J_3, $\hat{r}.\vec{S}$ and $\hat{r}.\vec{T}$ can be diagonalized simultaneously. The eigenvalues are denoted by J(J+1), M, σ,τ, respectively, and the eigenfunctions by $\Psi^J_{M\sigma\tau}$. The fermion field can be expanded in terms of these eigenfunctions:

$$\psi(\vec{r},t) = \sum_{J,M,\sigma,\tau} U^J_{M\sigma\tau}(r,t)\Psi^J_{M\sigma\tau}\ (\theta,\varphi)$$

After transforming to a gauge where $A_i = A_i^{cl}$, the Lagrangian becomes

$$L = \int \psi^{+}(\partial_t + A_0 + \frac{1}{r}D_\Omega + 2i\sigma(\partial_r + \frac{1}{r}))\psi d^3x \qquad (6)$$

where $\frac{1}{r}D_\Omega = i\sigma_i(\delta_{ij} - \hat{r}_i\hat{r}_j)(\partial_j + A_j) - i\frac{\vec{\partial}.\hat{r}}{r}$ is the angular part of the Dirac operator. To derive the exact Lagrangian requires a certain amount of algebra which has been outlined in Appendix A. Only the final result is given here. The J=1/2 and J > 1/2 pieces are written separately:

$$S = S_{\frac{1}{2}} + \sum_{J>\frac{1}{2}} S_J$$

$$S_{\frac{1}{2}} = S^{I}_{\frac{1}{2}} + S^{II}_{\frac{1}{2}} \qquad (7)$$

$$S^I = \sum_{M\sigma\tau} \int [\, U^*_{M\sigma\tau}(\partial_t + A_t + i\sigma(\partial_r + \frac{F}{r}))\, U_{M\sigma\tau} + \quad \frac{F}{r}\sqrt{1-\tau^2}\, U^*_{M\sigma\tau} U_{M-\sigma\tau}]\ r^2 dr dt$$

For J = 1/2, M takes the values $\pm\frac{1}{2}$, σ takes the values $\pm\frac{1}{2}$ and τ takes values 0, ± 1, with $|\sigma + \tau| \le 1$. Thus for each M, $(\sigma,\tau) = (\frac{1}{2},0), (\frac{-1}{2},0), (\frac{-1}{2},1), (\frac{1}{2},-1)$.

$$S^I = \sum_{M\sigma\tau} \int (1-F) \Big\{ U^*_{M\sigma\tau}\ (i\frac{\sigma}{r})\, U_{M\sigma\tau} + \sqrt{1-\tau^2}\ U^*_{M\sigma\tau} U_{M\sigma\tau}$$

$$+ \frac{\sqrt{2}i}{r}(U^{M*}_{\frac{-1}{2}0} U^{M}_{\frac{1}{2}-1} + U^{M*}_{\frac{1}{2}0} U^{M}_{\frac{-1}{2}1} + h.c.) \Big\}\ r^2 dr dt$$

$S^{II}_{\frac{1}{2}}$ vanishes asymptotically, i.e. as F -> 1.

$$S_{J>\frac{1}{2}} = \int [\, U^{J*}_{M\sigma\tau}(\partial_t + 2i\tau\partial_t\omega + 2i\sigma(\partial_r + \frac{1}{r}))\ U^{J}_{M\sigma\tau} \qquad (8)$$

$$+ \frac{1}{r} U^{J*}_{M\sigma\tau} U^{J}_{M-\sigma\tau} \sqrt{J(J+1) + \frac{1}{4} - \tau^2}\ ]\, r^2 dr dt$$

In equation (8), the limit F -> 1 has been taken. For J > 1/2, $\sqrt{J(J+1) + \frac{1}{4} - \tau^2} > 0$ because for isovector fermions $|\tau| \le 1$. This term acts like a centrifugal barrier and makes the wave function vanish at the origin. It will be argued later that for the existence of a zero mode and consequent condensate formation, it is essential that the wave function be non vanishing at the origin. Thus, for the moment, we concentrate on the J = 1/2 sector.

For simplicity we shall approximate F(r) by 1 outside a radius r_M. When M_X is very large this should be a good approximation because the distance over which F(r) changes from 0 to 1 is $\approx \frac{1}{M_X}$ which is << $\frac{1}{m_f}$ the Compton wavelength of the fermions, the only other scale in the problem.

At the origin, F=0, and the Lagrangian density $L^{II}_{\frac{1}{2}}$ where $S^{II}_{\frac{1}{2}} = \int L^{II}_{\frac{1}{2}} d^4x$ contains singular terms that blow up as 1/r, and the Hamiltonian density also blows up. To prevent this, some restrictions have to be imposed on the fields $U_{M\sigma T}$. The 1/r pieces can be collected and written as follows:(an index M that takes the values $\pm\frac{1}{2}$ is understood on all fields in the J = $\frac{1}{2}$ sector)

$$\begin{pmatrix} U^*_{\frac{1}{2}0} & U^*_{-\frac{1}{2}0} & U^*_{\frac{1}{2}-1} & U^*_{-\frac{1}{2}1} \end{pmatrix} \begin{bmatrix} 1 & -i & 0 & \sqrt{2} \\ -i & -1 & \sqrt{2} & 0 \\ 0 & -\sqrt{2} & 1 & 0 \\ -\sqrt{2} & 0 & 0 & -1 \end{bmatrix} \begin{pmatrix} U_{\frac{1}{2}0} \\ U_{-\frac{1}{2}0} \\ U_{\frac{1}{2}-1} \\ U_{-\frac{1}{2}1} \end{pmatrix}$$

We then require that at the origin the vector $\vec{U} = \begin{bmatrix} U_{\frac{1}{2}0} \\ U_{\frac{-1}{2}0} \\ U_{\frac{1}{2}-1} \\ U_{\frac{-1}{2}1} \end{bmatrix}$

be parallel to the null eigenvector which is :

$$\begin{bmatrix} 1 \\ i \\ \sqrt{2}i \\ -\sqrt{2} \end{bmatrix}$$

This can more conveniently be written as three boundary conditions :

$$U_{\frac{1}{2}-1}(0,t)+iU_{\frac{-1}{2}1}(0,t)=0 \tag{9a}$$

$$iU_{\frac{1}{2}0}(0,t)-U_{\frac{-1}{2}0}(0,t)=0 \tag{9b}$$

$$[U_{\frac{1}{2}-1}(0,t)-iU_{\frac{-1}{2}1}(0,t)]=\sqrt{2}[iU_{\frac{1}{2}0}(0,t)+U_{\frac{-1}{2}0}(0,t)] \tag{9c}$$

On studying $S^{I}_{\frac{1}{2}}$ and $S^{II}_{\frac{1}{2}}$, it is clear that in the limit $r_M \to 0$, the only remnant of the coupling between the $\tau = 1$ and $\tau = 0$ fields is the boundary condition (9c).

The next simplifying step is to take this limit $r_M \to 0$. Define new 'tilde' fields $\tilde{U}(r,t)$ generically in terms of the corresponding U(r,t) variables by

$$\frac{1}{\sqrt{4\pi}}\frac{1}{r}\tilde{U}\exp\left(\int_{\infty}^{r}\frac{1-F(r)}{r}\,dr\right) = U(r,t)$$

For $r < r_M$, $\tilde{U} = U$ and for $r > r_M$ $\tilde{U} = rU$. Thus $\tilde{U}$ interpolates smoothly between the fields U(r,t) inside the core and rU(r,t) outside, the latter of which is the convenient variable to work with. The Dirac equation written in terms of $\tilde{U}$ is smooth at $r = r_M$ and the boundary condition (9) is imposed on $\tilde{U}(r,t)$ at $r = r_M$, and in the limit $r_M \to 0$, this is identical to (9).

Anticipating this limit we concentrate from now onwards exclusively on the region $r > r_M$. The $\tau = \pm 1$ and $\tau = 0$ fields decouple in this region except for the boundary condition (9c). The $\tau = \pm 1$ Lagrangian is

$$L = \sum_M \int \tilde{U}\tau^3(\partial_t - 2i\tau^2(\partial_t\omega) + \tau^1\partial_r)\tilde{U}\,drdt \tag{10}$$

where

$$\tilde{U} = \begin{pmatrix}\tilde{u}_1\\ \tilde{u}_2\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\tilde{U}_{\frac{-1}{2}1} + i\tilde{U}_{\frac{1}{2}-1}\\ i\tilde{U}_{\frac{-1}{2}1} + \tilde{U}_{\frac{1}{2}-1}\end{pmatrix}$$

and boundary condition (9a) is $\tilde{U}_2(r_M) = 0$. This is exactly the same Lagrangian and boundary condition as in the isospinor case with two flavours which was worked out by Rubakov and Callan, except that the interaction term involves $2\partial_t\omega$ instead of $\partial_t\omega$. This corresponds to the fact that the charge of the charged isovector fermions is twice that of isospinor fermions. This Lagrangian, in fact, describes a massless Schwinger model with two fermion 'flavours' (M= $\pm\frac{1}{2}$) with $\tau^3 \approx \gamma^0$, $\tau^1 \approx \gamma^1$, $\tau^2 \approx \gamma^5$. The $\tau = \pm 1$ (J=1/2) part, by itself, is thus exactly solvable [17].

The $\tau = 0$ Lagrangian is

$$\int\left\{\tilde{U}^{*}_{\frac{1}{2}0}(\partial_t+i\partial_r)U_{\frac{1}{2}0} + \tilde{U}^{*}_{\frac{-1}{2}0}(\partial_t-i\partial_r)U_{\frac{-1}{2}0} + \frac{1}{r}\tilde{U}^{*}_{\frac{1}{2}0}U_{\frac{-1}{2}0} + \frac{1}{r}\tilde{U}^{*}_{\frac{-1}{2}0}U_{\frac{1}{2}0}\right\}drdt. \quad (11)$$

We define

$$v_1^M = \frac{1}{\sqrt{2}}(U^M_{\frac{1}{2}0} + iU^M_{\frac{-1}{2}0})$$
$$v_2^M = \frac{1}{\sqrt{2}}(U^M_{\frac{1}{2}0} - iU^M_{\frac{-1}{2}0})$$

so that the boundary condition (9b) simply becomes $v_1^M(r_M,t) = 0$.

Without boundary condition (9c), the $\tau = 0$ and $\tau = \pm 1$ sectors are completely decoupled, in which case the Green's functions for the fermions are 2 x 2 matrices for each sector ($\tau = 0$ and $\tau = \pm 1$). However (9c) requires that both sectors be treated simultaneously, i.e. with 4 x 4 matrices, and this means that the similarity of the $\tau = \pm 1$ sector to the Schwinger model cannot be exploited. At first sight, this problem appears to be considerably less tractable than the isospin 1/2 case. However, it will be shown that if r_M is small enough, errors made in treating the two sectors independently are negligible. This is because the terms that couple the two sectors in the 4 x 4 Green's function matrix vanish as $\sim \frac{r_M}{r}$. Thus as long as questions about short distance structure, $r \approx r_M \approx \frac{1}{M_X}$, are not being asked, these terms can be ignored and the fermion Green's function becomes block diagonal as $r_M \to 0$.

Consider the behaviour of the Green's function for the v_1^M, v_2^M system just outside the core. The equation of motion for the Green's function there can be written as

$$\begin{bmatrix} \partial_t & i(\partial_r+\frac{1}{r}) \\ i(\partial_r-\frac{1}{r}) & \partial_t \end{bmatrix}\begin{bmatrix} \partial_t G_1 & -i(\partial_r+\frac{1}{r})G_2 \\ -i(\partial_r-\frac{1}{r})G_1 & \partial_t G_2 \end{bmatrix} = \begin{pmatrix} \delta(r-r')\delta(t-t') & 0 \\ 0 & \delta(r-r')\delta(t-t') \end{pmatrix} \quad (12)$$

The Green's function has been parametrised by the functions G_1 and G_2. From (10) we obtain :

$$(\partial_t^2+\partial_r^2)G_1=\delta(r-r)\delta(t-t)$$

$$(\partial_t^2+\partial_r^2-\frac{2}{r^2})G_2=\delta(r-r)\delta(t-t)$$

so that $G_2 \sim r^2$ as $r_M \to 0$, and to ensure that $v_1^M(0,t)=0$, choose the solution where $G_1 \sim r$ as r ->0. (In the above equation, excitations of the gauge field of the form (2) that add terms of the form $\partial_t \omega$ to the equation of motion have not been considered. This is justified by the fact that the parameter ω of equation (2) satisfies $\omega(0,t)=0$ and therefore in the region of interest $r \approx r_M \approx 0$ one can safely neglect the effect of such interactions). Notice however that $v_2^M(0,t)$ also becomes zero with this choice of Green's function. Thus, while (9b) has been satisfied, something still has to be done about (9c) according to which $\lim_{r_M \to 0}(v_2^M(r_M)=\frac{1}{\sqrt{2}}\tilde{U}_1^M(r_M))$. If left as such, this would force $\tilde{U}_1^M(r_M)=0$, and since already $\tilde{U}_2^M(r_M)=0$, all fields would vanish at the origin. Since it will be argued later that a zero mode is necessarily non vanishing at the origin, the conclusion would seem to be that there is no condensate formation in the isovector case. However, we now show that this conclusion is not correct because the Green's function given above is not the most general that can be written down. One can add, as part of the full Green's function for the combined system of $\tau=0$ and $\tau=\pm 1$ fermions, a solution to the homogeneous equation (12) that enables (9c) to be satisfied in a trivial manner.

For convenience the t-variable is Fourier transformed and the frequency denoted by z. The functions

$$\begin{bmatrix} g(z,r,r_M) \\ f(z,r,r_M) \end{bmatrix} = \begin{bmatrix} r_M \dfrac{e^{-z(r-r_M)z}}{1+zr_M} \\ \dfrac{r_M}{r} e^{-z(r-r_M)} \dfrac{1+zr}{1+zr_M} \end{bmatrix}$$

satisfy the system (12) with the R.H.S. put equal to zero. Also

$$f(z,r_M,r_M)=1 \qquad g(z,r_M,r_M)=\frac{zr_M}{1+zr_M}.$$

Keeping these properties of f and g in mind, consider the Green's function for the combined $(\tilde{U}_1,\tilde{U}_2,v_1,v_2)$ system:

$$\begin{bmatrix} H_1 & H_2 & 0 & 0 \\ H_3 & H_4 & 0 & 0 \\ gA_1(z,r,t) & gA_2(z,r,t) & \partial_t G_1 & -i(\partial_r + \frac{1}{r})G_2 \\ fA_1(z,r,t) & fA_2(z,r,t) & -i(\partial_r - \frac{1}{r}) & \partial_t G_2 \end{bmatrix}$$

where H_i are the Green's functions for the $\tau = \pm 1$ system, which we assume satisfy the boundary condition (9a): $H_3(r_M,z,r',t') \approx H_4(r_M,z,r',t) \approx O(r_M) \approx 0$ and

$$A_1(z,r,t) = \frac{1}{\sqrt{2}} H_1(r_M,z,r,t)$$

$$A_2(z,r,t) = \frac{1}{\sqrt{2}} H_2(r_M,z,r,t)$$

At r = r_M , $G(r,z,r',t')$ becomes

$$\begin{bmatrix} H_1 & H_2 & 0 & 0 \\ \approx O(r_M) & \approx O(r_M) & 0 & 0 \\ \frac{\omega r_M}{1+\omega r_M}\frac{1}{\sqrt{2}}H_1 & \frac{\omega r_M}{\sqrt{2}1+\omega r_M}H_2 & \approx O(r_M) & \approx O(r_M) \\ \frac{1}{\sqrt{2}}H_1 & \frac{1}{\sqrt{2}}H_2 & \approx O(r_M) & \approx O(r_M) \end{bmatrix}$$

where the arguments (r_M,z,r',t') for H_1 and H_2 are understood. Now H_1 and H_2 are Dirac Green's functions which behave as $\frac{1}{z}$ as z-> ∞*. Therefore we can safely say that $\frac{zr_M}{\sqrt{2}(1 + zr_M)}H_{1,2}(r_M,z,r',t') \rightarrow 0$ as $z \rightarrow \infty$. This Green's function thus satisfies all the boundary conditions (9). The functions f and g have another interesting property : in the limit $r_M \rightarrow 0$, they vanish for any value of r > 0. Thus the Green's function also becomes block diagonal in this limit and the $\tau = 0$ and $\tau = \pm 1$ sectors are decoupled.

Let us summarize briefly, what we have done so far in the isovector case. First, the core of the monopole is removed from the problem by imposing some boundary conditions at $r = r_M$ so that the equations of motion need be

* This is strictly true for the free Green's function. The short distance behaviour of the Green's function is not affected by interactions.

solved only in the region $r > r_M$. This approximation becomes exact in the limit $r_M \to 0$. Second, the Green's functions are modified by the addition of homogeneous solutions of the equations of motion in such a way that all the boundary conditions at $r = r_M$ are satisfied. Furthermore, in the region $r > r_M + \varepsilon$, where $\varepsilon > 0$, but small, and ε goes to zero as $r_M \to 0$, these homogeneous solutions are negligible and may be ignored. In the limit $r_M \to 0$, $\varepsilon \to 0$, this approximation also becomes exact.In fig. 13 the situation for finite r_M is depicted. The $\tau = \pm 1$ zero mode (condensate) is non vanishing at the origin. Since the core of the monopole is a region where SU(2) is unbroken, one expects that there should be a certain amount of the $\tau = 0$ condensate as well. This is also what the boundary conditions seem to indicate. However the $\tau = 0$ condensate vanishes quite rapidly over a distance of $O(r_M)$. Now we can proceed to treat the $\tau = 0$ and $\tau = \pm 1$ parts separately.

The $\tau = 0$ Lagrangian outside the core was given in equation (11). The $\tau = 0$ fermions do not couple to the gauge field excitations, and hence do not have any zero modes. They do not form any condensate, except in a region around the core of extent (r_M), as mentioned in the previous paragraph.

Consider the $\tau = \pm 1$ Lagrangian (equation (10)):

$$L = \sum_M \int \bar{\tilde{U}} \tau^3 (\partial_t - 2i\tau^2(\partial_t \omega) + \tau^1 \partial_r) \tilde{U} \, drdt$$

with boundary condition $\tilde{U}_2^M(0) = 0$*. We can proceed just as in the isospinor case. The free fermion equation is

$$(\tau^3 \partial_t + \tau^1 \partial_r) G(\xi, \xi') = \delta(\xi - \xi')$$

where $\xi = (r, t)$. The solution to this is $G_0(\xi, \xi') = \dfrac{1}{2\pi} \dfrac{\tau^3(t - t') + \tau^1(r - r')}{(r - r')^2 + (t - t')^2}$. The boundary condition $\tilde{U}_2(0, t) = 0$ has to be satisfied, so the same boundary condition is imposed on the Green's function. It can be seen that the Green's

* The two values of M should be considered to be flavour indices for the purposes of this calculation.

function $G_0((r - r'),(t - t')) + G_0((r + r'),(t - t'))\tau^3$ automatically satisfies this. As in the Schwinger model, one can now solve for the Green's function in the presence of a gauge field excitation parametrised by $\omega(r,t)$. The solution is [17,1,2]

$$G(\xi,\xi') = \exp[2i\tau^2(\alpha(\xi)) + 2(\sigma(\xi) - \sigma(\xi'))]G_0(\xi,\xi')\exp 2i\tau^2\alpha(\xi') \qquad (13)$$

where

$$\sigma(\xi') = \int \Box^{-1}\partial_r\partial_t\omega(r,t)drdt \qquad (14)$$

$$\alpha(\xi') = \int_0^{r'} \partial_t\sigma(r,t)dr$$

$$\Box^{-1} = \Box^{-1}((r - r'),(t - t')) + \Box^{-1}((r + r'),(t - t'))$$

i.e. it is the Green's function for a scalar field in two dimensions that satisfies $\Box\varphi = 0$ and has boundary condition $\partial_r\varphi(0,t) = 0$. Also note that $\alpha(0) = 0$. This is required because imposing $(1-\tau^3)G_0(0,t,r',t') = 0$ is not sufficient for $(1-\tau^3)G(0,t;r',t') = 0$ to be satisfied because the matrix $\alpha(0)\tau^2$ in the exponent can mix the upper and lower components. $\Box^{-1}(r,t) = \frac{1}{4\pi}\log\mu^2(r^2+t^2)$ where μ is an arbitrary mass scale. Now the fermions can be integrated out exactly. This is worked out in Appendix B. The result is

$$S_{fermi} = \frac{4}{\pi}\int(\partial_\mu\sigma)^2 drdt. \qquad (15)$$

The corresponding expression in the isospinor case with two flavours was

$$S_{fermi} = \frac{1}{\pi}\int(\partial_\mu\sigma)^2 drdt. \qquad (16)$$

The factor of 4 is easily understood. The gauge field excitation term in the isovector case is $2(\partial_t\omega)$, versus $(\partial_t\omega)$ in the isospinor case. This reflects the relative charges of the fermions in the two cases. The number of flavours is two (M $= \pm\frac{1}{2}$), just as in the isospinor calculation. Thus simply replacing σ by 2σ gives us the required factor of four.

To get the full action, the gauge field kinetic term has to be added. Ignoring the constant background of the monopole, we are left with the electric field due to the excitations which is given by $\partial_r \partial_t \omega$. Thus the term to be added is:

$$\frac{1}{g^2}\int d^4x \, |E|^2 = \frac{1}{g^2}\int 4\pi r^2 dr dt \, (\partial_r \partial_t \omega)^2.$$

Using (14), this can be written as

$$\frac{1}{g^2}\int 4\pi r^2 (\Box \sigma)^2 dr dt.$$

Thus we finally get:

$$S_{eff}(\sigma) = 4\frac{\pi}{e^2}\int r^2 (\Box \sigma)^2 dr dt \; + \; \frac{4}{\pi}\int (\partial_\mu \sigma)^2 dr dt. \tag{17}$$

The propagator for the σ field and some of its asymptotics are given in Appendix C.

We can now calculate the expectation value of an operator of the form

$$U^{\frac{1}{2}}_{\frac{1}{2}-1} U^{-\frac{1}{2}}_{\frac{-1}{2}1} - U^{-\frac{1}{2}}_{\frac{1}{2}-1} U^{\frac{1}{2}}_{\frac{-1}{2}1}. \tag{18}$$

Letting

$$U^{m} = \begin{pmatrix} u^{m}_{1} \\ u^{m}_{2} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} U^{m}_{\frac{-1}{2}1} + iU^{m}_{\frac{1}{2}-1} \\ iU^{m}_{\frac{-1}{2}1} + U^{m}_{\frac{1}{2}-1} \end{pmatrix}$$

$$= \frac{(1+i\tau^1)}{\sqrt{2}} \begin{pmatrix} U^{M}_{\frac{-1}{2}1} \\ U^{M}_{\frac{1}{2}-1} \end{pmatrix}$$

(18) becomes $U^{\frac{1}{2}T} U^{-\frac{1}{2}}$. In terms of the tilde variables, $\widetilde{U}^{\frac{1}{2}}$, (see equation 10) this becomes

$$\frac{1}{r^2}\widetilde{U}^{\frac{1}{2}T}(r,t)\widetilde{U}^{-\frac{1}{2}}(r,t) \equiv Q(r,t). \tag{19}$$

Thus we have to calculate

$$\langle M | Q(r,t) | M \rangle. \tag{20}$$

First evaluate

$$<M\,|\,Q(\tau,t)Q^{\dagger}(\tau',t')\,|\,M>. \tag{21}$$

In the limit (t - t') -> ∞, this should become, by the cluster decomposition property,

$$<M\,|\,Q(\tau,t)\,|\,M><M\,|\,Q^{\dagger}(\tau',t')\,|\,M>. \tag{22}$$

(21) is easier to evaluate because it can be written directly in terms of the fermion Green's functions (13) which are known exactly.

$$Q(\tau,t)Q^{\dagger}(\tau',t')= \frac{1}{\tau^2}\frac{1}{\tau^2}:\tilde{U}^{1T}\tilde{U}^{2}::\tilde{U}^{2\dagger}\tilde{U}^{1*}: \tag{23}$$

$$=\frac{1}{\tau^2}\frac{1}{\tau^2}\tilde{U}^{1T}\tilde{U}^{1*}\;\;\tilde{U}^{2}\tilde{U}^{2\dagger}$$

$$=\frac{1}{\tau^2}\frac{1}{\tau^2}[\tilde{U}^{1\dagger}\tilde{U}^{1}]^{T}[\tilde{U}^{2}\tilde{U}^{2\dagger}]$$

$$=\frac{1}{\tau^2}\frac{1}{\tau^2}Tr[G^{1T}(\tau,t;\tau',t')G^{2}(\tau,t;\tau',t')]$$

$$<M\,|\,Q(\tau,t)Q^{\dagger}(\tau',t')\,|\,M> = \frac{1}{\tau^2}\frac{1}{\tau^2}\int[d\sigma]Tr[G^{1T}G^{2}]e^{-S_{eff}(\sigma)}$$

Substituting the expressions for G from equation (13) and $S_{eff}(\sigma)$ from equation (17) one finds:

$$<M\,|\,Q(\tau,t)Q(\tau',t')\,|\,M> = \frac{1}{\tau^2}\frac{1}{\tau^2}\int[d\sigma]TrG_0^{1T}(\xi,\xi')G_0^{2}(\xi,\xi')e^{-S_{eff}(\sigma)+4\sigma(\xi)-4\sigma(\xi')} \tag{24}$$

Using $G_0 = \dfrac{1}{2\pi}\dfrac{\tau^3(t-t')+\tau^1(r-r')}{(t-t')^2+(r-r')^2} + \dfrac{1}{2\pi}\dfrac{(t-t')-i\tau^2(r+r')}{(t-t')^2+(r+r')^2}$, we get

$$Tr\;\;G_0^{T}G_0 = \frac{1}{2\pi^2}\left[\frac{1}{(t-t')^2+(r-r')^2} + \frac{1}{(t-t')^2+(r+r')^2}\right] \tag{25}$$

To evaluate

$$\int[d\sigma]e^{-S_{eff}(\sigma)+4\sigma(\xi)-4\sigma(\xi')},$$

observe that the terms linear in σ can be treated as source terms with the source being $J(\xi'') = 4\delta(\xi''-\xi) - 4\delta(\xi''-\xi')$. Then the result of doing the Gaussian integral is just:

$$e^{\frac{1}{2}\int J(\xi)K(\xi,\xi''')J(\xi''')d^2\xi\, d^2\xi'''} \tag{26}$$

where K is the σ-field propagator described in Appendix C. Equation 26 then becomes

$$e^{8K(\xi,\xi) + 8K(\xi',\xi') - 16K(\xi,\xi')}.$$

Using the asymptotic expressions from Appendix C :

$$4K(\xi,\xi') = -\frac{\pi}{2}[\frac{1}{4\pi}\log\mu^2((r - r')^2 + (t - t')^2)$$

$$+ \frac{1}{4\pi}\log\mu^2((r + r')^2 + (t - t')^2)] - \frac{1}{4\pi}\log\frac{(t - t')^2 + (r - r')^2}{8rr'}.$$

Letting $(r - r') = \varepsilon$ and t=t' , we get after taking the limit $\varepsilon \to 0$,

$$4K(\xi,\xi) = \frac{-1}{4}\log\mu^2 - \frac{1}{8}\log 32 - \frac{1}{4}\log r^2$$

and

$$4K(\xi',\xi') = \frac{-1}{4}\log\mu^2 - \frac{1}{8}\log 32 - \frac{1}{4}\log r'^2$$

and

$$4K(\xi,\xi') = \frac{-1}{4}\log\mu^2 - \frac{1}{8}\log[(r - r')^2 + (t - t')^2] - \frac{1}{8}\log[(r + r')^2 + (t - t')^2].$$

Putting everything together one finds

$$e^{8K(\xi,\xi) + 8K(\xi',\xi') - 16K(\xi,\xi')} = \frac{1}{r}\frac{1}{r}(\sqrt{(r - r')^2 + (t - t')^2}\sqrt{(r + r')^2 + (t - t')^2}). \tag{27}$$

Substituting (25) and (27) in (24) we get

$$<M\,|\,Q(r,t)Q^{\dagger}(r',t')\,|M> =$$

$$\frac{1}{2\pi^2}\frac{1}{r^3r'^3}[\frac{1}{(t - t')^2 + (r - r')^2} + \frac{1}{(t - t')^2 + (r + r')^2}]$$

$$\times\sqrt{(r - r')^2 + (t - t')^2}\sqrt{(r + r')^2 + (t - t')^2}.$$

In the limit (t-t')-> ∞, this expression simplifies to $\frac{1}{2\pi^2}\frac{2}{r^3r'^3}$. Using the cluster property one finds

$$<M\,|\,Q(r,t)\,|\,M> \approx \frac{1}{r^3} \qquad (28)$$

The important points to note are that the expression is independent of the coupling constant, and that it has no dimensionful parameter in it. Dependence on the coupling constant enters only in higher order corrections. Thus we have shown that the operator in equation (18) gets a vev in the presence of the monopole. The calculation is almost identical to that done originally by Rubakov for the isospinor case. Let us pause for a minute and summarize what we have learnt from this calculation. The isotriplet fermions are expanded in spherical harmonics of total angular momentum J, with J= 1/2, 3/2, 5/2... There is a centrifugal barrier proportional to $\sqrt{J(J+1)+1/4-\tau^2}$ which prevents the harmonics with J > 1/2 from participating in the condensate (this is still to be shown). In the J=1/2 sector the centrifugal barrier vanishes for $\tau = \pm 1$, but not for $\tau = 0$. Consequently $\tau = 0$ fermions also do not participate except in an infinitesimal region around the core.* The $\tau = \pm 1$ fermions behave like an isodoublet, but with twice the charge of a genuine isodoublet. We refer to these as "isodoublets". Also since J =1/2, there are two values of M which, for all calculational purposes, are like two flavours. In the presence of configurations with non zero winding number, these "isodoublets" have zero modes. To count the number of zero modes invoke the index theorem, which says that if the winding number is one, an isodoublet has one zero mode (since these are Weyl fermions). There are two "isodoublets", but each one is doubly charged. This means that the effective winding number seen by the "isodoublet" is twice that seen by the usual isodoublet. Thus, in a configuration with winding number one, there are four zero modes and a four-fermi operator should condense. However, explicit calculation showed that it was a two fermion operator which got a vev. This could only mean that the gauge field configuration that was responsible for the zero modes had a

* Another reason for not participating in the condensate, special to $\tau = 0$ fermions, is that they are uncharged and do not see the monopole in the limit $r_M \to 0$.

winding number of 1/2. In fact, it has already been argued that there is no reason for the winding number to be an integer, and so there is no cause for surprise. The winding number can be checked as follows. It was pointed out that the terms linear σ in equation (24) can be treated as source terms where $J(\xi) = 4\delta(\xi - \xi) - 4\delta(\xi - \xi')$. The classical solution to the equations of motion is then given by the configuration

$$\sigma(\xi^2) = \int d\xi^1 K(\xi^2,\xi^1) J(\xi^1) \tag{29}$$

$$= 4K(\xi^2,\xi) - 4K(\xi^2,\xi').$$

Usually, this is only a saddle point of the integrand, but in our case, because the action is quadratic this gives the exact result of the functional integral. Knowing the exact form of the configuration, one can evaluate the value of $\omega(\infty,+\infty) - \omega(\infty,-\infty)$, which gives us the winding number. It is found to be a combination of winding numbers $\pm\frac{1}{2}$ (the + 1/2 part, + K , being responsible for $<M| Q(r,t)|M>$ and the other, - K , for $<M|Q^\dagger(r',t')|M>$). Thus, whenever the fermions have integral charge (T = 1, 2..), we can expect to have winding numbers that are half integral. However, the addition of two flavours of isospinors (so that now there is one isovector along with two isospinors) modifies the picture once again. One can re-do the calculation (this requires only minor modifications and is done in Appendix D) to find that the non-vanishing condensate has four fermions from the T=1 multiplet and one each from the T=1/2 multiplets. This is consistent with a winding number of one. In this case, one cannot have a winding number of half, because that would require half a fermion each, from the isospinor multiplets. This feature generalizes to arbitrary representations, i.e. if there is even one fermion representation that has half integral isospin, the winding number has to be integral.

To complete this analysis from the point of view of instanton physics, the following question should be answered: the index theorem states that $N_L - N_R = c\,\nu[N_L$ = number of negative chirality zero modes; N_R = number of positive frequency zero modes ; ν = winding number ;c a numerical constant]. What can be said about N_L and N_R individually ? In the case of the instanton solutions, the self duality (or anti self duality) ensured that N_L (or N_R) was non zero. In the present case, it was implicitly assumed in the discussions of the previous paragraphs that a similar property held. We would like now to substantiate this assumption by showing that, for the gauge field configuration in equation (29), N_L or N_R is zero. We use the notation of ref. [18]. Then

$$\bar{\alpha}^{\mu}(\partial_{\mu}+A_{\mu})\psi_R = 0$$

where $\bar{\alpha}^{\mu} = (i\vec{\sigma},1)$

$$\alpha^{\mu}(\partial_{\mu}+A_{\mu})\psi_L = 0$$

where $\alpha^{\mu} = (-i\vec{\sigma},1)$. Applying $\alpha^{\mu}(\partial_{\mu}+A_{\mu})$ and $\bar{\alpha}^{\mu}(\partial_{\mu}+A_{\mu})$ respectively gives

$$[(\partial_{\mu}+A_{\mu})^2 + 2i\bar{\sigma}_{\mu\nu}F^{\mu\nu}]\psi_R = 0$$

$$[(\partial_{\mu}+A_{\mu})^2 + 2i\sigma_{\mu\nu}F^{\mu\nu}]\psi_L = 0$$

where $\sigma^{i4} = -\bar{\sigma}^{i4} = \frac{1}{2}\sigma^i$, $\sigma^{ij} = \bar{\sigma}^{ij} = \frac{1}{2}\varepsilon^{ijk}\sigma^k$. In our case, $2i\sigma_{\mu\nu}F^{\mu\nu} = -\frac{1}{r^2}(\hat{r}.\vec{T})(\hat{r}.\vec{\sigma}) - (\hat{r}.\vec{T})(\hat{r}.\vec{\sigma})(\partial_r\partial_t\omega)$. Also $(\hat{r}.\vec{T})(\hat{r}.\vec{\sigma})$ is negative. This was seen explicitly in the isovector case, and also in the isospinor case [1,2]. This is true in general also because $\hat{r}.\vec{T} + \hat{r}.\vec{\sigma} = \hat{r}.\vec{J}$. For a zero mode, $\sqrt{J(J+1) + 1/4 - \tau^2} = 0$, which requires $J = \tau - \frac{1}{2}$. This implies $|\hat{r}.\vec{T}| > |\hat{r}.\vec{J}|$, and therefore $(\hat{r}.\vec{\sigma})$ has to have a sign opposite that of $\hat{r}.\vec{T}$, which proves the statement. Also, in the case where ω increases from 0 to $\frac{\pi}{2}$, it can be shown from the formula in equation (29) that $\partial_t\partial_r\omega$ is positive. To see this, note that $\partial_t\partial_r\omega \approx \frac{g^2}{2}Q_k[1 + \frac{(r-r)^2 + (t-t)^2}{2rr'}]$ (see Appendix C), and k is a function of the coupling constant but > 0 always. Q_k then has the property that $Q_k(x) > 0$ if x > 1. So $\partial_t\partial_r\omega$ is positive everywhere. Thus $(\partial_{\mu}+A_{\mu})^2 + 2i\bar{\sigma}^{\mu\nu}F_{\mu\nu}$ is a positive definite operator for winding number 1/2. There are therefore no normalizable zero modes for ψ_R if the winding number is 1/2. Similar arguments hold for winding number -1/2, and also for higher winding numbers.

We now demonstrate as promised earlier that the condensate consists entirely of J = 1/2 fermions. The J > 1/2 fermions do not participate in the condensate because there are no zero mode solutions with J >1/2. The J > 1/2 Lagrangian in the limit $r_M \to 0$ is (8)

$$L = \int U^J_{M\sigma\tau}(\partial_t + 2i\tau\partial_t\omega + 2i\sigma(\partial_r + \frac{1}{r}))U^J_{M\sigma\tau} + \frac{\sqrt{J(J+1) + \frac{1}{4} - \tau^2}}{r} U^J_{M\sigma\tau} U^J_{M-\sigma\tau} d^2r$$

Define $U^J_{M\sigma\tau} = rU^J_{M\sigma\tau}$. The Lagrangian becomes

$$L = \int U^J_{M\sigma\tau}(\partial_t + 2i\tau\partial_t\omega + 2i\sigma\partial_r)U^J_{M\sigma\tau} + \frac{\sqrt{J(J+1) + \frac{1}{4} - \tau^2}}{r} U^J_{M\sigma\tau} U^J_{M-\sigma\tau} dr$$

Define $U^J_{M\tau} = \begin{pmatrix} U^J_{M\sigma\tau} \\ U^J_{M-\sigma\tau} \end{pmatrix}$.

Then

$$L = \int [U^J_{M\tau}(\partial_t + 2i\tau\partial_t\omega + 2i\sigma_3\partial_r + \frac{c(J,\tau)}{r}\sigma_1)U^J_{M\tau}]dr \qquad (30)$$

where $c(J,\tau) = \sqrt{J(J+1) + \frac{1}{4} - \tau^2}$. A further redefinition $V^J_{M\tau} = \begin{pmatrix} U^J_{M\sigma\tau} + iU^J_{M-\sigma\tau} \\ U^J_{M\sigma\tau} - iU^J_{M-\sigma\tau} \end{pmatrix}$

allows us to write the equation for the Green's function in the following form:

$$\begin{pmatrix} \partial_t & i(\partial_r + \frac{c}{r}) \\ i(\partial_r - \frac{c}{r}) & \partial_t \end{pmatrix} \begin{pmatrix} \partial_t G_1 & -i(\partial_r + \frac{c}{r})G_2 \\ -i(\partial_r - \frac{c}{r})G_1 & \partial_t G_2 \end{pmatrix} = \begin{pmatrix} \delta(r-r')\delta(t-t') & 0 \\ 0 & \delta(r-r')\delta(t-t') \end{pmatrix} \qquad (31)$$

where G_1 and G_2 satisfy:

$$[\partial_t^2 + \partial_r^2 - \frac{c(c-1)}{r^2}]G_1(r,t;r',t') = \delta((r-r))\delta((t-t)) \qquad (32)$$

$$[\partial_t^2 + \partial_r^2 - \frac{c(c+1)}{r^2}]G_2(r,t;r',t') = \delta((r-r))\delta((t-t)).$$

If J $> \frac{3}{2}$, then $c \geq \sqrt{3}$. Thus both $c(c+1)$ and $c(c-1)$ are necessarily > 0. In that case, it is easy to see that the Green's function has to vanish at the origin. In the case of G_1 for small r, $v \sim r^{\sqrt{3}}$ or $r^{1-\sqrt{3}}$. The latter solution would require $U \approx r^{-\sqrt{3}}$ which is not square integrable at the origin. Therefore we are left with $r^{\sqrt{3}}$ which vanishes at the origin. The same is true for G_2. This can be made use of as follows*. The vanishing of the Green's function at the origin allows us to perform a singular gauge transformation on the gauge fields - gauge transformations of the form $A_\mu \to g^{-1}A_\mu g + g^{-1}\partial_\mu g$ where $g = \exp(i\Lambda(r,t)\,\hat{r}.\vec{T})$, $\Lambda(0,t) \neq 0$. The fermion determinant in terms of diagrams is shown in fig. 14. In coordinate space it is of the form

$$\int d^2y_1 d^2y_2 \ldots G^J(y_1,y_2)a(y_2)G^J(y_2,y_3)\ldots G^J(y_N,y_1) \tag{33}$$

$a = \gamma^\mu a^\mu$, a_μ is the excitation of the background monopole configuration. In the calculations so far $a_1 = 0, a_0 = 2\tau\partial_t\omega$, $\gamma^0 = iI$, $\gamma^1 = i\sigma^3$. Consider the part $\int d^2y_2 G^J(y_1,y_2)\gamma^\mu G^J(y_2,y_3)a_\mu$. A gauge transformation adds to a_μ the piece $\partial_\mu\Lambda$

$\int d^2y_2 G^J(y_1,y_2)\gamma^\mu G^J(y_2,y_3)\partial_\mu\Lambda(y_2)$ can be shown to be zero. Integrating by parts we get

$$\int d^2y_2 \partial_\mu[G^J(y_1,y_2)\gamma^\mu G^J(y_2,y_3)\Lambda(y_2)]$$

$$-\int d^2y_2[\partial_\mu G^J(y_1,y_2)]\gamma^\mu G^J(y_2,y_3)\Lambda(y_2) - \int d^2y_2 G^J(y_1,y_2)\gamma^\mu[\partial_\mu G^J(y_2,y_3)]\Lambda(y_2).$$

The last two terms cancel against each other on using the equation of motion. The first term vanishes being a surface term, since $G^J(y,y')$ vanishes both at r=0 and r= ∞. If G had not vanished at r=0 we would have had to require $\Lambda(0,t)$ =0. Thus the terms shown in fig. 16, and hence the fermion determinant, are invariant under gauge transformations, for which $\Lambda(0,t) \neq 0$. But by means of such singular gauge transformations one can change the winding number of a configuration. Consider a gauge transformation with $\Lambda(r,-\infty) = 0$, $\Lambda(r,+\infty) = \pi$.

* This argument was made by Rubakov [1]. We have given it a physical interpretation

This has a winding number of one. This can be seen by noting that on the surface at infinity of Euclidean space, S_3 , this has the same form as an instanton configuration (in the Lorentz gauge) of winding number one. Since the winding number of a configuration depends only on the values on this surface at infinity, the result follows. The fact that the winding number can be changed at will, and in particular be made zero, without affecting the fermion determinant implies that there are no zero modes. These J > 1/2 fermions do not 'see' any winding number, and therefore the condensate consists entirely of J = 1/2 fermions. This argument applies to other fermion representations also, and will be made use of in generalizing the results of this section.

The calculation of the condensate formation in the monopole isovector fermion system is now complete up to finite mass effects. We also have a simple instanton picture to understand the condensate formation. These results along with results from the isodoublet calculation of Rubakov and Callan enable us to discuss the general features of the fermion monopole system. This is done in the next section.

5. The Monopole Fermion System

In the presence of massless fermions the monopole is surrounded by a condensate of J-invariant multilinear fermion operators. We know the composition of this condensate in the case of the fermions which are isodoublets or isotriplets. (The more general representation is treated in the next section). For an isotriplet (ψ^+,ψ^0,ψ^-), the condensate is $\psi^+\psi^-$ (if no isodoublet is present). If isodoublets (χ^{+1}, χ^{-1}) and (χ^{+2}, χ^{-2}) are also present, the condensate is $\psi^+\psi^-\psi^+\psi^-\chi^{+1}\chi^{-2}$ as discussed in the last section. These condensates fall off as $r^{-3n/2}$ where n is the number of fermion operators in the condensate. What symmetries do these condensates violate ? The answer is as follows. Symmetries that are preserved by the Lagrangian and that are anomaly free are not violated by the condensate. Of course this is not a coincidence. It is clear from the instanton picture that the distribution of fermions in the condensate is governed by the index theorem which counts the number of zero modes for each representation. Also, as previously mentioned, the anomaly equation is just a local version of the index theorem. Thus the condensate formation has to be consistent with the anomaly equation. In particular, if a symmetry is anomaly free, the anomaly equation requires that the condensate be neutral under that symmetry.* On the other hand, if a symmetry is anomalous, the anomaly equation also tells you how many units of that quantum number is violated by the condensate. Let us illustrate this with a simple example. Consider a theory with two left handed isodoublets $\psi^\pm$, $\chi^\pm$. The minimal condensate formed is $\psi^+\chi^- - \psi^-\chi^+$.** The superscripts $\pm$ refer to charge $\hat{r}.\vec{T}$. This charge is conserved and anomaly free, as is reflected by the manifest charge neutrality of the condensate. It is also J invariant, which is a

*Global symmetries corresponding to exactly conserved currents can get spontaneously broken by the vacuum e.g. chiral symmetries in QCD. But these are strong coupling effects. Here we are talking about effects that exist for arbitrarily weak coupling.

**Note that we are always referring to a condensate with the least possible number of fermions in it. Thus condensates of the form $(\psi^+\chi^- - \psi^-\chi^+)^n$ for arbitrary n are formed but they fall off much faster than the minimal one.

statement of the conservation of angular momentum. For the isodoublet case this is trivially true, since each fermion field in the condensate has J=0 ***, and is therefore invariant under J. It does not have to be invariant under the full Lorentz group, because the monopole background is not invariant under boosts. This condensate carries -2 units of chirality. This is easily seen to be consistent with the anomaly equation. The $U_A(1)$ (A stands for axial) current satisfies : $\partial_\mu j^\mu{}_5 = N\frac{e^2}{64\pi^2}\varepsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma}$. (N is the number of flavours $\frac{1}{64\pi^2}\int d^4x\,\varepsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} = \nu$ winding number). For $\nu = 1$, N =2 we get ΔQ_5=2. Similarly one can define conserved fermion numbers or any other quantum number and verify that there is always complete agreement between the violation (or conservation) of quantum numbers predicted by the anomaly equation prediction and the violation (or conservation) by the condensate.

Motivated by the approximate baryon number conservation in SU(5), we would like now, to understand what happens to quantum numbers that are only approximately conserved. In the SU(5) example, at low energies where the gauge group is SU(3) X SU(2) X U(1), baryon number is conserved approximately. Any perturbative violation is proportional to $\frac{1}{M_X}$, where M_X is the scale of SU(5) breaking. In the limit $M_X \to \infty$ one expects this symmetry to be conserved exactly. Yet, this need not be true in the presence of a grand unified monopole, since the very existence of such a monopole is a reminder of the underlying grand unification. (Of course as $M_X \to \infty$, the monopole itself becomes infinitely heavy, but if they had been produced in sufficient numbers in the early universe this does not cost us anything). Rubakov and Callan have argued, that when the results for SU(2) monopoles are extended to SU(5) monopoles, one should expect to see strong baryon number violation, even in the limit $M_X \to \infty$. In order to examine approximately conserved symmetries, let us go back to our SU(2) monopole with two isodoublets, $\psi_L^\pm$, $\chi_L^\pm$. Let us

***When we say that $\psi^+\chi^- - \psi^-\chi^+$ is the condensate, it is understood that only those harmonics of ψ,χ with the appropriate J (J=0 for isospin 1/2), actually participate in this.

define some U(1) quantum number (G) as follows :

$$\psi^+(1) \qquad \chi^+(2)$$

$$\psi^-(-1) \qquad \chi^-(-2).$$

This global U(1) symmetry is clearly anomaly free, and also conserved in the low energy theory by the photon interactions. But there is no consistent assignment of U(1) charges possible for the $W^\pm$ bosons. Thus the full Lagrangian has interactions terms involving the $W^\pm$ fields that explicitly violate this symmetry. In the absence of a monopole, any violation of this baryon number G is suppressed by powers of $\frac{1}{M_X^2}$, because of the W propagators. This symmetry G is exactly like baryon number in SU(5) (except for the fact that G is anomaly free even in the full theory, unlike baryon number). However, in the presence of the monopole, the condensate $\psi^+\chi^- - \psi^-\chi^+$ is formed which violates G by ± 1. Furthermore, this violation is strong, because the condensate falls off only as $\frac{1}{r^3}$, and exists even in the limit $M_X \to \infty$. Although this symmetry is anomaly free, the current associated with this symmetry is not conserved. It satisfies $\partial_\mu j^{\mu G} \approx e D_\mu \vec{\phi} \bar{\psi} \vec{\tau} \gamma^\mu \psi$. (We assign quantum numbers to $W^\pm$ so that the term $\bar{\chi} D \chi$ is invariant under this symmetry). The R.H.S. reflects the explicit violation of this symmetry. Integrating over space-time we get $\Delta G \approx \int d^4x D_\mu \vec{\phi} \bar{\psi} \vec{\tau} \gamma^\mu \psi$. Since this quantity is a pure dimensionless number and there is only one scale, M_X in the problem, it cannot depend on this scale (the fermions being massless). Substituting for ψ the J=0 zero mode of ref.[1], one finds for finite M_X, a non zero contribution from the core where $D_\mu \varphi \neq 0$. Since it does not depend on M_X this contribution remains non vanishing in the limit $M_X \to \infty$. Therefore it is not surprising that the vacuum around the monopole violates this symmetry even in this limit.

Let us turn to a discussion of the scattering process. Only the J=0 sector undergoes a chirality-flip scattering in the isospinor case. Since $\hat{r}.\vec{T} + \hat{r}.\vec{S} = 0$, a positively charged particle ($\hat{r}.\vec{T} > 0$) has to have its spin pointing inward. So there are positively charged, incoming right-handed and outgoing left-handed particles, and the reverse for negatively charged particles. If the condensate is $\psi^+\chi^-$, then $\chi^-{}_L$ can annihiliate a left-handed incoming particle and replace it with an outgoing right-handed negatively charged particle $(\psi^+{}_L)^*$. Charge is conserved and chirality violated in the process. It also violates the baryon number G introduced earlier. We will discuss the SU(5) case briefly in sec. 2.7.

We have been describing till now the chirality violating scattering as taking place around the monopole. However, just as in the analysis of the Dirac monopole (sec 1.3), one can study the process purely quantum mechanically, without introducing the quantum field-theoretic concept of a condensate. This has been done by many authors. The picture that emerges is that all the action (chirality flipping, baryon number violation) takes place as the particle goes through the centre of the monopole, just as in the Abelian case. In this picture, the condensate of $W^\pm$ bosons in the centre of the monopole, is responsible for the violation of baryon number G rather than the condensate of fermions surrounding the monopole. (This process cannot, of course, be the perturbative process $\psi^+ \rightarrow \psi^-$ (fig. 15) which involves a charge exchange, but has to be something more complicated). One might think at first that the two approaches would predict dramatically different values for the cross section, because in one case the particle has to go through the centre of the monopole whereas in the other case it scatters off the condensate, which extends a great distance outside the core, and therefore provides a bigger target. This is not the case. For example, in the isospin 1/2 case, in the quantum mechanical picture, only the J=0 wave sees the core and undergoes chirality flip. In the field theoretic calculation also, the condensate consists of J=0 fermions and therefore only J=0 fermions scatter off the condensate and flip chirality. In both cases chirality changes with unit probability. Thus the geometric size of the target is irrelevant. Any difference between the two pictures is necessarily of a more subtle nature. One possible difference could lie in the following effect.

Imagine a wave packet (in the J=0 state) concentrated around r=R and moving inwards. In the quantum mechanical picture we can expect to see chirality flip scattering only after a time interval of $\frac{2R}{v}$, where v is the velocity of the wave packet because it has to go through the core. In the other picture, we expect to see scattering long before that because the condensate exists for macroscopic distances outside the monopole. Thus, in principle, one should expect to see some difference in the two pictures. This requires further investigation and will not be pursued here.

6. Fermions in Arbitrary Representations

In this section we shall extend the results of Rubakov and Callan for isodoublet fermions, and that of section 2.4 for isovector fermions, to fermions in arbitrary representations of the SU(2) group. The generalization is straightforward and introduces no new concepts.

Consider a fermion with isospin T. We can expand the fermion field operator into eigenstates of $\vec{J}.\vec{J}, J_3, \hat{r}.\vec{S}, \hat{r}.\vec{T}$ as in eq. (6). The Lagrangian is of the form given in eq. (8) for all J's once the limit $r_M \to 0$ is taken. The centrifugal barrier is proportional to $\sqrt{J(J+1)+\frac{1}{4}-\tau^2}$. This vanishes for $J = \tau - \frac{1}{2}$. Thus if |J| > |T|, the Green's function necessarily vanishes at the origin, and by the same arguments as before, it can be shown that there are no normalizable zero modes and hence no condensate formation. We therefore concentrate on $|J| \le |T-\frac{1}{2}|$. Let us count the number of zero modes. For any given value of τ, the value of J for which the centrifugal barrier vanishes is $\tau - \frac{1}{2}$. If $J = \frac{1}{2}$, then $(\hat{r}.\vec{J})_{\max} = \tau - \frac{1}{2}$. Also, $\hat{r}.\vec{J} = \hat{r}.\vec{S} + \hat{r}.\vec{T} = \sigma + \tau$. Of the two values for σ $(\pm\frac{1}{2})$, we are forced to have $\sigma = -\frac{1}{2}$. The only remaining label for the wave function is M, which takes on $2J+1 = 2\tau$ values. So we have 2τ such functions. Furthermore, for a given winding number, say ν, as ω varies from 0 to $\nu\pi$, the effective variation of ω as seen by this particle is τ times as much, since it has charge τ. Thus the number of zero modes is $2\tau^2\nu$. We have to sum this over all possible values of τ from -T to +T. So the total number of zero modes is $2\sum_{\tau=-T}^{\tau=+T}\tau^2\nu = 2/3T(T+1)(2T+1)\nu$. This is a standard result, but we have obtained it merely by studying the form of the Lagrangian. For an application of these ideas to a few cases see Table 1.

The picture which emerges then is the following: a fermion representation of isospin T behaves like a union of decoupled doublets, the different doublets having $\tau = \pm T, \pm(T-1), \ldots$ with $J = \tau - \frac{1}{2}$. in each case. Also the number of doublets with a given τ is 2τ. The Lagrangian for each doublet is exactly the same, except for the value of the charge, which is τ. A word about the

boundary conditions. For the T=1/2 case there is a boundary condition that equates the values of the $\tau = +\frac{1}{2}$ and $\tau = -\frac{1}{2}$ components at the origin. The physical reason for this is quite clear. The Hamiltonian contains terms that connect these two states and these terms go as 1/r. This forces the components to be exactly equal at the origin. For the isovector case also, all the components had to have values in some definite proportion at the origin. Furthermore, the $\tau = +1$ component was equal to the $\tau = -1$ component. This should be the case for higher values of T also and one expects to have a definite relation between the $+\tau$ and $-\tau$ components of the doublets. In fact, by symmetry they have to be equal as in the isovector case. Again as in the isovector case, it should be possible to satisfy the boundary conditions in a trivial way. Thus we expect no difference (except for the details already mentioned) as we go to higher representations. Furthermore, as in the isovector case, if all representations have integral isospin, one can have half integral winding numbers and the minimal condensate contains half the number of fermion operators one would put in naively. Putting all these facts together, one finally arrives at the following prescription for the form of the condensate. If ψ_i are left handed fermion fields in the theory, the operator that gets a vev is of the form $\psi_1^{m_1}\psi_2^{m_2}\cdots$ with m_i being integers. The operator satisfies the following conditions:

(a) It is neutral under all anomaly free symmetries of the theory.

(b) It is a singlet under $\vec{L} + \vec{S} + \vec{T}$

(c) The m_i are such, that a fermion with charge τ and J $=\tau-\frac{1}{2}$ occurs $4\tau^2\nu$ times, ν being the winding number. ν is an integer if there is at least one representation with half integral isospin and half integer otherwise.

7. Conclusions

In this chapter we have described in some detail the fermion monopole system. The discussion was essentially restricted to the 't Hooft Polyakov monopole in an SU(2) gauge theory. We conclude now with a brief discussion of the possible relevance the phenomena described might have to the real world. Since, thus far, no one has succeeded in doing any realistic calculations, the discussion shall necessarily be of a conjectural nature.

We had mentioned earlier the possibility of baryon number violation being catalysed by grand unified monopoles. Let us look at the fundamental monopole in the SU(5) grand unified theory [6]. The SU(2) embedding in the SU(5) can be described by identifying the location of the 2x2 Pauli matrices within the 5x5 SU(5) generators.

$$\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & (SU(2)) & \\ & & & 0 \end{bmatrix}$$

It acts on the third and fourth indices of the 5. The fermions of one family that interact with this monopole can be grouped into four doublets.

$$\begin{pmatrix} -\bar{u}^2 \\ u^1 \end{pmatrix}_L \begin{pmatrix} \bar{u}^1 \\ u^2 \end{pmatrix}_L \begin{pmatrix} d^3 \\ e^+ \end{pmatrix}_L \begin{pmatrix} e^- \\ -\bar{d}^3 \end{pmatrix}_L$$

In accordance with the rules discussed the minimal condensate consists of one fermion from each doublet and also conserves $\hat{r}.\vec{T}$. One of the possibilities is the combination $u_L^1 u_L^2 d_L^3 e_L^-$. This combination has the quantum numbers of $p\ e^-$ and violates baryon number. Note that it preserves B-L which is an exact symmetry in SU(5). The scattering process has u^1,u^2 as incoming particles in J=0 waves so that $\hat{r}.\vec{T} + \hat{r}.\vec{S} = 0$, and d^3,e^+ as outgoing particles. Thus the process is (fig. 16):

$$u_L^1 + u_L^2 \longrightarrow d^3 + e^+$$

or

$$p \longrightarrow \pi^0 + e^+.$$

This suggests that monopoles catalyse baryon number violation at strong interaction rates. The effect this has on neutron stars has been examined and used to place strong bounds on the flux of monopoles in the galaxy [7]. These results should be taken as tentative because there has not been any quantitative estimate of the process $p \longrightarrow \pi^0 + e^+$. The effects of fermion masses have to be taken into account. The confining effects of gluon interactions have also been neglected. These are topics that require further investigation.

Another interesting application of the ideas discussed here is the U(1) problem of QCD. QCD has an $SU(N)_L \times SU(N)_R \times U(1)_V \times U(1)_A$ symmetry. When chiral symmetry breaks down spontaneously and the quarks get their masses, the remaining symmetry is $SU(N)_V \times U(1)_V$. The pions and their extensions to SU(N) are the pseudo- Goldstone bosons associated with $\dfrac{SU(N)_L \times SU(N)_R}{SU(N)_V}$. (They are pseudo-Goldstone bosons because the bare mass terms for the quarks explicitly break $SU(N)_L \times SU(N)_R$.) There should be a pseudo goldstone boson for the $U(1)_A$ also. It was first thought to be the η' but it was shown that the relation $m_\eta' \leq \sqrt{3} m_\pi$ had to hold in that case. This is not true experimentally ($m_\pi \approx 139 Mev$, $m_\eta' \approx 958 Mev$). This is the U(1) problem. The solution to this invokes instanton effects. It is known that $U(1)_A$ had an anomaly. This has a non trivial effect when instanton configurations are turned on. It can be shown first of all, that because of the anomaly, the goldstone boson is not a physical particle, and does not show up as a physical pole in gauge invariant Green's functions and therefore the η is not the pseudo-Goldstone boson for the $U(1)_A$. Secondly, the instanton effects cause there to be a condensate of fermion multilinear operators in the QCD ground state which could give a mass to the η particle. This mass is, of course, incalculable, due to the infrared divergences of QCD.

The phenomenon discussed in this paper allows us to link the mass of the η more directly to the confinement mechanism in the following way. It is known that the confining vacuum of QCD can be described as a condensate of monopoles. Then there should also exist a condensate of fermion operators around each microscopic monopole due to the phenomenon described in this chapter. This condensate is exactly the same as the one for which the instantons were responsible before and therefore gives a mass to the η! It should be emphasized that although there was no need to invoke instantons in this picture, this mechanism and the instanton mechanism are essentially the same. The word 'instanton' refers to a particular approximation scheme for some QCD calculations. The monopole mechanism is a different approximation scheme. Since the effects of infrared divergences, which plague instanton calculations, have presumably all been lumped into the parameters describing the monopole condensate, this should provide a better starting point for calculations. This concludes our discussion of the monopole fermion system.

Table I

T	τ	J	$J^2+1/4-\tau^2$	No. of zero modes	$\tau.\sigma$
1/2	1/2	0	0	1	-ve
1	1	1/2	0	4	-ve
1	0	1/2	1	0	?
3/2	1/2	0	0	1	-ve
3/2	3/2	1	0	9	-ve
3/2	1/2	1	2	0	?
2	1	1/2	0	4	-ve
2	0	1/2	1	0	?
2	2	3/2	0	16	-ve
2	1	3/2	3	0	?
2	0	3/2	4	0	?

Note:'?' stands for no definite sign.

APPENDIX A:

We would like to outline the calculation leading to the Lagrangian for T=1 fermions. We are only concerned with excitations of the type

$$A_0 = 2i\,\partial_0\omega\hat{r}.\vec{T}, \qquad A_i = A_i^{cl}, \qquad \varphi=\varphi^{cl}$$

The Lagrangian of the theory is

$$\bar{\psi}\gamma_E^\mu(\partial_\mu + A_\mu)\psi$$

where $\gamma_E^0=1$, $\gamma_E^i = i\sigma^i$. Substituting the form of A_i^{cl} given in equation (1) it takes the form :

$$L = \psi^\dagger[\partial_0+A_0 + \frac{1}{r}D_\Omega + i(\vec{\sigma}.\hat{r})(\partial_r+\frac{1}{r})]\psi$$

where

$$D_\Omega = D_{\Omega_0} + \Delta$$

$$D_{\Omega_0} = ir\sigma^i(\delta_{ij}-\hat{r}_i\hat{r}_j)\partial_j - \sigma^r\varepsilon_{rat}T^a\hat{r}^t - i\hat{r}.\vec{\sigma}$$

$$\Delta = \sigma^r\varepsilon_{rat}T^a\hat{r}^t(1-F)$$

$$D_\Omega = D_{\Omega_0} + \Delta$$

In the limit F->1 only D_{Ω_0} survives. D_{Ω_0} has the nice commutation and anti- commutation rules:

$$[D_{\Omega_0},\hat{r}.\vec{T}] = 0 \qquad \text{(A1)}$$

$$[D_{\Omega_0},\hat{r}.\vec{\sigma}]_+ = 0 \qquad \text{(A2)}$$

Also

$$D_{\Omega_0}{}^2 = \vec{J}.\vec{J} + 1/4 - (\hat{\tau}.\vec{T})^2 \tag{A3}$$

We split the calculation into two parts. First we evaluate the Lagrangian using D_{Ω_0}. This is very straightforward if we use (A1),(A2) and (A3) with the expansion

$$\psi = \sum_{J,M,\sigma,\tau} U^J_{M\sigma\tau}(r,t)\Psi^J_{M\sigma\tau}(\theta,\varphi) \tag{A4}$$

where the $\Psi^J_{M\sigma\tau}$ satisfy

$$\int \Psi^{J*}_{M\sigma\tau}\Psi^{J'}_{M'\sigma'\tau'}d\Omega = \delta^{JJ'}\delta^{MM'}\delta^{\sigma\sigma'}\delta^{\tau\tau'}$$

and $\hat{\tau}.\vec{J} = \sigma + \tau$ This gives for the J=1/2, T=1 case

$$\sum_{M,\sigma,\tau}\int[U^*_{M\sigma\tau}(\partial_0 + A_0 + 2i\sigma(\partial_r + 1/r))U_{M\sigma\tau} + \frac{\sqrt{1-\tau^2}}{r}U^*_{M\sigma\tau}U_{M-\sigma\tau}]r^2drdt \tag{A5}$$

To evaluate the contribution due to Δ, it is more convenient to use explicit expressions for $\Psi^{\frac{1}{2}}_{M\sigma\tau}$

$$\psi_{ija} = S_{ijaM}\theta^M + V_{ijaM}\chi^M + T_{ijaM}\xi^M + A_{ijaM}\lambda^M \tag{A6}$$

where

$$S_{ijaM} = \varepsilon_{ia}\varepsilon_{jM} + \varepsilon_{ja}\varepsilon_{iM}$$

$$V_{ijaM} = (\hat{\tau}.\vec{\tau}\varepsilon)_{ij}\varepsilon_{aM}$$

$$T_{ijaM} = (\hat{\tau}.\vec{\tau}\varepsilon)_{ij}(\hat{\tau}.\vec{\tau}\varepsilon)_{aM}$$

$$A_{ijaM} = i\varepsilon^{abc}\hat{\tau}^a(\tau^b\varepsilon)_{ij}(\tau^c\varepsilon)_{aM}$$

i,j =1-2 are isospin indices and a=1-2 is a spin index, M =1-2 is the index corresponding to $J_3=\pm 1/2$, τ^a with a=1-3 are the Pauli matrices and S,V,T,A are meant to suggest scalar, vector, tensor and axial vector respectively.

Using this expansion one finds that the contribution to L proportional to (1-F) is

$$\frac{4i}{r}[(\chi^\dagger\theta - \theta^\dagger\chi) + (\lambda^\dagger\theta - \theta^\dagger\lambda) + (\xi^\dagger\lambda - \lambda^\dagger\xi)] \tag{A7}$$

We now have to express this in terms of the $U^M_{\sigma\tau}$. By acting with the generators $\hat{\tau}.\vec{S}$ and $\hat{\tau}.\vec{T}$ one can verify that the tensors S,V,T and A combine into eigenfunctions of the above generators in the following way :

$$A+S+T : \sigma = \frac{-1}{2}, \ \tau = +1$$

$$T-V : \sigma = \frac{-1}{2}, \ \tau = 0$$

$$A-S-T : \sigma = +\frac{1}{2}, \tau = -1$$

$$T+V : \sigma = +\frac{1}{2}, \tau = 0$$

For each (σ,τ) pair there are two values of M. We have to normalize these before identifying them with the $\Psi_{M\sigma\tau}$. So we require $\int (A+S+T)^{*\,ijaM}(A+S+T)_{ijaN} d\Omega = \delta_{MN}$ and the same for the rest which lead to

$$\Psi_{\frac{-1}{2}1} = \frac{(A+S+T)}{2\sqrt{2}}$$

$$\Psi_{\frac{-1}{2}0} = \frac{(T-V)}{2}$$

$$\Psi_{\frac{1}{2}-1} = \frac{(A-S-T)}{2\sqrt{2}}$$

$$\Psi_{\frac{1}{2}0} = \frac{(T+V)}{2}$$

The following identities were useful in the calculations described above

$$(\tau^a \varepsilon)_{ij}(\tau^a \varepsilon)_{aM} = -(\varepsilon_{ia}\varepsilon_{jM} + \varepsilon_{ja}\varepsilon_{iM})$$

$$i\varepsilon^{abc}\hat{r}^a(\tau^b\varepsilon)_{ij}(\tau^c\varepsilon)_{aM} = (\hat{r}.\vec{\tau}\varepsilon)_{ia}\varepsilon_{jM} + (\hat{r}.\vec{\tau}\varepsilon)_{ja}\varepsilon_{iM} - (\hat{r}.\vec{\tau}\varepsilon)_{ij}\varepsilon_{aM}$$

$$= \frac{1}{2}[(\hat{r}.\vec{\tau}\varepsilon)_{ia}\varepsilon_{jM} + (\hat{r}.\vec{\tau}\varepsilon)_{ja}\varepsilon_{iM} + (\hat{r}.\vec{\tau}\varepsilon)_{iM}\varepsilon_{ja} + (\hat{r}.\tau\varepsilon)_{jM}\varepsilon_{ia}]$$

So expansions (A4) and (A6) become identical if we identify (suppressing the index M)

$$\lambda = \frac{1}{2\sqrt{2}}(U_{\frac{-1}{2}1} + U_{\frac{1}{2}-1})$$

$$\theta = \frac{1}{2\sqrt{2}}(U_{\frac{-1}{2}1} - U_{\frac{1}{2}-1})$$

$$\xi = \frac{1}{2\sqrt{2}}(U_{\frac{-1}{2}1} - U_{\frac{1}{2}-1}) + \frac{1}{2}(U_{\frac{-1}{2}0} + U_{\frac{1}{2}0})$$

$$\chi = \frac{1}{2}(U_{\frac{1}{2}0} - U_{\frac{-1}{2}0})$$

Plugging all this into (A7) yields the result

$$(1-F)\frac{\sqrt{2}i}{r}[U^{\dagger}_{\frac{-1}{2}0}U_{\frac{1}{2}-1} + U^{\dagger}_{\frac{1}{2}0}U_{\frac{-1}{2}1} - U^{\dagger}_{\frac{1}{2}-1}U_{\frac{-1}{2}0} - U^{\dagger}_{\frac{-1}{2}1}U_{\frac{1}{2}0}]$$

which added to (A5) gives us the full T=1 Lagrangian.

Appendix B

The details of the evaluation of the fermion determinant is given here.

We have

$$\int [d\widetilde{U}][d\widetilde{U}]^M e^{\sum_M \int \widetilde{\bar{U}}(\tau^3(\partial_t - 2i\tau^2(\partial_t\omega)) + \tau^1\partial_r)\widetilde{U}^M dr dt} \tag{B1}$$

$$= [DetD]^2$$

where $D = \tau^3(\partial_t - 2i\tau^2(\partial_t\omega)) + \tau^1\partial_r = \gamma^0(\partial_t - i2\gamma^5\partial_t\omega) + \gamma^1\partial_r$

Use the fact that

$$\delta[\ln DetD] = \int dr dt\, \delta A. \langle J \rangle_A \tag{B2}$$

We have to evaluate $\langle J^0 \rangle$ since $A_1 = 0$

$$\langle j_\mu(x) \rangle = \langle 0 | \bar{\widetilde{U}} \gamma^\mu \gamma^5 \widetilde{U}(x) | 0 \rangle_A \tag{B3}$$

The subscript A indicates that there is a background A field. Since the product of operators at the same space time point is not well defined we use the point splitting procedure. Let x' and x be at equal times so x'-x is purely spacelike.

$$\langle J_\mu(x) \rangle = \lim_{x' \to x} Tr[\gamma^\mu\gamma^5 G(x,x') e^{-i2e\gamma^5 \int_{x'}^{x} A^\mu dx_\mu}] \tag{B5}$$

The factor $e^{-ie\gamma^5 \int_{x'}^{x} A.dx}$ is inserted in the usual way to make it gauge invariant. In our case $A_1 = 0$ and $A_0 = 2\partial_t\omega$.

$$\langle J^\mu(x) \rangle = \lim_{x' \to x} Tr[\gamma^\mu\gamma^5 e^{2i\gamma^5(\alpha(x)-\alpha(x')) - 2\sigma(x) + 2\sigma(x')} G_0(x,x') e^{-i\gamma^5 \int_{x'}^{x} A.dx}] \tag{B6}$$

We have used equation (13) for the full Green's function.

$$\langle J^\mu(x) \rangle = \lim_{x' \to x} Tr[[\gamma^\mu\gamma^5(1 + 2i\gamma^5\partial_1\alpha(x) + 2\partial_1\sigma(x))\varepsilon]\frac{1}{2\pi}[\frac{\gamma^1}{\varepsilon} - i\gamma^1\frac{2}{r}\gamma^0] e^{-ie\gamma^5 \int_{x'}^{x} A.dx} \tag{B7}$$

where $x' - x = \varepsilon$ and we have used the fact that t'-t = 0. Also $\int A.dx = \int_{x'}^{x} A^1 dx_1 = 0$ since $A^1 = 0$ Taking the limit $\varepsilon \to 0$ we are left with

$$\langle J^0 \rangle = \frac{1}{2\pi} Tr[2\gamma^0\gamma^5\gamma^1\partial_1\sigma(x)] = \frac{2}{\pi}\partial_1\sigma(x) \tag{B8}$$

Substituting B8 into B2

$$\delta[\ln Det D] = \frac{2}{\pi}\int drdt\, \delta A_0 \partial_1 \sigma(x)$$

$$= +\frac{4}{\pi}\int drdt\, \delta(\partial_t \omega)\partial_1 \sigma(x)$$

$$= \frac{-4}{\pi}\int drdt\, \delta(\partial_1 \partial_t \omega)\sigma(x) \quad + \quad \frac{4}{\pi}\int dt\, (\partial_t \omega)\sigma \Big]_{r=0}^{\infty}$$

The surface term vanishes because $\partial_t \omega(r=0) = 0$ and $\sigma(r=\infty) = 0$ Using (14)

$$\delta[\ln Det D] = \frac{-4}{\pi}\int drdt\, \delta(\partial^2 \sigma)\sigma \tag{B9}$$

$$= \frac{+2}{\pi}\int drdt\, \delta[(\partial_\mu \sigma)(\partial^\mu \sigma)]$$

$$\ln Det D = \frac{2}{\pi}\int drdt\, (\partial_\mu \sigma)^2$$

$$\ln Det D^2 = \frac{4}{\pi}\int drdt\, (\partial_\mu \sigma)^2$$

which is equation (15).

Appendix C

We reproduce here the equation for the σ propagator and its asymptotics. The action for the σ field is (eq. 17)

$$S(\sigma) = \frac{4\pi}{e^2}\int r^2(\Box\sigma)^2 dr dt + \frac{4}{\pi}\int(\partial_\mu\sigma)^2 dr dt \tag{17}$$

The propagator $K(\xi,\xi')$ satisfies:

$$[\frac{8\pi}{e^2}\Box(r^2\Box) - \frac{8}{\pi}\Box]K(\xi,\xi') = \delta^2(\xi-\xi') \tag{C1}$$

with boundary conditions $\partial_r K(r,t;r',t')|_{r=0} = 0$ Let us write $\frac{4}{e'^2} = \frac{1}{e^2}$ so that (C1) becomes

$$[8\frac{\pi}{e'^2}\Box(r^2\Box) - \frac{2}{\pi}]4K(\xi,\xi') = \delta(\xi-\xi') \tag{C2}$$

This equation for 4K is identical to the equation for the propagator in the isospinor case of Rubakov. We proceed in the same way: Let

$$4K(r,t;r',t') = -\frac{\pi}{2}[\Box^{-1}((r-r'),(t-t')) + \Box^{-1}((r+r'),(t-t')) - \tilde{K}(r,t;r',t')] \tag{C3}$$

Then $\tilde{K}$ satisfies

$$(\Box - \frac{e'^2}{4\pi r^2})\tilde{K}(r,t,',t') = \delta((r-r'))\delta((t-t')) \tag{C4}$$

The solution is

$$\tilde{K} = \frac{-1}{2\pi}Q_\nu[\frac{(r-r')^2 + (t-t')^2}{2rr'} + 1] \tag{C5}$$

with $\nu = \frac{1}{2}[(1+\frac{e'^2}{\pi^2})^{\frac{1}{2}} - 1]$ Some asymptotics are:

$$\tilde{K} \to \frac{1}{4\pi}\log[\frac{(t-t')^2 + (r-r')^2}{8rr'}] + const.$$

for r' -> r , t' -> t

$$\tilde{K} \to const.\,[\frac{rr'}{r'^2 + (t-t')^2}] \qquad r \to 0$$

$$\to const.\,(\frac{r}{r'})^{-1-\nu} \qquad r \to \infty$$

$$\rightarrow \left(\frac{(t-t')^2}{2rr'} \right)^{-1-\nu} \qquad |(t-t')| \rightarrow \infty$$

Appendix D

We would like to calculate the condensate formation in the case where we have two flavours of isodoublets and one isotriplet. In the notation of equation (19) define

$$Q_1(r,t) = \frac{1}{r^2}\tilde{U}^{\frac{1}{2}T}(r,t)\tilde{U}^{-\frac{1}{2}}(r,t) \tag{D1}$$

where $\pm\frac{1}{2}$ refers to the values of M. Similarly let

$$Q_{\frac{1}{2}}(r,t) = \frac{1}{r^2}\varphi^{aT}(r,t)\varphi^{b}(r,t) \tag{D2}$$

where φ^a and φ^b are two isodoublets and a,b stand for the two flavours. If the winding number is one the operator that gets an expectation value will be :

$$F(r,t) = [Q_1(r,t)]^2[Q_{\frac{1}{2}}(r,t)] \tag{D3}$$

We shall evaluate this by first evaluating as usual <0|F(r,t) F(r',t') |0> then letting $|(t-t\ )|\to\infty$ and using the cluster property. Let us write FF in terms of the Green's functions:

$$FF = Q_1(r,t)Q_1(r,t)Q_1(r',t')Q_1(r',t')Q_{\frac{1}{2}}(r,t)Q_{\frac{1}{2}}(r',t') \tag{D4}$$

Substituting for $Q_1, Q_{\frac{1}{2}}$ the expressions (D1) and (D2) we find on performing the Wick contractions:

$$\frac{1}{r^6 r'^6}\{2Tr[G_1G_1^T]Tr[G_1G_1^T] - 2Tr[G_1G_1^TG_1G_1^T]\}\{Tr[G_{\frac{1}{2}}G_{\frac{1}{2}}^T]\} \tag{D5}$$

Where G_1 is the Green's function for the isovector field and $G_{\frac{1}{2}}$ for the isospinor field φ. We find

$$Tr[GoGo^T] = 2[\frac{1}{(t-t\ )^2+(r-r\ )^2} + \frac{1}{(t-t\ )^2+(r+r\ )^2}] \tag{D6}$$

$$Tr[G_0G_0^TG_0G_0^T] = 2[(\frac{1}{(t-t\ )^2+(r-r\ )^2} + \frac{1}{(t-t\ )^2+(r+r\ )^2})^2$$
$$+ 4\frac{1}{((t-t\ )^2+(r-r\ )^2)((t-t\ )^2+(r+r\ )^2)}]$$

where

$$G_0 = \frac{\tau^3(t-t\ ) + \tau^1(\tau-\tau\ )}{(t-t\ )^2 + (\tau-\tau\ )^2} + \frac{(t-t\ ) - i\tau^2(\tau+\tau\ )}{(t-t\ )^2 + (\tau+\tau\ )^2}$$

is the free particle Green's function. Also

$$G_1 = e^{2i(\alpha(\xi) - \alpha(\xi'))\tau^2 + 2(\sigma(\xi) - \sigma(\xi'))} G_0(\xi,\xi') \tag{D7}$$

$$G_{\frac{1}{2}} = e^{i(\alpha(\xi) - \alpha(\xi'))\tau^2 + (\sigma(\xi) - \sigma(\xi'))} G_0(\xi,\xi')$$

Using (D6) and (D7)

$$FF = \frac{1}{(\tau\tau')^6} 2\left[\frac{1}{(t-t\ )^2 + (\tau-\tau\ )^2} - \frac{1}{(t-t\ )^2 + (\tau+\tau\ )^2}\right]^2 \tag{D8}$$

$$X\left[\frac{1}{(t-t\ )^2 + (\tau-\tau\ )^2} + \frac{1}{(t-t\ )^2 + (\tau+\tau\ )^2}\right] e^{10\sigma(\xi) - 10\sigma(\xi')}$$

We have to evaluate

$$\int [d\sigma] FF e^{S_{eff}(\sigma)}$$

In this case

$$S_{eff}(\sigma) = \frac{4\pi}{e^2}\int \tau^2(\Box\sigma)^2 d\tau dt + \frac{5}{\pi}\int (\partial_\mu \sigma)^2 d\tau dt$$

The coefficient 5 in the second term gets a contribution 1 from the isodoublets and 4 from the isotriplet. As in Appendix C we find the equation for K in this case is the same as in the isospinor case. Thus our K merely differs by a factorof 5 from the K of the isospinor case. Using this we find

$$\langle M|FF|M\rangle = \frac{1}{(\tau\tau')^6} 2\left[\frac{1}{(t-t\ )^2 + (\tau-\tau\ )^2} - \frac{1}{(t-t\ )^2 + (\tau+\tau\ )^2}\right]^2$$

$$X\left[\frac{1}{(t-t\ )^2 + (\tau-\tau\ )^2} + \frac{1}{(t-t\ )^2 + (\tau+\tau\ )^2}\right] e^{10(K(\xi,\xi) + 10K(\xi',\xi') - 20K(\xi,\xi'))}$$

As $T\to\infty$ the exponential factor goes as T^{10}. The net result is $\langle M|FF|M\rangle \approx \frac{1}{(\tau\tau')^9}$ Using the cluster property we get $\langle M|F|M\rangle \approx \frac{1}{\tau^9}$ as expected.

References

[1] V. A. Rubakov, Nucl. Phys. B203 (1982) 311

[2] C. Callan, Phys. Rev. D26 (1982) 2058; Phys. Rev. D25 (1982) 2141

[3] G. 't Hooft, Nucl. Phys. B79 276 ; A. Polyakov, JETP Lett 20 (1974) 194

[4] J. Preskill, Phys. Rev. Lett. 43 (1979) 1365

[5] B. Cabrera, Phys. Rev. Lett. 48 (1982) 1378

[6] C. Dokos, T. Tomaras, Phys. Rev. D21 (1980) 2940

[7] E. Kolb, Colgate, J. Harvey, Phys. Rev. Lett. 49 (1982) 1373

[8] A. A. Belavin et al. Phys. Lett. 59B 85, G. 't Hooft Phys. Rev. D14 3432

[9] J. Gervais, B. Sakita, Phys. Rev. D11 2943

[10] B. Julia, A. Zee, Phys. Rev. D11 (1975) 2227

[11] Y. Kazama, C.N. Yang, A. Goldhaber, Phys. Rev. D15 (1977) 2287

[12] A. Goldhaber, Phys. Rev. D16 (1977) 1815

[13] A. Blaer, N. Christ, J-F. Tang, Phys. Rev. Lett. 47 (1981) 1364

[14] E. Witten, Phys. Lett. 86B (1979) 283

[15] A. Schwartz et al. see ref. 8

[16] M. F. Atiyah, I. Singer, Ann. Math. 87,485

[17] J. Schwinger, Phys. Rev. 128 (1962) 2425

[18] R. Jackiw, C. Rebbi, Phys. Rev. D16 1052

[19] V. N. Gribov, Nucl. Phys. B206 (1982) 103

Figure Captions

[8] The 't Hooft-Polyakov monopole located at 'O' in the spherically symmetric gauge. The Higgs vev is in the isospin direction $\hat{r}.\hat{T}$.

[9] The functions H(r) and F(r) both go to 1 asymptotically. The core radius $r_M \approx \frac{1}{M_X}$.

[10] The heavy arrow indicates the helicity and the light arrow, the momentum. The helicity flips as the particle goes through the centre of the monopole.

[11] This configuration has winding number k.

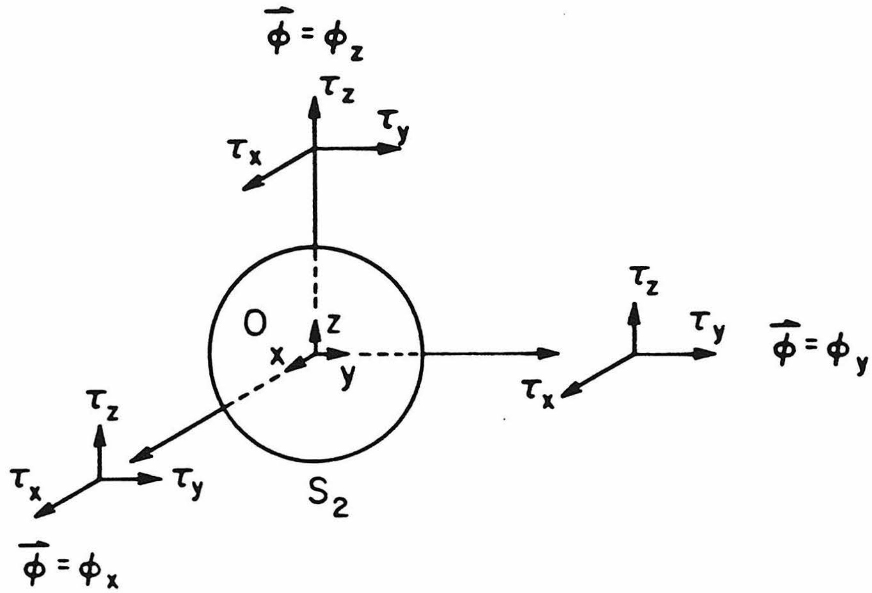
[12] The first column is a plot of the (Euclidean) time evolution of the function hg^{-1} (sec.1.3, eq.(5)) describing an instanton. The second column is the corresponding plot for a configuration with non-zero winding number in the presence of a monopole.

[13] The form of the condensate for the $T = 1$ system.

[14] The fermion determinant expanded in powers of the externa; photon field.

[15] A process that does not take place because it involves charge exchange and is energetically unfavourable.

[16] Proton decay. Note that p-> e^+ is not allowed by J conservation.



$\vec{\phi} = \phi_z$
τ_z
τ_y
τ_x
O
z
x
y
τ_z
τ_y
τ_x
$\vec{\phi} = \phi_y$
τ_z
τ_x
τ_y
S_2
$\vec{\phi} = \phi_x$

Fig. 8.

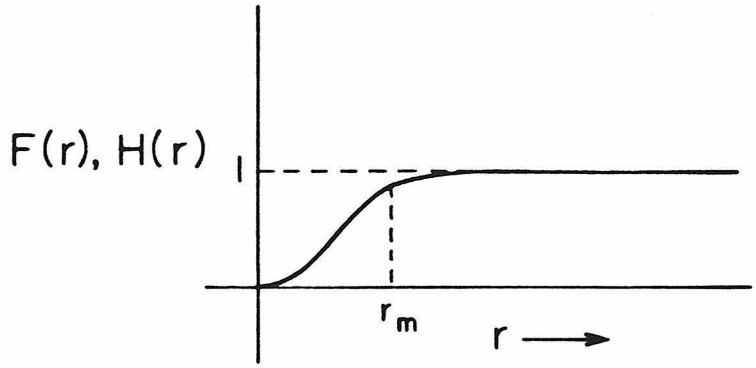


Fig. 9.

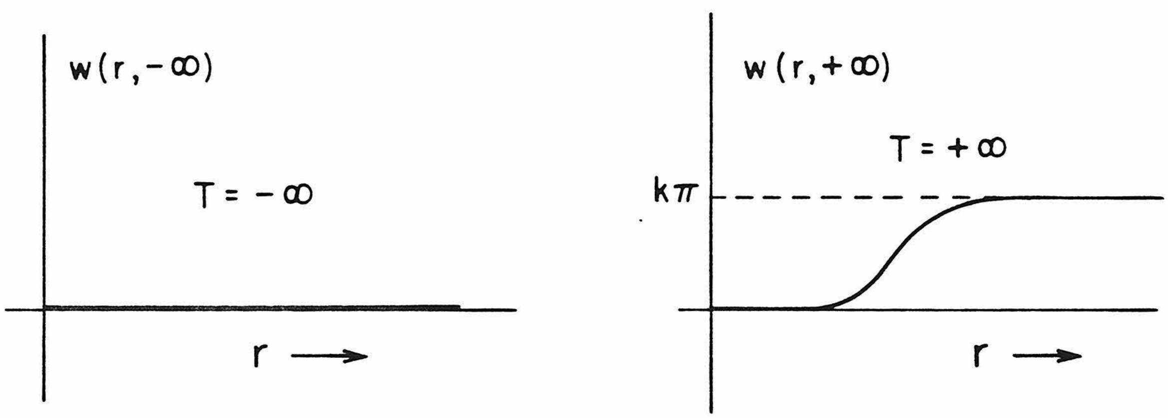


Fig. 11.

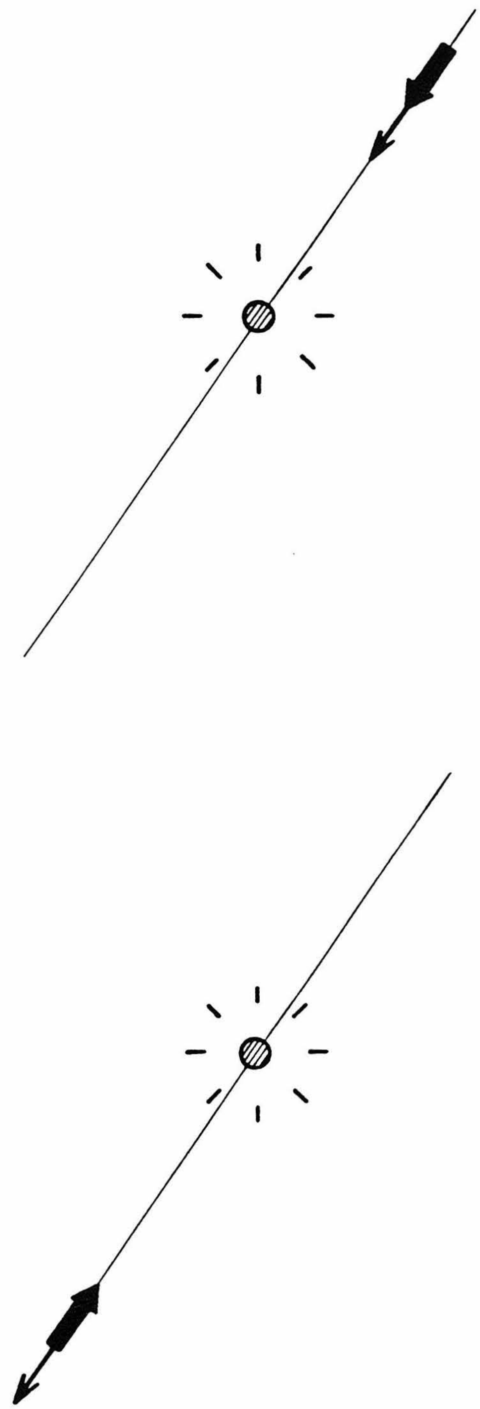


Fig. 10.

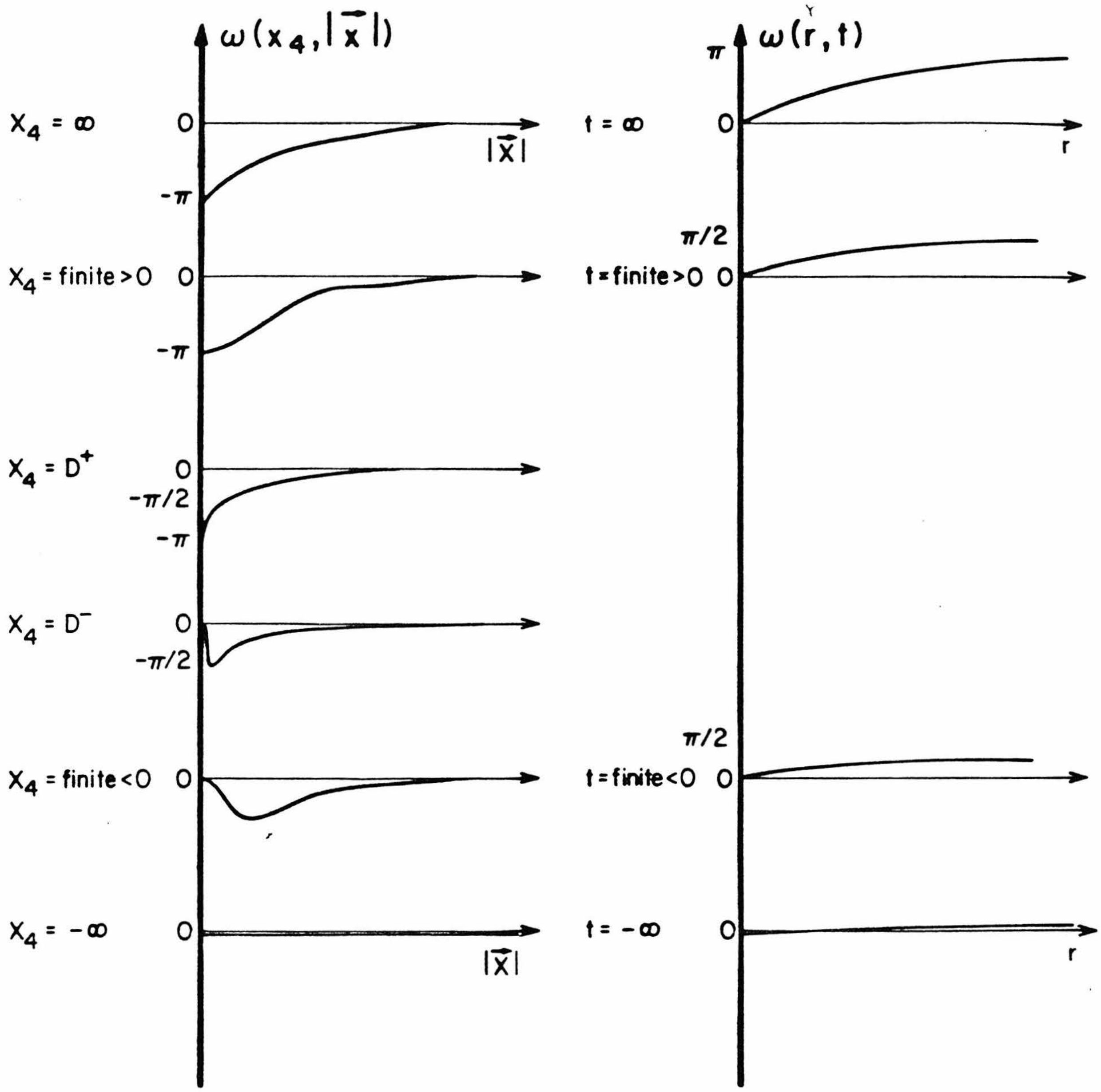

$\omega(x_4,|\vec{x}|)$
$X_4 = \infty$
0
$-\pi$
$|\vec{X}|$
X_4 = finite > 0
0
$-\pi$
$X_4 = D^+$
0
$-\pi/2$
$-\pi$
$X_4 = D^-$
0
$-\pi/2$
X_4 = finite < 0
0
$X_4 = -\infty$
0
$|\vec{X}|$
$\omega(r,t)$
π
$t = \infty$
0
r
$\pi/2$
t = finite > 0
0
$\pi/2$
t = finite < 0
0
$t = -\infty$
0
r

Fig. 12

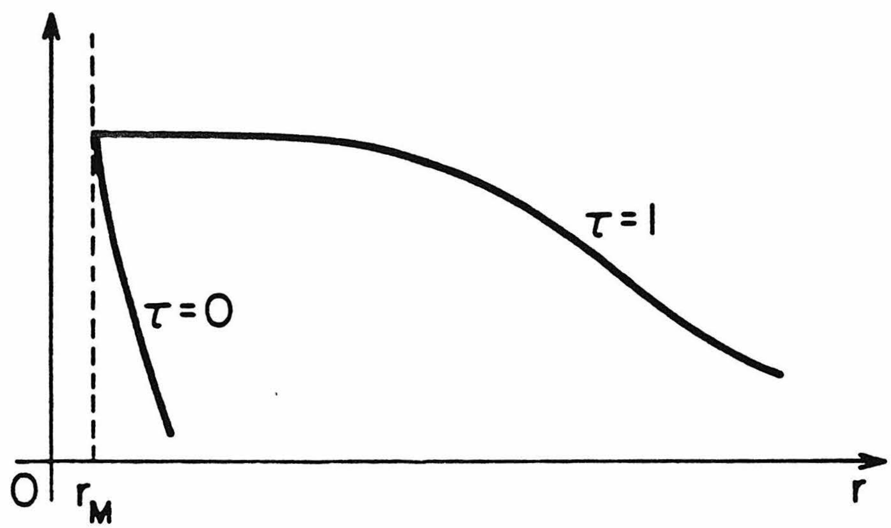


Fig. 13

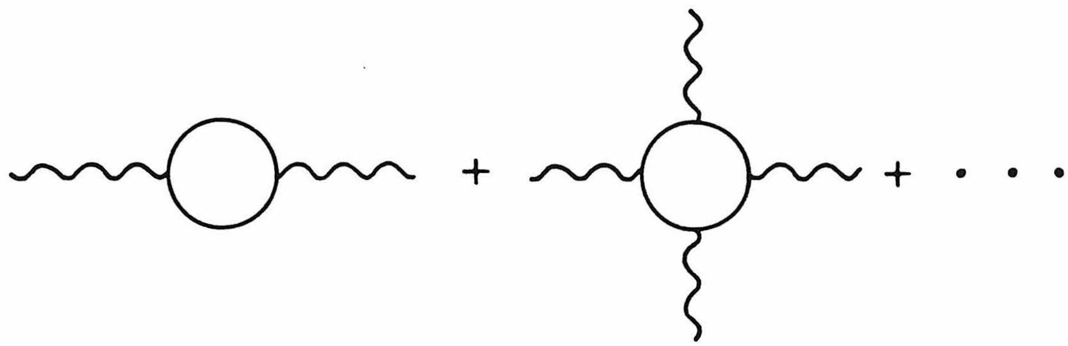


Fig. 14.

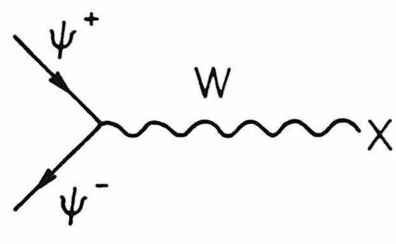


Fig. 15.

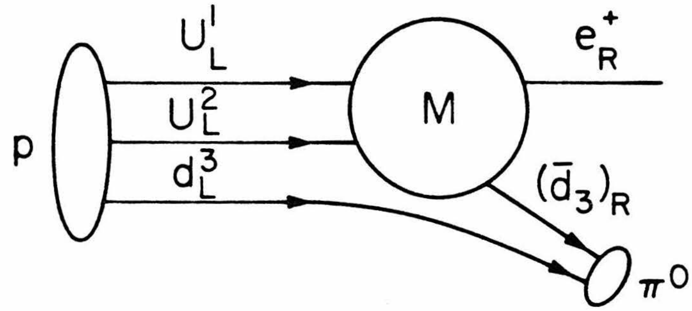


Fig. 16.