Nuclear Physics B185 (1981) 513–554

DYNAMICAL BREAKING OF SUPERSYMMETRY*

Edward WITTEN

Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544, USA

Received 29 April 1981

General conditions for dynamical supersymmetry breaking are discussed. Very small effects that would usually be ignored, such as instantons of a grand unified theory, might break supersymmetry at a low energy scale. Examples are given (in $0+1$ and $2+1$ dimensions) in which dynamical supersymmetry breaking occurs. Difficulties that confront such a program in four dimensions are described.

1. Introduction

Supersymmetry has fascinated particle physicists since it was first discovered [1]. It is an outstanding example of a known mathematical structure which may plausibly be absorbed in the future into our understanding of particle physics.

Of course, if nature really is described by a supersymmetric theory, the symmetry must be spontaneously broken [2]. At what energies does the symmetry breaking occur? It might very well occur at energies of order the Planck mass. In that case supersymmetry would be relevant to particle physics at "ordinary" energies only indirectly, in as much as the broken supersymmetry might make predictions concerning particle quantum numbers and relations among masses and coupling constants.

On the other hand, it is possible that supersymmetry breaking occurs at "ordinary" energies like a few hundred GeV or a few TeV. In this case, ordinary particle physics, at energies much less than the Planck mass, is presumably described by a renormalizable, globally supersymmetric model. There has been some success [3] in constructing realistic models of this sort for ordinary particle interactions.

If supersymmetry breaking does occur at ordinary particle physics energies, we must ask why the energy scale of supersymmetry breaking is so tiny compared to the natural energy scale of gravity and supergravity [4], which is presumably the Planck mass. This is a variant of the "hierarchy problem" [5]: why is the mass scale of ordinary particle physics so much less than the mass scale of grand unification or gravitation?

* Supported in part by NSF Grant PHY80-19754.

Presumably, if supersymmetry is spontaneously broken at the tree level, the breaking will have a strength of the same order as the natural mass scale of the theory. For supersymmetry to be broken only at, say, 10^3 GeV, which is 10^{-16} times the Planck mass of 10^{19} GeV, we require a theory in which supersymmetry is unbroken at the tree level and is broken only by extremely small corrections. These quantum corrections are presumably non-perturbative. We are looking for a theory in which supersymmetry is "dynamically broken" by non-perturbative effects.

If dynamical breaking of supersymmetry can occur, this could not only explain how supersymmetry could survive down to low energies and then be spontaneously broken. It might also resolve the usual hierarchy problem of understanding why the W and Z mesons are so light compared to the mass scale of possible grand unification and to the Planck mass. Once one can understand the existence of a "'low" mass scale of supersymmetry breaking, of order perhaps 10^3 GeV, it is perfectly possible that $SU(2) \times U(1)$ breaking could be part of this low energy symmetry breaking. In fact, one of the rather few phenomenological motivations for supersymmetry is precisely this $SU(2) \times U(1)$ hierarchy problem. For $SU(2) \times U(1)$ breaking to occur at a low energy scale, we need the usual Higgs doublet to be massless on the scale of grand unification or the Planck mass. According to our best understanding, masslessness of elementary charged scalars is not natural, except in supersymmetric theories. In supersymmetric theories (with supersymmetry not spontaneously broken) massless scalars occur naturally because scalars that are in the same supermultiplet with massless fermions or vector mesons must be massless. We may therefore imagine that the $SU(2) \times U(1)$ Higgs doublet is kept massless down to low energies by unbroken supersymmetry. Once supersymmetry is spontaneously broken the Higgs doublet need no longer be massless. The same non-perturbative effects that trigger supersymmetry breaking could therefore give the Higgs doublet a mass squared, which if it is negative will lead to the spontaneous breaking of $SU(2) \times U(1)$.

This scenario is made slightly more plausible by the fact that the Higgs doublet $\begin{pmatrix} \phi^0 \\ \phi^- \end{pmatrix}$ has the same $SU(3) \times SU(2) \times U(1)$ quantum numbers as the lepton doublet $\begin{pmatrix} \nu \\ e^- \end{pmatrix}_L$. This suggests that they may be supersymmetry partners. The fact that the leptons get their masses only from $SU(2) \times U(1)$ breaking is, of course, a consequence of gauge invariance provided that, as suggested by experiment, right-handed leptons transform differently under $SU(2) \times U(1)$ from the way the left-handed leptons transform. Therefore, if the Higgs doublet is related by supersymmetry to the leptons, it must be massless as long as supersymmetry and $SU(2) \times U(1)$ are unbroken, and its expectation value cannot be larger than the mass scale of supersymmetry breaking.

(I have tacitly used the fact that in global supersymmetry, which we are presuming to be what is relevant at low energies, a boson and its fermionic partner always have the same quantum numbers under gauge transformations. This is so because in global supersymmetry the supercharges always commute with gauge symmetries; the commutator, if not zero, would be a supersymmetry transformation depending on the infinite number of parameters of the gauge symmetry, so would have to be a local supersymmetry.)

In this paper I will discuss in a general way what would be involved in dynamical breaking of supersymmetry. I wish to warn the reader in advance that no essential problems are solved in this paper. I do not have a realistic model or even a realistic mechanism by which dynamical supersymmetry breaking can occur in four dimensions, and the discussion will be very qualitative. I hope, however, to raise some relevant issues. The plan of this paper is as follows. In sect. 2 some well-known facts are reviewed. In sect. 3 it is argued that unbroken supersymmetry is potentially unstable. In sect. 4 some grand unified models are described which could be realistic if dynamical supersymmetry breaking occurs. In sect. 5 the "breakdown" of naturalness in supersymmetric theories is discussed. In sect. 6 two models are presented (in less than four dimensions) in which dynamical supersymmetry breaking occurs. In sect. 7 the grand unified models of sect. 4 are re-examined, in the light of lessons from the discussions of naturalness and of soluble models. In sect. 8 it is suggested that gravity could play a role in dynamical supersymmetry breaking. In sect. 9 it is suggested that supersymmetry could cure the problems that plague theories of dynamical $SU(2) \times U(1)$ symmetry breaking. Some conclusions are drawn in sect. 10.

2. Supersymmetry and the vacuum energy

In this section some long-established aspects of spontaneously broken supersymmetry will be reviewed.

One of the central features of globally supersymmetric theories is that the hamiltonian H is the sum of the squares of the supersymmetry charges. In the simplest supersymmetry algebra there is a single Majorana spinor of four hermitian supersymmetry charges Q_α, $\alpha = 1 \ldots 4$. In terms of these we may write

$$H = \sum_{\alpha=1}^{4} Q_\alpha^2, \tag{1}$$

provided that a possible additive constant in H is chosen properly.

Since H is the sum of squares of hermitian operators, the energy of any state is positive or zero [6]. A state can have zero energy only if it is annihilated by each of the Q, since if $H|0\rangle = 0$, then $0 = \langle 0|H|0\rangle = \Sigma_\alpha \langle 0|Q_\alpha^2|0\rangle = \Sigma_\alpha |Q_\alpha|0\rangle|^2$, which is possible only if $Q_\alpha|0\rangle = 0$. If, conversely, a state is annihilated by the Q_a, then its energy is zero because $Q_\alpha|0\rangle = 0$ implies $H|0\rangle = \Sigma_\alpha Q_\alpha^2|0\rangle = 0$.

If there exists a supersymmetrically invariant state – that is, a state annihilated by the Q_α – then it is automatically the true vacuum state, since it has zero energy and any state that is not invariant under supersymmetry has positive energy. Thus, if a supersymmetric state exists, it is the ground state and supersymmetry is not spontaneously broken. Only if there does not exist a state invariant under supersymmetry is supersymmetry spontaneously broken. In this case the ground-state energy is positive.

In this respect supersymmetry is quite different from ordinary symmetries. With an ordinary internal symmetry, a symmetric state may exist without being the ground state. The situation has been illustrated [7] by fig. 1. In fig. 1a is shown a scalar potential which describes a theory with two ground states. In each ground state the scalar field has a vacuum expectation value, possibly breaking some internal symmetry. However, supersymmetry is not spontaneously broken, because the ground-state energy, the minimum value of the potential, is zero for each of the two possible states.

In fig. 1b, on the other hand, there is a unique ground state, the scalar field does not have an expectation value, and no internal symmetry is spontaneously broken. However, supersymmetry is spontaneously broken because the minimum value of the potential is positive. A state with $V=0$ does not exist.

Notice that despite the fact that supersymmetry is a Fermi-Bose symmetry, a non-zero vacuum expectation value of a scalar field does not necessarily mean that supersymmetry is spontaneously broken. Supersymmetry is spontaneously broken if the anticommutator of the supercharge Q_α with some operator X is non-zero:

$$\langle 0|\{Q_\alpha, X\}|0\rangle = \langle 0|(Q_\alpha X + XQ_\alpha)|0\rangle \neq 0, \tag{2}$$

since if $Q_\alpha|0\rangle = 0$ the expression in (2) would evidently have to vanish. However, in supersymmetric theories the elementary scalars ϕ can never be written as $\{Q_\alpha, X\}$, so they can obtain vacuum expectation values without necessarily breaking supersymmetry. (It is possible to write the derivatives $\partial_\mu\phi$ of the elementary scalars in the form $\{Q, \psi\}$, where the ψ are elementary fermions. However, whether supersymmetry is spontaneously broken or not, Lorentz invariance prevents $\partial_\mu\phi$ from having a vacuum expectation value.)

The fact that a positive vacuum energy indicates supersymmetry breaking is a special case of the fact that supersymmetry is spontaneously broken if any $\{Q, X\}$

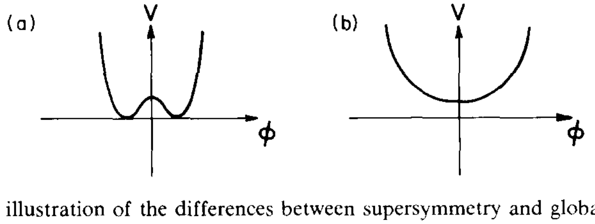

Fig. 1. A classical illustration of the differences between supersymmetry and global symmetries. In (a), the expectation value of the scalar field breaks an internal symmetry, but does not break supersymmetry, because the vacuum energy is zero. In (b), supersymmetry is spontaneously broken.

has a vacuum expectation value. In fact, the vacuum energy density E is defined in terms of the expectation value of the energy-momentum tensor $T_{\mu\nu}$:

$$\langle 0|T_{\mu\nu}|0\rangle = Eg_{\mu\nu}. \tag{3}$$

But in a supersymmetric theory $T_{\mu\nu} = (\gamma_\mu)_{\alpha\beta}\,\{Q_\alpha, S_{\nu\beta}\}$, where Q_α is the supercharge and $S_{\nu\beta}$ the supersymmetry current, so

$$Eg_{\mu\nu} = (\gamma_\mu)_{\alpha\beta}\langle 0|\{Q_\alpha, S_{\nu\beta}\}|0\rangle, \tag{4}$$

and a non-zero value for this expression means that supersymmetry is spontaneously broken.

The same sort of reasoning leads to a simple proof [8] of the analogue of Goldstone's theorem for supersymmetry. In the standard fashion of current algebra, one writes

$$\langle 0|\{Q_\alpha, S_{\nu\beta}\}|0\rangle = \int \mathrm{d}^4x \frac{\partial}{\partial x^\sigma}\langle 0|\mathrm{T}(S_{\sigma\alpha}(x)S_{\nu\beta}(0))|0\rangle, \tag{5}$$

where the equality depends upon the fact that $\partial_\sigma S^\sigma = 0$ and upon the definition that $Q_\alpha = \int \mathrm{d}^3x\, S_{0\alpha}$. In view of the anticommutation relation (4), the left-hand side of (5) is equal to $(\gamma_\nu)_{\alpha\beta}$ times the vacuum energy E. If (and only if) supersymmetry is spontaneously broken, E is non-zero and the right-hand side of (5) must be non-zero. Being the integral of a total divergence, the right-hand side of (5) can be non-zero only if there is a surface contribution. This means that a massless particle must be present.

A surface contribution can appear in (5) only if the two-point function $\langle 0|\mathrm{T}(S(x)S(0))|0\rangle$ vanishes for large $|x|$ only as $1/|x|^3$. The only intermediate state whose contribution would fall off so slowly would be a one-particle state containing a single massless fermion of spin one half. The massless fermion whose existence is so established is known as the Goldstone fermion.

Let us define the coupling f of the supercurrent to the Goldstone fermion by

$$\langle 0|S_{\mu\alpha}|\psi_\beta\rangle = f(\gamma_\mu)_{\alpha\beta}, \tag{6}$$

where $|\psi_\beta\rangle$ is a one fermion state with spin state $|\beta\rangle$. Then from eqs. (4) and (5) and the fact that only the one-fermion state contributes to the right-hand side of (5) follows a simple and fundamental formula,

$$E = f^2, \tag{7}$$

relating the vacuum energy density E to the coupling f.

For the sake of clarity, it should be noted that the supercurrent may create a massless fermion from the vacuum by means of a derivative coupling, $\langle 0|S_{\mu\alpha}|\psi_\beta\rangle =$

$gp_\mu\delta_{\alpha\beta}$, without supersymmetry being broken. Such a derivative coupling gives a matrix element that vanishes too rapidly to give a surface contribution in (5). Supersymmetry breaking is related specifically to the existence of the non-derivative coupling of eq. (6). In this paper, the phrase "the current creates a massless fermion from the vacuum" always refers specifically to a non-derivative coupling, as in eq. (6).

3. Supersymmetry and internal symmetries

The fact reviewed in the last section, that supersymmetry is spontaneously broken if and only if the vacuum energy is non-zero, leads to some basic differences between spontaneous breaking of supersymmetry and spontaneous breaking of ordinary (local or global) internal symmetries.

Let us ask under what conditions quantum corrections can change the pattern of symmetry breaking that one finds at the tree level. We will confine our attention primarily to theories in which the quantum corrections are weak – characterized by small, well-defined coupling constants. Can weak quantum effects change the pattern of symmetry breaking that one observes at the tree level?

A simple argument [9] shows that this ordinarily cannot happen in the case of internal symmetries. If an internal symmetry is unbroken at the tree level, this means (fig. 2a) that the classical potential has its minimum at a symmetrical point. Broken symmetry means (fig. 2b) that the classical potential has its minimum away from the origin. It is not possible by means of arbitrarily small corrections to turn a potential of type 2a into one of type 2b, or vice versa. Therefore, sufficiently weak quantum corrections will not break a symmetry that is unbroken at the tree level, nor will they restore a broken symmetry.

Although there is much truth in the above reasoning, some possible exceptions should be noted. An exception can arise if the tree potential (fig. 3) has a degeneracy between states of broken and unbroken symmetry. In this case the quantum corrections, no matter how small, are crucial in lifting the degeneracy and determining what is the true ground state. However, except by artificially adjusting the parameters, there is no known way to obtain this sort of degeneracy at the tree level. As shown by Coleman and E. Weinberg [10], an exception can also arise if the

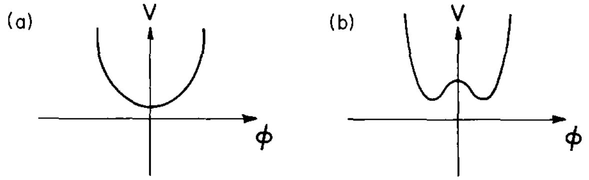

Fig. 2. In the case of internal symmetries, an arbitrarily small change in the parameters cannot produce broken symmetry, as in (b), from unbroken symmetry, as in (a).

curvature of the potential at the origin in field space is constrained to vanish. This vanishing curvature corresponds to an approximate degeneracy, since the energy is almost independent of the field. However, no known mechanism leads to vanishing curvature of the potential in a natural way. An exception might also arise if the quantum corrections to the effective potential were sufficiently singular near the origin in field space. This can occur in two dimensions in theories with massless fermions because of $\phi^2 \ln\phi$ terms in the effective potential, but does not seem to occur in four dimensions. Finally, one might wonder about theories like QCD. In QCD with massless quarks, chiral symmetry is unbroken at the tree level but is believed to be broken by quantum corrections. In this theory, because of the masslessness of the quarks, there are at the tree level states of non-zero chirality with arbitrarily low (although not zero) energy. It is this approximate degeneracy which makes it possible for quantum corrections to spontaneously break the symmetry, no matter how small the coupling is initially.

In the absence of the sort of degeneracy or approximate degeneracy just discussed, small quantum corrections (characterized by sufficiently small coupling constants) cannot change the pattern of symmetry breaking, because they cannot change a potential of type 2a into one of type 2b, or vice versa. Symmetries that are unbroken at the tree level are really unbroken.

In supersymmetry, these issues must be reconsidered, because the criteria for supersymmetry breaking are rather different. Let us refer back to fig. 1. In fig. 1a a scalar field has a vacuum expectation value, possibly breaking some internal symmetry. However, supersymmetry is *not* spontaneously broken, because the vacuum energy, the value of the potential at its minimum, is zero. In fig. 1b, on the other hand, the scalar field has zero vacuum expectation value and internal symmetries are not spontaneously broken. But supersymmetry is spontaneously broken, because the ground-state energy, the minimum of the potential, is positive.

It follows from this that if supersymmetry is broken in the tree approximation, then it really is broken in the exact theory, at least if the coupling is weak enough. Arbitrarily weak corrections cannot shift the minimum of the potential from the non-zero value of fig. 1b to zero. This could occur, if at all, only if the coupling constant exceeds some critical value. In this respect supersymmetry resembles ordinary internal symmetries.

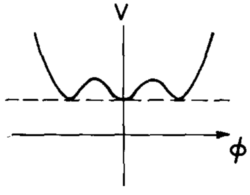

Fig. 3. If the potential possesses an accidental degeneracy at the tree level, quantum corrections will determine what is the ground state.

If we assume instead that supersymmetry is *not* broken at the tree level, the situation is very different, and arbitrarily weak quantum corrections could conceivably induce supersymmetry breaking. If in some approximation the minimum of the potential is zero, an arbitrarily small effect, shifting the potential by a tiny amount, could shift the minimum to a small but non-zero value (fig. 4). Then supersymmetry is spontaneously broken. It is potentially very delicate to claim that, in a given theory, supersymmetry is not spontaneously broken. An approximate calculation including many effects and showing that the vacuum energy is zero in a certain approximation always leaves open the possibility that even smaller effects that have been neglected could raise the ground-state energy slightly above zero. Unbroken symmetry could be unstable.

The above statement actually requires an important qualification. We know that if supersymmetry is spontaneously broken, there must exist a massless fermion, the Goldstone fermion. Weak quantum corrections will not bring into being a massless fermion if one does not already exist. If all fermions have non-zero mass at the tree level, weak corrections will not shift any of the fermion masses to zero. Consequently, in any theory in which supersymmetry is not broken at the tree level and in which all fermions have non-zero masses at the tree level, the supersymmetry must be truly unbroken, at least for weak enough coupling.

However, there are many reasons that a fermion might be massless other than its being a Goldstone fermion. Fermions might be massless because of unbroken chiral symmetries. As long as supersymmetry is unbroken, fermions may be massless because they are related by supersymmetry to massless gauge mesons or to massless Goldstone bosons.

The mere fact that a massless fermion exists does not mean *ipso facto* that supersymmetry is spontaneously broken. A Goldstone fermion is not simply a massless fermion. It is specifically a massless fermion that can be created from the vacuum by the supersymmetry current,

$$\langle 0|S_{\mu\alpha}|\psi_\beta\rangle = f(\gamma_\mu)_{\alpha\beta}, \tag{8}$$

with some non-zero f.

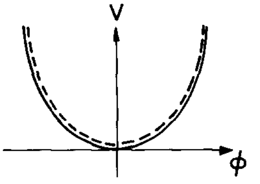

Fig. 4. In the case of supersymmetry, an arbitrarily small distortion of a potential with unbroken symmetry (solid line) can give a positive vacuum energy and therefore broken symmetry (dashed line).

Let us consider theories in which, at the tree level, supersymmetry is unbroken and there exist massless fermions which are not Goldstone fermions; that is, $f=0$ for each massless fermion. Theories of this sort would seem to be potentially unstable against supersymmetry breaking. Could not tiny quantum corrections give a non-zero value of f and thus make the massless fermion that already exists at the tree level into a Goldstone fermion? At the same time, these tiny quantum corrections would shift the minimum of the potential to a small non-zero level, which would be $E=f^2$ in view of the comments at the end of sect. 2.

For instance, even if f vanishes at the tree level, could not one-loop effects give a non-zero f of order α (and hence E of order α^2)? Or if f vanishes at the one-loop order, could not f arise as a two-loop effect, of order α^2?

The question is particularly interesting because any attempt at a realistic supersymmetric theory of nature would have, at the tree level, if supersymmetry is unbroken, massless fermions that could conceivably play the role just described. For instance, the supersymmetric partner of the photon would be a massless neutral fermion which might become a Goldstone fermion. If $SU(2)\times U(1)$ is unbroken at the tree level – as it must be if we are to use supersymmetry to solve the hierarchy problem – the supersymmetric partner of the Z meson is also massless and is another candidate Goldstone fermion. Realistic models might contain still more candidates.

A candidate Goldstone fermion must, of course, have the same quantum numbers as the supersymmetry current under all unbroken symmetries; otherwise the matrix element in eq. (8) would have to vanish. In particular, the would-be Goldstone fermion must be neutral under any gauge symmetries which are to remain unbroken. (Recall that in global supersymmetry, which we assume to be the relevant "low energy" approximation, the supersymmetry current is neutral under all gauge symmetries.) However, the candidate Goldstone fermions suggested above, the supersymmetry partners of the photon and Z meson, satisfy this condition.

The fact that unbroken supersymmetry could be unstable against weak corrections only if a massless neutral fermion exists at the tree level is analogous to the fact that unbroken internal symmetry is unstable only if there are degeneracies or approximate degeneracies at the tree level. But in practice there is a crucial difference. The difference is that massless neutral fermions inevitably exist in supersymmetric models of realistic interest.

Now we return to the previous question. If the coupling f of the supersymmetry current to a massless fermion vanishes at the tree level, can it receive a non-zero contribution from loop corrections?

The answer to this question is a quite remarkable surprise. By a detailed study of Feynman diagrams (done most conveniently in terms of superspace diagrams) it is possible to show that the answer is no. If f vanishes at the tree level, then f remains zero to all finite orders of perturbation theory. In the literature this has been stated in terms of the effective potential; it is stated that if the effective potential vanishes at some point in field space, then it vanishes at that point to all finite orders [11].

Since the effective potential at its minimum is the square of f, the two statements are equivalent.

The fact that f remains zero to all finite orders if it vanishes at the tree level is a special case of a more general phenomenon. In supersymmetric theories, the usual concept of naturalness does not apply, as long as supersymmetry is not spontaneously broken. The known facts in this area, and some corollaries, will be discussed in sect. 5.

No field theoretic reason is known for the fact that f remains zero to all finite orders. The existing proofs are based on details of perturbation theory (and are most tractable by use of the beautiful and efficient method of superspace perturbation theory [12]). Since the known proofs are based on details of perturbation theory, the result is not necessarily valid at the non-perturbative level.

If f became non-zero at, say, the one- or two-loop level, we would obtain a gauge hierarchy of some sort. The ratios of the squared masses of gauge bosons would be proportional to f and would be of order α or α^2. Such a hierarchy would not nearly be great enough to account for the observed ratio of energy scales in physics.

It is therefore very exciting that f is known to vanish to all finite orders, on the basis of arguments that do not necessarily apply non-perturbatively. If non-perturbative effects can give a non-zero f, we may obtain the desired enormous hierarchy. In this paper, an attempt will be made to explore this possibility.

4. "Realistic" grand unified models

Before discussing the status of naturalness in supersymmetric theories, and some possible mechanisms for dynamical breakdown of supersymmetry, let us first discuss some issues that arise in constructing realistic models of supersymmetry and grand unification.

To make a realistic model, we must first choose the symmetry algebra. So far we have discussed the simplest algebra, with a single spinor supercharge Q_α. More generally, there may be several spinor supercharges $Q_{\alpha i}$, $i = 1 \ldots N$. In this case the hamiltonian is obtained by summing the Q^2 over α, for fixed i. Specifically,

$$H = \sum_\alpha Q_{\alpha i}^2 = \sum_\alpha Q_{\alpha j}^2, \tag{9}$$

where one sums over α, but i or j is fixed.

In global supersymmetry (presumably what is relevant at low energies), realistic models are possible only for $N = 1$. This follows from one of the basic observations in particle physics: the massless fermions of helicity $\frac{1}{2}$ do not transform under $SU(3) \times SU(2) \times U(1)$ the same way the helicity $-\frac{1}{2}$ fermions transform. (Equivalently, the massless fermions of given helicity transform in a "complex" representation of the gauge group.) It is easy to see that in global supersymmetry with $N > 1$, the helicity $\frac{1}{2}$ and helicity $-\frac{1}{2}$ fermions necessarily transform identically. For instance, for $N=2$, the supersymmetry representations with massless particles of helicity $\frac{1}{2}$ contain the three helicities $(\frac{1}{2},0,-\frac{1}{2})$ or $(1,\frac{1}{2},0)$. Moreover, all particles in a given multiplet of global supersymmetry transform the same way under $SU(3)\times SU(2)\times U(1)$ (recall that in global supersymmetry, the supersymmetry charges commute with the group generators). The $(\frac{1}{2},0,-\frac{1}{2})$ multiplet relates fermions of helicity $\frac{1}{2}$ and $-\frac{1}{2}$, which would have to have the same quantum numbers. The $(1,\frac{1}{2},0)$ multiplet relates fermions of helicity $\frac{1}{2}$ to massless bosons of helicity one. But massless bosons of helicity one are always gauge bosons, transforming in the adjoint representation, which is real. (And there always are helicity -1 bosons transforming the same way; their partners would have helicity $-\frac{1}{2}$.) So whether we consider the $(1,\frac{1}{2},0)$ or the $(\frac{1}{2},0,-\frac{1}{2})$ multiplet, the fermions in $N=2$ (or $N>2$) global supersymmetry transform in a real representation of the gauge group; helicity $\frac{1}{2}$ and helicity $-\frac{1}{2}$ fermions transform equivalently.

To form a realistic model, we therefore should assume $N=1$.

Is it possible that N is greater than one microscopically, and that a "realistic" $N=1$ theory would be only a low energy approximation? Evidently, in this case, the fermions of the effective $N=1$ theory, transforming in a complex representation of the gauge group, must be absent in the microscopic lagrangian (they do not form representations of the $N>1$ algebra; and they cannot be put into such representations by adding additional particles that could receive mass at energies above $SU(2)\times U(1)$ breaking). If N is really greater than one in nature, one must suppose that the observed fermions have been generated dynamically at energies at which only the $N=1$ algebra (or no supersymmetry at all) is relevant.

Actually, in global supersymmetry, it is impossible for supersymmetry with $N>1$ to be spontaneously broken down to supersymmetry with $N=1$. This follows from eq. (9). If there is an unbroken supersymmetry, say $Q_{\alpha 1}|0\rangle=0$, then $H|0\rangle=\Sigma_\alpha Q_{\alpha 1}^2|0\rangle=0$. From this it follows that all of the supersymmetries are unbroken, because, for any i, $\Sigma_\alpha Q_{\alpha i}^2|0\rangle=H|0\rangle=0$.

In supergravity, it is possible to spontaneously break some supersymmetries without breaking all of them. Examples were first given by Scherk and Schwarz [13]. The possibility of doing this is related to the delicacy in defining global space-time transformations in general relativity. For some additional discussion, including some efforts to generate fermions in a complex representation of $SU(3)\times SU(2)\times U(1)$ after spontaneous breaking of some supersymmetries, see ref. [14].

The net conclusion is that supersymmetry with $N>1$ may be relevant to physics, but only at energies at which gravitation is important. $N=1$ supersymmetry may be relevant at ordinary energies.

Let us now turn to a consideration of some semi-realistic models. We must first recall some facts about construction of models with $N=1$ supersymmetry.

We may in general have an arbitrary gauge group, with gauge mesons A_μ^a and fermionic partners λ^a. In addition, we may introduce left-handed fermions ψ_L^i in an arbitrary multiplet of the gauge group. They form supersymmetry multiplets $\begin{pmatrix} \psi^i_{\mathrm{L}} \\ \phi^i \end{pmatrix}$ with complex scalar bosons ϕ^i. The right-handed fermion fields are the complex conjugates of the left-handed fermion fields, $\psi_{j\mathrm{R}} = (\psi^j_{\mathrm{L}})^*$ and their supersymmetry partners are the complex conjugates of ϕ^*_j of the ϕ^i.

The scalar potential of this theory is a sum of two terms. One term is derived from the gauge couplings. Let us group all of the scalar fields in a vector $\boldsymbol{\phi}$. Let T^a be the generators of the gauge group acting on the (possibly reducible) scalar representation. Then the term in the scalar potential coming from the gauge couplings is

$$V_1\left(\phi^i, \phi^*_j\right) = \sum_a \left(e_a(\boldsymbol{\phi}^*, T^a\boldsymbol{\phi})\right)^2, \tag{10}$$

where the sum runs over all generators, and e_a is the coupling constant associated with the generator T^a.

(In one case this formula must be generalized. If the gauge group is not semisimple, but contains a U(1) generator, say Y, with charge e, then it is possible [15] to define a supersymmetric theory in which the contribution $e^2(\boldsymbol{\phi}^*, Y\boldsymbol{\phi})^2$ in (10) is generalized to be $e^2\,((\boldsymbol{\phi}^*, Y\boldsymbol{\phi}) + \mu^2)^2$, where μ^2 is an arbitrary constant. This "D term" has been used in constructing realistic models [3].)

The second term in the scalar potential is related by supersymmetry not to the gauge couplings but to the Yukawa couplings. One begins by introducing a new function, sometimes called the superspace potential. The superspace potential W is a function of the ϕ^i but not of their complex conjugates the ϕ^*_j. For a renormalizable theory W should be at most cubic in the ϕ^i; otherwise W is restricted only by gauge invariance. The general form of W is

$$W(\phi^i) = a_i\phi^i + a_{ij}\phi^i\phi^j + a_{ijk}\phi^i\phi^j\phi^k,$$

where a_i, a_{ij}, and a_{ijk} are gauge-invariant tensors.

In terms of W, the Yukawa couplings of ϕ and ψ are defined by

$$L^{\mathrm{Yuk}} = \left(\partial^2 W/\partial\phi^i\partial\phi^j\right)\psi^i_{\mathrm{L}}\psi^j_{\mathrm{L}}.$$

In addition, there is a new contribution to the scalar potential,

$$V_2 = \sum_i \left|\frac{\partial W}{\partial\phi_i}\right|^2. \tag{11}$$

The total scalar potential is

$$V(\phi, \phi^*) = V_1 + V_2, \tag{12}$$

with V_1 and V_2 given in (10) and (11). Under what conditions is supersymmetry unbroken? It is required that V vanishes when evaluated at the expectation value of the scalar fields. Since V_1 and V_2 are each non-negative, they must both be zero. The condition for supersymmetry to be unbroken at the tree level is that for each a and each i, we require

$$(\phi^*, T^a\phi) = 0, \qquad \partial W/\partial\phi^i = 0. \tag{13}$$

If these equations have a simultaneous solution, supersymmetry is unbroken at the tree level. Otherwise, supersymmetry is spontaneously broken.

Now let us consider some more or less realistic examples of grand unified supersymmetric models. For simplicity we will take the gauge group to be SU(5). We wish a scalar potential which will at the tree level spontaneously break SU(5) down to SU(3) × SU(2) × U(1). If we wish to solve the hierarchy problem via dynamically broken supersymmetry, then supersymmetry should be unbroken at the tree level. However, models in which supersymmetry is broken at the tree level will also be considered below.

It is very easy to make a model in which, at the tree level, SU(5) is spontaneously broken to SU(3) × SU(2) × U(1) but supersymmetry is unbroken. This can be done rather simply if the fields ϕ^i discussed above consist of a single traceless complex matrix A^i_j transforming in the adjoint representation of SU(5). (The hermitian and antihermitian parts of A transform separately under SU(5), but must both be present because of supersymmetry.)

In this case the most general choice for W is

$$W = \tfrac{1}{2}M\,\mathrm{Tr}\,A^2 + \tfrac{1}{3}\lambda\,\mathrm{Tr}\,A^3, \tag{14}$$

where M and λ are constants, which can be taken to be real by redefining the overall phases of A and of W. [The phase of W always cancels out in expressions such as (11) for the physical scalar potential.] The first equation in (13) says that the hermitian and antihermitian parts of A commute and so can be diagonalized simultaneously by an SU(5) transformation.

The second equation in (13) gives

$$\lambda\left(A^2 - \tfrac{1}{5}\,\mathrm{Tr}\,A^2\right) + MA = 0. \tag{15}$$

There are three solutions, up to a gauge transformation:

(i) $$A = 0,$$

(ii) $$A = \frac{M}{3\lambda}\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & -4 \end{pmatrix}, \tag{16}$$

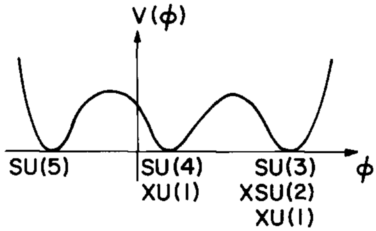

Fig. 5. A schematic illustration of a potential that does not determine the unbroken symmetry group.

(iii)
$$A=\frac{2M}{\lambda}\begin{pmatrix}1&&&&\\&1&&&\\&&1&&\\&&&-\frac{3}{2}&\\&&&&-\frac{3}{2}\end{pmatrix}.$$

The unbroken gauge symmetry is SU(5), SU(4) × U(1), or SU(3) × SU(2) × U(1), depending on which solution one considers.

Since these three states solve the equations (13) of unbroken supersymmetry, they all have zero energy at the tree level and so are completely degenerate. This is schematically depicted in fig. 5; a Higgs potential is sketched which has three precisely degenerate minima, each at zero energy. To finite orders in perturbation theory, as discussed in the last section, the degeneracy will not be lifted. The three states each remain at zero energy.

It is possible that non-perturbative effects do not change this picture at all. Perhaps supersymmetry remains unbroken – when all non-perturbative effects are included – in each of the three states. In that case this theory has three equally valid vacuum states, with different unbroken gauge symmetries and different spectra of the elementary particles.

It is possible that non-perturbative effects leave supersymmetry unbroken in one or two of the ground states but not in the others. If (fig. 6a) there is precisely one state in which non-perturbative effects leave supersymmetry unbroken, then this is uniquely selected as the true ground state of lowest energy.

On the other hand, it may be (fig. 6b) that non-perturbative effects lift the ground-state energy to tiny but positive values in each of the three vacuum states, so that supersymmetry is spontaneously broken for any choice of the ground state. The true ground state will then be the one in which supersymmetry is broken most weakly and the energy is smallest. If this is the state in which the unbroken gauge symmetry at the tree level was SU(3) × SU(2) × U(1), and if the non-perturbative effects that break supersymmetry also break SU(2) × U(1), one may obtain a good description of nature with a large gauge hierarchy.

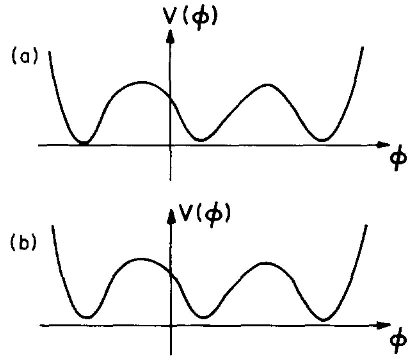

Fig. 6. Small quantum corrections to the potential of the previous drawing have determined the true ground state. If dynamical supersymmetry breaking occurs in each minimum but one, that one is the true vacuum (a). If supersymmetry is broken in each minimum, then the true vacuum (b) is one in which supersymmetry is broken most weakly.

Of course, to make the model realistic we must introduce quarks and leptons as well as the fields considered above. In accordance with the standard SU(5) phenomenology, we should introduce matter multiplets transforming in the 10 and $\bar{5}$ representations of SU(5). So for each quark-lepton generation we have a supersymmetry multiplet $\Phi^{ij} = \begin{pmatrix} \phi^{ij} \\ \psi^{ij} \end{pmatrix}$ in the 10 representation ($\Phi^{ij} = -\Phi^{ji}$), and an additional $\bar{5}$ multiplet $\Phi_i = \begin{pmatrix} \phi_i \\ \psi_i \end{pmatrix}$. In view of some phenomenological problems that will be mentioned shortly, it is probably necessary to include some additional matter multiplets transforming in a real representation of SU(5). One then generalizes the function W of eq. (14) to a gauge-invariant function $W(A^i_j, \phi^{kl}, \phi_m)$ with cubic and lower powers of the fields. This need not disturb the picture described above.

Apart from the question of whether non-perturbative effects really do spontaneously break supersymmetry, there are two rather serious phenomenological problems that arise in a model of this kind.

(i) It may be awkward to obtain a realistic quark-lepton mass spectrum. For instance, if the only matter multiplets other than A are the SU(5) 10 and $\bar{5}$ fields ϕ^{ij}_a and ϕ_{kb} (a and b are flavor indices), then the only term involving these fields that can be added to W is $\Delta W = g_{abc}\phi^{ij}_a\phi_{ib}\phi_{jc}$, where the g_{abc} are arbitrary coupling constants (restricted only by $g_{abc} = -g_{acb}$). If one assumes that along with dynamical supersymmetry breaking one of the $SU(2)\times U(1)$ doublet fields in ϕ_{kb} will get an expectation value, then the Yukawa couplings associated with ΔW will give masses to the down quarks and charged leptons. However, the up quarks will be massless at the tree level; the Yukawa couplings required to give them masses are forbidden by supersymmetry. Even if the up quarks receive masses in higher orders, these masses would be unacceptably small, at least in the case of the top quark. Because of the antisymmetry of g_{abc} in the last two indices, in this model the down quark and charged lepton in the same generation as the scalar that has an expectation value would also be massless at the tree level; in view of the lightness of the down quark and electron, this is phenomenologically acceptable.

(ii) It is difficult in this sort of model to obtain a realistically long proton lifetime unless symmetries are assumed that make the proton stable. In ordinary SU(5), the color triplet scalars that are related by SU(5) to the Higgs doublet can mediate baryon decay. However, one can assume an arbitrarily large mass for these scalars. Here, these color triplet scalars cannot be assumed to be arbitrarily heavy, for they are supersymmetry partners of the down, strange, and bottom quarks, and we are assuming that supersymmetry is unbroken down to low energies. The baryon non-conserving couplings of these scalars – which are related by supersymmetry to the ΔW term discussed above – are a serious problem.

To overcome these problems it is necessary to introduce additional matter fields beyond the $\bar{5}$ and 10, and to assume either some global symmetries or a larger gauge group.

We will return in sects. 6, 7 to the question of whether in a model like this one supersymmetry really is dynamically broken. Here let us consider the other logical possibility that arises when one considers supersymmetry together with grand unification. Perhaps supersymmetry is spontaneously broken at the tree level: at the energies of grand unification.

This will occur if and only if eqs. (13) are inconsistent and have no solution. Actually, $\phi = 0$ always satisfies the first equation, so as a necessary condition for supersymmetry to be broken at the tree level, it is necessary that $\partial W/\partial\phi_i = 0$ is not satisfied at $\phi = 0$. This is possible only if W contains a term which is linear in the ϕ fields, which is in turn possible only if at least one of the ϕ fields is a singlet under the gauge group.

For a "minimal" model of this kind, we may introduce one complex singlet field X and two complex fields A^i_j and B^i_j, each in the adjoint representation of SU(5). For W many choices are possible. The following choices, among many others, are interesting. One may take

$$W(X, A, B) = M^2X - gX\,\mathrm{Tr}\,A^2 + \lambda\,\mathrm{Tr}\,AB, \tag{17}$$

$$W(X, A, B) = M^2X - gX\,\mathrm{Tr}\,A^2 + \lambda\,\mathrm{Tr}\,A^2B, \tag{18}$$

where M, g, and λ are parameters which by redefining fields can be assumed to be real and positive. These choices of W give models analogous to the O'Raifeartaigh model [2].

The choices of W in (17) and (18) give theories that are technically natural because of global symmetries analogous to those in the O'Raifeartaigh model. One must bear in mind that any change in the fields under which W changes only by an overall phase is a symmetry of the theory, because the phase of W cancels out in the Higgs potential [eq. (11)] and can be removed from the Yukawa couplings by chiral rotations of the Fermi fields. Eq. (17) is technically natural because of symmetries under $X \to e^{i\alpha}X$, $B \to e^{i\alpha}B$ and under $X \to -X$, $A \to -A$. Eq. (18) is technically natural because of symmetry under $X \to e^{i\alpha}X$, $B \to e^{i\alpha}B$ and under $A \to -A$. However, as will be discussed in the next section, the concept of naturalness does not apply in the usual way to supersymmetric theories. One could just as well consider choices of W that are not technically natural.

It is easy to see that either choice of W given above leads to a theory in which supersymmetry is spontaneously broken at the tree level, because eqs. (13) for unbroken supersymmetry are inconsistent and do not have a solution. The equation $\partial W/\partial X = 0$ requires $M^2 = g\,\mathrm{Tr}\,A^2$, but the equation $\partial W/\partial B^i_j = 0$ requires $A^i_j = 0$. These are obviously incompatible, so supersymmetry is spontaneously broken.

The next step is to minimize the potential of the scalar fields, which can be read off from eqs. (10)–(12). For the choice of W in eq. (17) the potential is

$$V = |M^2 - g\,\mathrm{Tr}\,A^2|^2 + \lambda^2 \mathrm{Tr}\,AA^* + \mathrm{Tr}(\lambda B - 2gXA)(\lambda B^* - 2gX^*A^*)$$
$$+ e^2 \mathrm{Tr}(i[A, A^*] + i[B, B^*])^2, \tag{19}$$

where e is the SU(5) coupling constant. Minimization of this potential does not uniquely determine the pattern of symmetry breaking. The general minimum of the potential is as follows: A is any hermitian matrix with $\mathrm{Tr}\,A^2 = (2gM^2 - \lambda^2)/(2g^2)$, X is arbitrary, and $B = 2gXA/\lambda$. There are many massless scalars at the tree level. A one-loop calculation must be performed to lift the degeneracy and determine the expectation values of A and X. While the unbroken symmetry may turn out to be $SU(3) \times SU(2) \times U(1)$, there are many other possibilities, such as $SU(4) \times U(1)$.

If one considers instead the choice of W in eq. (18), the potential is

$$V = |M^2 - g\,\mathrm{Tr}\,A^2|^2 + \lambda^2\left(\mathrm{Tr}\,A^2A^{*2} - \tfrac{1}{5}\mathrm{Tr}\,A^2\,\mathrm{Tr}\,A^{*2}\right)$$
$$+ \mathrm{Tr}|\lambda(AB + BA - \tfrac{2}{5}\mathrm{Tr}\,AB) - 2gXA|^2 + e^2\mathrm{Tr}(i[A, A^*] + i[B, B^*])^2. \tag{20}$$

One now finds that A is uniquely determined to be

$$A = \frac{M}{\sqrt{30g + \lambda^2}} \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & -3 & \\ & & & & -3 \end{pmatrix} \tag{21}$$

in a suitable basis. The unbroken gauge group is thus uniquely determined to be $SU(3) \times SU(2) \times U(1)$. The vacuum expectation value of X is again not determined at the tree level, and B is proportional to A.

It should be obvious that a theory of this kind, once quarks and leptons are introduced, can give just as good an account of nature as is given by standard grand unified theories. In fact, since supersymmetry is broken at very high energies, the difficulty is really whether it is possible to distinguish these theories in terms of low energy predictions from standard grand unified theories without supersymmetry.

One possibility for distinguishing this sort of theory from conventional grand unified theories arises from the fact that supersymmetry forbids certain Yukawa couplings and therefore one may obtain new relations among quark and lepton masses. (Even though supersymmetry is badly broken, the corrections to supersymmetry constraints on Yukawa couplings, arising from loop diagrams, will be of order α.) However, somewhat as in conventional grand unification, the simplest models give mass relations that are too restrictive. For instance, if one enlarges the models described above by adding left-handed fields in the $\bar{5}$ and 10 representations only (and no other new fields), one finds that (as in the models discussed previously with supersymmetry unbroken at the tree level) all up quarks are massless at the tree level.

The key hurdle that this type of theory must face is, of course, whether the hierarchy problem can be solved. Having obtained strong breaking of supersymmetry and SU(5), can one also obtain the extremely weak breaking of $SU(2) \times U(1)$ that is needed to describe nature?

The models considered above illustrate the fact that in theories with supersymmetry breaking at the tree level, there typically are massless scalars at the tree level. The X particle was massless in both models, and in the first model there also were massless charged scalars. When quarks and leptons are incorporated in the model, the situation becomes far more dramatic. If one includes quarks and leptons in the minimal fashion described above, by introducing supermultiplets $\begin{pmatrix} \psi_L \\ \phi \end{pmatrix}$ in the $\bar{5}$ and 10 representations, it is almost obvious that *all* the scalar partners of the quarks and leptons are automatically massless at the tree level. This is so because it is impossible to write a gauge invariant quadratic or cubic term in W that will couple $\bar{5}$ and 10 fields to the fields A, B, and X considered previously. So at the tree level, even though supersymmetry has been spontaneously broken, the scalar partners of quarks and leptons do not "know" this; they remain degenerate with the fermions, and therefore massless.

Thus, the supersymmetry partners of $\begin{pmatrix} \nu \\ e \end{pmatrix}_L$, which have the quantum numbers of the usual Weinberg-Salam Higgs doublet, are massless at the tree level. Unfortunately, there also are at the tree level many massless colored scalars – partners of the quarks – which are dangerous because their exchange could violate baryon number. One would ordinarily assume that when loop corrections are considered, all of the massless scalars (with or without color) will acquire large mass. If this is so, the danger of scalars mediating baryon decay will be eliminated. But the possibility of low energy $SU(2)\times U(1)$ breaking would also be eliminated.

To make a successful model of this sort one must find a reason that the usual Higgs doublet – but not the colored scalars – remains massless in finite orders of perturbation theory. This would violate usual concepts of naturalness. In the next section we will discuss some respects in which usual concepts of naturalness are violated by supersymmetric theories. However, as far as I know, this is restricted to theories in which supersymmetry is not spontaneously broken. I cannot suggest any solution of the hierarchy problem in models where supersymmetry is broken at the tree level.

5. Naturalness in supersymmetric theories

We have mentioned before that, when there are massless fermions at the tree level, unbroken supersymmetry would seem to be potentially unstable against perturbative corrections. Nevertheless, it is known that this does not occur. It is known that if the matrix elements $\langle 0|S_{\mu\alpha}|\psi_\beta\rangle$ vanish at the tree level, they vanish to all finite orders, even if some of the $|\psi_\beta\rangle$ share all quantum numbers with the $S_{\mu\alpha}$.

Stated in this way, the fact that unbroken supersymmetry is stable against perturbative corrections seems to violate our usual concepts of naturalness. Indeed, there are a variety of respects in which supersymmetric theories do not satisfy properties that one would ordinarily regard as consequences of naturalness. The known arguments concerning the breakdown of naturalness in supersymmetric theories all depend on details of perturbation theory. No general argument is known that is based purely on symmetry and invariance principles. It is therefore an open question whether the breakdown of naturalness (in several respects that will be discussed) and the stability of unbroken supersymmetry are valid at the non-perturbative level.

Let us now consider the basic facts in this area. By far the most efficient way to construct supersymmetric invariants in $N=1$ global supersymmetry is the superspace formalism [12]. One has matter superfields

$$\Phi(x)=\phi(x)+\theta^\alpha\psi_\alpha(x)+\theta_\alpha\theta^\alpha F(x),$$

and spinor superfields

$$F_\alpha(x)=\lambda_\alpha(x)+F_{\alpha\beta}(x)\theta^\beta+\theta_\alpha D(x)+\theta^\beta\theta_\beta D_{\alpha\alpha'}\bar{\lambda}^{\alpha'}$$

for the gauge fields. The most general superfield invariant under arbitrary superspace gauge transformations is a product of superfields of the above-mentioned type and their covariant derivatives.

Given a Lorentz-invariant superfield Q, one can always make a supersymmetric invariant (perhaps zero) by integrating it over all of superspace:

$$I=\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\tilde{\theta}\,Q(x,\theta,\tilde{\theta}\,). \tag{22}$$

In this expression θ and $\tilde{\theta}$ are the anticommuting coordinates of negative and positive chirality. In addition, if one has a Lorentz-invariant superfield $R(x,\theta)$ which is a function of x and θ only (but not $\tilde{\theta}$) one can form a supersymmetric invariant by integrating over x and θ:

$$I=\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,R(x,\theta). \tag{23}$$

Obviously, any invariant which can be written in the form (22) can also be written in the form (23). It is enough to define

$$R=\int \mathrm{d}^2\tilde{\theta}\,Q. \tag{24}$$

However, it is important to realize that there exist supersymmetry invariants which can be written in the form (23) but cannot be written in the form (22) with any superfield Q.

For instance, all mass terms and all Yukawa couplings in renormalizable supersymmetric theories come from operators of type (23) which cannot be written in the form (22). (The integrand is the function W, discussed in the previous section. The only exception is the D term, to be discussed.) On the other hand, the kinetic energy terms, both for matter fields and for gauge fields, can be written in the form (22). The status of the gauge field kinetic energy is actually somewhat delicate. It can be written in the form (23) with a gauge-invariant integrand; it can also be written in the form (22), but with an integrand that in this case is not gauge invariant under arbitrary superspace gauge transformations, but only under those that preserve the Wess-Zumino gauge condition (these include ordinary space-dependent gauge transformations). To complete our survey of supersymmetric invariants, it should be noted that there is one operator whose status is somewhat anomalous. This is the "D term" which can appear when the gauge group is not semisimple but has a U(1) factor. It cannot be written in the form (23), and when it is written in the form (22), the integrand is only invariant under Wess-Zumino gauge transformations. We will at first assume that the gauge group is semisimple, and postpone to the end the complications associated with the D term.

Now, it has been proved [16] that to any finite order of perturbation theory, quantum corrections to the effective potential generate only operators of type (22), never operators that can only be written in the form (23). (The technique of superspace perturbation theory [12] is a great aid in simplifying these arguments.) The fact that only operators of type (22) appear in the effective potential has a number of interesting corollaries, many of which have been noted in the literature. (In discussing these corollaries, it must be kept in mind that the gauge kinetic energy, and also the D term, to which we will return later, are to be regarded as operators of type (22); this depends on the details of how the Wess-Zumino gauge enters in the formalism used in proving that only operators of type (22) are generated.)

First of all, because the kinetic energy operators are of type (22), wave-function renormalization occurs for both matter and gauge fields. But because the mass terms and Yukawa couplings come from operators of type (23), there is no renormalization of these parameters independent of the wave-function renormalization. The statement "no renormalization" refers to finite as well as infinite contributions. The bare mass terms and Yukawa couplings come from terms in the bare potential of type (23), and no additional finite or infinite contributions of type (23) are generated in perturbation theory. In supersymmetric theories one can, if one wishes, impose arbitrary relations among mass and Yukawa coupling parameters or set some of them equal to zero.

The fact that only operators of type (22) are generated in perturbation theory has another consequence that is very important for our purposes. This is the fact that supersymmetry remains unbroken to any finite order if it is unbroken at the tree level. This has often been expressed [6] as the statement that the vacuum energy is zero to all orders if it is zero at the tree level.

Let us review a few facts. Given any elementary fermion ψ_α, from its commutator we can form a scalar field $\{Q_\alpha, \psi^\alpha\}$ which is known as the auxiliary field and is denoted as F or D depending on whether ψ is related by supersymmetry to a boson of spin zero or spin one. If the auxiliary field has a vacuum expectation value supersymmetry is spontaneously broken, since obviously supersymmetry is spontaneously broken if $\langle\{Q_\alpha, \psi^\alpha\}\rangle \neq 0$. The converse is also true in the context of perturbation theory. If supersymmetry is spontaneously broken in finite orders of perturbation theory, one of the elementary fermions must be a Goldstone fermion, $\langle 0|S_{\mu\alpha}|\psi_\beta\rangle = \gamma_{\mu\alpha\beta} f$, with $f \neq 0$. (If supersymmetry breaking is a non-perturbative effect, the Goldstone fermion might be a bound state.) But if ψ appears as a one-particle pole in the two-point function $\langle 0|\mathrm{T}(S_{\mu\alpha}\psi_\beta)|0\rangle$, then standard current algebra (by considering the divergence of that matrix element) implies that the auxiliary field has an expectation value (equal, in fact, to f).

To obtain supersymmetry breaking in finite orders of perturbation theory, we must give an expectation value to F or D. This means that we must obtain in the supersymmetric effective potential an operator linear in F or D, times fields with vacuum expectation values. The only fields with vacuum expectation values are elementary scalar fields, since neither gauge fields nor Fermi fields nor derivatives of fields have expectation values. When one constructs superfields whose integrals would be gauge invariant, if one ignores terms involving Fermi fields, gauge fields, and derivatives, each power of F is accompanied by at most two powers of the anticommuting coordinates θ^{α}. The situation for D is more complicated and will be discussed momentarily.

Operators linear in F, which are at most quadratic in the θ_{α}, can survive an integral of the type $\int d^2\theta$ that appears in eq. (23). However, as we have noted, no operators of (23) are generated by loop corrections to the effective potential, to any finite order. Non-zero contributions in the integral $\int d^2\theta d^2\tilde{\theta}$ appearing in (22), if they involve F at all, involve F quadratically (or involve F times D) so as to be quartic in θ and $\tilde{\theta}$. So a term linear in F does not appear, and F does not get an expectation value.

Let us now consider D. While there is an auxiliary field for every generator of the gauge group, the D fields that could get vacuum expectations values are those associated with *unbroken* gauge symmetries. Only those D fields are related by supersymmetry to massless gauge mesons and to massless fermions that might become Goldstone fermions. A D field associated with a broken symmetry is not in lowest order related by supersymmetry to any massless fermion and therefore it could not get a vacuum expectation value in the lowest order of perturbation theory in which supersymmetry is spontaneously broken and some fermion becomes a Goldstone fermion.

Let us consider first semisimple groups (an important complication that arises otherwise will be described later). Operators of type (22) can contain terms linear in D (the matter field kinetic energy is an example). However, the linear terms have the following structure: a D field associated with a given generator of the gauge group always multiplies scalar fields that are non-singlets under the action of that generator. A D field associated with an unbroken gauge symmetry is always multiplied by charged scalar fields that have zero expectation value precisely because the symmetry is unbroken. Hence no expectation value of the D term is induced.

The non-generation of operators of type (23) has some other interesting corollaries. In a broad class of theories, one can prove [11] that if supersymmetry is unbroken at the tree level, then loop corrections do not induce shifts in the vacuum expectation values of the scalar fields*. This can be proved, in those cases where it is true, by showing that all operators of type (22) are stationary when the fields are set equal to their tree approximation vacuum expectation values. For instance, in all theories of spin 0 and spin $\frac{1}{2}$ fields only, we have seen that the effective potential is at least quadratic in the F fields. So it is stationary at any point where the F fields vanish. But the vanishing of the F fields is the condition that defines the tree level vacuum expectation values of the fields.

Another interesting corollary of the non-generation of operators of type (23) is that in a broad class of supersymmetric theories, one can show that if supersymme-

* However, B. Ovrut and J. Wess have recently found a class of theories in which this is not so (private communication).

try is unbroken at the tree level, then any particle massless at the tree level remains massless to all finite orders of perturbation theory, even if its masslessness resulted from arbitrary adjustment of parameters. (I do not know if this is true in all supersymmetric theories.) Let us again for simplicity concentrate on the case of theories with spin 0 and spin $\frac{1}{2}$ fields only. The F fields in general contain terms linear and quadratic in the elementary scalar fields. But *massless* scalars do not appear in the linear terms in F; if they did they would not be massless, since at the tree level the potential is $\frac{1}{2}\Sigma|F_i|^2$. Since the effective potential, at every order, is at least quadratic in the F fields, the massless scalars appear only in terms that are at least quartic, and do not get masses.

So far we have been discussing theories with semisimple gauge groups. We must now finally discuss how this picture changes if the gauge group is not semisimple. Assume that the gauge group contains a U(1) factor. Let D be the auxiliary field related by supersymmetry to the U(1) gauge boson. Then D itself is gauge invariant and $\int d^4xD(x)$ is supersymmetric. It can be written in the form (22), although with an integrand that has only restricted (Wess-Zumino) gauge invariance. Being of the form (22) this "D term" can be generated by loops, even if it is absent at the tree level. In fact, having dimension two it can be generated with a coefficient that is quadratically divergent. A simple calculation shows that if the U(1) symmetry is not spontaneously broken at the tree level, then the D term is generated with coefficient

$$\sum_i e_i \int d^4k \frac{1}{k^2}, \tag{25}$$

where the sum runs over all massless left-handed Fermi fields, e_i being the U(1) quantum number of the ith such field. In any theory with a U(1) factor in the gauge group, the D term will be generated in loops unless it is forbidden by a discrete parity symmetry (D is pseudoscalar).

The D term violates many of the statements made above and thereby restores naturalness. The D term can break supersymmetry, can give mass to massless particles, and can shift vacuum expectation values of scalars. Because the D term is subject to renormalization, however, no particular value of its coefficient is natural. If a small change in the coefficient can trigger supersymmetry breaking, the unbroken supersymmetry was not natural. Simply generating a D term by loops is not the way to get a natural gauge hierarchy.

What about theories that we may be most interested in, in which the gauge group is semisimple (perhaps simple) but is spontaneously broken to a non-semisimple group? For instance, suppose SU(5) is spontaneously broken to $SU(3)\times SU(2)\times U(1)$ at the tree level, but without breaking supersymmetry. One might then expect that the D term of the unbroken U(1) subgroup could be generated in loops, restoring naturalness (at least in some respects) and perhaps spontaneously breaking supersymmetry. This does not occur. To any finite order of perturbation theory we can work with an SU(5) invariant effective potential. Remarkably, there does not exist a supersymmetric and SU(5) invariant operator that after supersymmetry breaking reduces to the ordinary D term. In fact we have already seen that supersymmetric, SU(5) invariant operators that can be generated in loops and are linear in the D field associated with an unbroken symmetry always contain charged fields with zero vacuum expectation value. They cannot reduce after breaking of SU(5) (but not supersymmetry) to the simple D term.

If SU(5) (or some other group) is spontaneously broken down to $SU(3)\times SU(2)\times U(1)$, then it would seem natural, from the point of view of the low energy physics, for the D term to be generated. The masslessness of fields which would get mass if the D term appeared, and also unbroken supersymmetry, if this would be spoiled by appearance of D, would not seem natural from the point of view of the low energy theory. Nonetheless, the mere fact that at some energy $SU(3)\times SU(2)\times U(1)$ have been unified in SU(5) is enough to ensure that the D term is not generated in any finite order. It simply does not have a supersymmetric SU(5) invariant generalization. An illustration of this is the one-loop expression of eq. (25) which clearly vanishes if U(1) is part of a semisimple group.

This is possibly a unique case in physics in which an operator allowed by the unbroken symmetry group of a theory is prevented from appearing by a *broken* symmetry that is relevant only at higher energies. It sounds, superficially, like precisely "what the doctor ordered" in order to solve the gauge hierarchy problem. Nevertheless, in trying to apply this idea we will run into difficulties.

6. Some models

The previous sections have aimed at convincing the reader that dynamical breaking of supersymmetry is plausible. In this section, two simple models will be presented in which dynamical supersymmetry breaking really does occur. The models involve systems of less than four dimensions. As will become clear in the following sections, I do not know whether a workable mechanism for dynamical supersymmetry breaking exists in four dimensions.

The first model is not a field theory model at all but a model in potential theory – supersymmetric quantum mechanics. A supersymmetric quantum mechanical system is one in which there are operators Q_i that commute with the hamiltonian,

$$[Q_i, H] = 0, \qquad i = 1\cdots N, \tag{26}$$

and satisfy the algebra

$$\{Q_i, Q_j\} = \delta_{ij} H. \tag{27}$$

The simplest such system has $N = 2$ and involves a spin one half particle moving on the line. The wave function is therefore a two-component Pauli spinor,

$$\psi(x)=\begin{pmatrix}\phi_1(x)\\ \phi_2(x)\end{pmatrix}.$$

The Q_i are defined as

$$Q_1=\tfrac{1}{2}(\sigma_1 p+\sigma_2 W(x)),$$

$$Q_2=\tfrac{1}{2}(\sigma_2 p-\sigma_1 W(x)), \tag{28}$$

where the σ_i are the usual Pauli spin matrices, where as usual $p=-i\hbar\mathrm{d}/\mathrm{d}x$, and where W is an arbitrary function of x. It is straightforward to check that the algebra (26) and (27) is satisfied with

$$H=\tfrac{1}{2}\left(p^2+W^2(x)+\hbar\sigma_3\frac{\mathrm{d}W}{\mathrm{d}x}\right). \tag{29}$$

We will assume $|W|\to\infty$ as $|x|\to\infty$ so that the spectrum of the hamiltonian is discrete.

Now let us ask under what conditions supersymmetry is spontaneously broken. In the weak coupling (small $\hbar$) limit, one ignores the zero-point motion, and one ignores the last term, which is explicitly proportional to $\hbar$ (and which is a potential theory analogue of what in field theory would be a Yukawa coupling). At the tree level, the ground-state energy is therefore simply the minimum of W^2. The number of supersymmetrically invariant, zero-energy states is therefore, at the tree level, equal to the number of solutions of the equation $W(x)=0$.

Several interesting choices of W can be considered. In fig. 7 several choices of $W(x)$ and of the corresponding potential energy $V(x)=W^2(x)$ have been plotted. In fig. 7a W has a single zero, so the potential has a single supersymmetric minimum of zero energy. In fig. 7b, W has two zeros, so there are at the tree level two supersymmetrically invariant states. In fig. 7c, W has no zeros, the minimum of the potential is at non-zero energy, and supersymmetry is spontaneously broken at the tree level.

Before considering the exact spectrum, it is interesting to discuss the $\mathrm{O}(\hbar)$ corrections to the vacuum energy, in the case that this energy vanishes classically. Suppose that W has a simple zero at $x=x_0$, so $W(x)=\lambda(x-x_0)+\mathrm{O}((x-x_0)^2)$ for some λ. The hamiltonian is then $H=\frac{1}{2}(p^2+\lambda^2(x-x_0)^2+\hbar\lambda\sigma_3)$ plus higher order terms in $(x-x_0)$ that will not affect the $\mathrm{O}(\hbar)$ terms in the ground-state energy. The first two terms are a harmonic oscillator hamiltonian with zero-point energy $\frac{1}{2}\hbar|\lambda|$. The last term, $\frac{1}{2}\hbar\lambda\sigma_3$, has eigenvalues $\pm\frac{1}{2}\hbar|\lambda|$. Choosing the eigenvalue of σ_3 to minimize the energy, the ground-state energy is $\frac{1}{2}\hbar|\lambda|-\frac{1}{2}\hbar|\lambda|=0$.

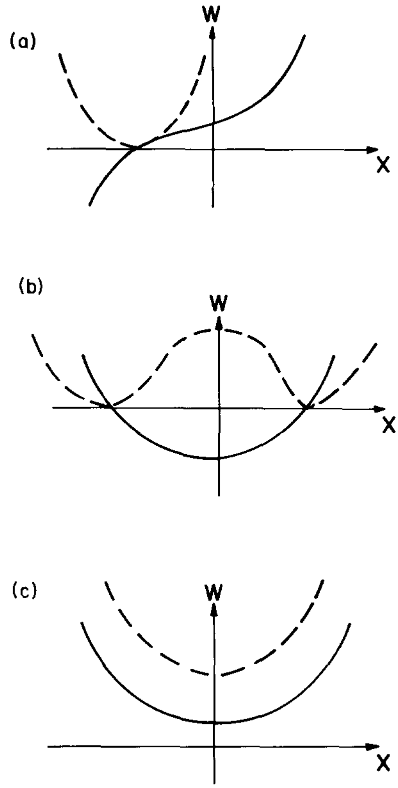

Fig. 7. Some choices of W (solid line) and W^2 (dashed line). In (a) W has a single zero and there is a single supersymmetric state in perturbation theory. In (b) W has two zeros and there are two such states. In (c) W has no zeros and supersymmetry is spontaneously broken at the tree level.

This is an example of the famous cancellation between the Bose and Fermi contributions to the ground-state energy of supersymmetric theories.

Rather than considering higher order terms in perturbation theory, let us now consider the exact spectrum of the theory. We will find that while at the tree level the number of supersymmetric states equals the number of zeros of W, in the exact spectrum the number of supersymmetric states is equal to one if W has an odd number of zeros, and is equal to zero if W has an even number of zeros. Therefore, in the theory of fig. 7a the exact ground state is supersymmetric, and in fig. 7c supersymmetry is really spontaneously broken. But in fig. 7b non-perturbative effects dynamically break the supersymmetry, which is unbroken at the tree level.

To establish these results, we will simply study the equation $Q_i\psi = 0$ which must be satisfied by a supersymmetric state. Actually, because of the general relation $Q_1^2 = Q_2^2 = \frac{1}{2}H$ (or the fact that in this particular model $Q_2 = -i\sigma_3 Q_1$, as can easily be checked), it is enough to satisfy $Q_1\psi = 0$, or $\sigma_1 p\psi = -\sigma_2 W\psi$. Multiplying by σ_1

and using the facts that $p = -i\hbar\,\mathrm{d}/\mathrm{d}x$, $\sigma_1\sigma_2 = i\sigma_3$, this equation becomes

$$\frac{\mathrm{d}\psi}{\mathrm{d}x} = \frac{1}{\hbar} W(x)\sigma_3\psi(x), \tag{30}$$

and the solution is

$$\psi(x) = \left(\exp \int_0^x \mathrm{d}y \frac{W(y)}{\hbar}\sigma_3\right)\psi(0). \tag{31}$$

This defines a supersymmetric state provided that it is a state – provided that ψ is normalizable. We must recall that in non-relativistic quantum mechanics, the quantization of energy levels and eigenvalues holds only because wave functions must be normalizable.

We have assumed that $|W(x)| \to \infty$ as $|x| \to \infty$. If W has an odd number of zeros, then the sign of W for $x \to +\infty$ is opposite from the sign of W for $x \to -\infty$. In this case there is a unique choice of $\psi(0)$ that makes (31) normalizable. For instance, if $W(x) \sim x^3$ for large x, then $\int_0^x \mathrm{d}y W(y) \sim \frac{1}{4}x^4$, and if $\sigma_3\psi(0) = -\psi(0)$, then (31) is normalizable. Since in fig. 7a W has precisely one zero, in this theory there is a supersymmetric ground state, given by (31) with $\sigma_3\psi(0) = -\psi(0)$.

On the other hand, if W has an even number of zeros, then the sign of W is the same for $x \to +\infty$ as for $x \to -\infty$. In this case no choice of $\psi(0)$ makes (31) normalizable. For instance, if $W(x) \sim x^4$ for large x, then $\int_0^x \mathrm{d}y W(y) \sim \frac{1}{5}x^5$. In this case, if $\sigma_3\psi(0) = -\psi(0)$, then (31) behaves badly for large negative x, and if $\sigma_3\psi(0) = +\psi(0)$, then (31) behaves badly for large positive x. Thus we find that both in figs. 7b and c, where W has an even number of zeros (two or zero, respectively), the exact spectrum does not contain a state invariant under supersymmetry. In the case of fig. 7c, where supersymmetry is spontaneously broken even at the tree level, this is no surprise. But in the case of fig. 7b, we have learned that the exact spectrum does not contain a state invariant under supersymmetry even though two such states exist at the tree level. Thus, quantum corrections dynamically break supersymmetry in this model.

We have not yet established explicitly that it is *non-perturbative* effects (rather than effects of some finite order in perturbation theory) that break supersymmetry in the model of fig. 7b, but this can easily be seen. In finite orders of perturbation theory one expands around a zero of W (if one exists). We have seen that supersymmetry is always unbroken if W possesses precisely one zero. Perturbation theory, in any finite order, is not sensitive to the existence or non-existence of a second zero of W. The vacuum energy, therefore, must vanish to all finite orders if it vanishes at the tree level.

Only a calculation which reveals the existence of more than one zero of W can possibly reveal the dynamical breaking of supersymmetry, in the model of fig. 7b. Such a calculation would be an instanton calculation – a calculation in which one evaluates the tunneling from one zero of W (one minimum of the potential) to the other. An instanton calculation in the model of fig. 7b does, in fact, demonstrate the dynamical symmetry breaking. This calculation will not be considered here in detail, because this would lead to the sort of considerations that we will consider anyway in the model discussed next.

Our next and last model of dynamically broken supersymmetry is a field theory in $2+1$ dimensions. It will be, in fact, a minimal supersymmetric generalization of one of the original instanton models considered by Polyakov [17].

This will be a model with a SU(2) gauge symmetry spontaneously broken down to U(1) at the tree level. Supersymmetry will be unbroken at the tree level but will be spontaneously broken due to non-perturbative effects, instantons.

The Bose fields of this model will be the gauge field $\boldsymbol{A}_\mu$ and a Higgs triplet $\boldsymbol{\phi}$; both are isovectors. The supersymmetry partners of the bose fields are Majorana Fermi fields $\boldsymbol{\psi}_A$ and $\boldsymbol{\chi}_A$, respectively. A suitable supersymmetric lagrangian involves, in addition to the conventional gauge-invariant kinetic energy, the following Yukawa couplings and scalar self-couplings:

$$L_{\text{int}} = -\tfrac{1}{2}\lambda^2\phi^2(\phi^2-a^2)^2 + e\varepsilon_{ijk}\bar{\psi}^i\chi^j\phi^k$$

$$-\lambda\bar{\chi}\chi(\phi^2-a^2) - 2\lambda(\bar{\chi}\cdot\phi)(\phi\cdot\chi), \tag{32}$$

where e is the gauge coupling and λ is a Yukawa coupling parameter. (A simple quartic potential for the scalars is forbidden by supersymmetry unless additional fields are introduced. The sixth-order potential considered above is renormalizable in three dimensions.)

The supersymmetry transformation laws can be written down most easily if one bears in mind that in three dimensions a vector V_i is equivalent to a second-rank symmetric spinor V_{AB} $(=V_{BA})$. The connection between V_{AB} and V_i is as follows: $V_{12}=V_z$, $V_{11}=\sqrt{\tfrac{1}{2}}(V_x+iV_y)$, $V_{22}=\sqrt{\tfrac{1}{2}}(V_x-iV_y)$. Since the gauge field $\boldsymbol{A}_i$, the field strength $\boldsymbol{B}_i$ and the covariant derivative of the Higgs field $\boldsymbol{D}_i\boldsymbol{\phi}$ are all vectors, they can be represented as symmetric spinors $\boldsymbol{A}_{AB}$, $\boldsymbol{B}_{AB}$, and $\boldsymbol{D}_{AB}\phi$. With this understanding, the supersymmetry transformation laws are

$$\delta\boldsymbol{A}_{AB} = \tfrac{1}{2}(\boldsymbol{\psi}_A\varepsilon_B + \boldsymbol{\psi}_B\varepsilon_A),$$

$$\delta\boldsymbol{\psi}_A = \boldsymbol{B}_{AB}\varepsilon^B,$$

$$\delta\boldsymbol{\phi} = \boldsymbol{\chi}_A\varepsilon^A,$$

$$\delta\chi_A = D_{AB}\phi\varepsilon^B + \varepsilon_A\lambda\phi(\tfrac{1}{2}\phi^2-a^2), \tag{33}$$

where ε_A is a parameter; indices are raised or lowered with the antisymmetric spinor ε^{AB}.

Looking now at the potential $\frac{1}{2}\lambda^2\phi^2(\phi^2 - a^2)$ of this theory, we note that there are two possible choices for the ground state: $\phi = 0$ and $|\phi| = a$. The first choice leads to a strongly coupled theory with an SU(2) gauge group. Dynamical supersymmetry breaking may or may not occur in this theory; tools for investigating this do not exist. We will instead analyze the vacuum state with $|\phi| = a$ at the tree level. We will see that in this vacuum, dynamical supersymmetry breaking occurs.

The $|\phi| = a$ vacuum has SU(2) broken down to U(1). The only massless boson is the U(1) gauge boson. The only massless fermion is the partner of the U(1) gauge boson. It can be represented in a gauge invariant way as $\phi \cdot \psi_A$. If supersymmetry is to be spontaneously broken, this massless fermion must become a Goldstone fermion; we must find a non-zero coupling of the massless fermion to the supersymmetry current. Actually, it follows from current algebra that $\phi \cdot \psi_A$ creates a Goldstone fermion if and only if the operator $\varepsilon^{AB}\{Q_A, \phi \cdot \psi_B\} = \{Q_A, \phi \cdot \psi^A\}$ has a non-zero vacuum expectation value. (Here ε^{AB} is the invariant, anti-symmetric second-rank spinor.) From (34) it is easy to see that $\{Q_A, \phi \cdot \psi^A\} = \chi_A \cdot \psi^A$ (note that $\{Q_A, \psi^A\} = 0$ because B_{AB} is symmetric and $\varepsilon^{AB}B_{AB} = 0$). We will demonstrate supersymmetry breaking by showing that the order parameter $\chi_A \cdot \psi^A$ has a vacuum expectation value.

After spontaneous breaking of SU(2) to U(1), this model has instantons, as in the model (without fermions) originally discussed by Polyakov. The instantons are just 't Hooft-Polyakov magnetic monopoles in one dimension less. They have a non-zero value of the topological charge $(e/2\pi a)\int d^3x\, \partial_i(\boldsymbol{B}_i \cdot \phi)$, which in $3+1$ dimensions would be regarded as magnetic charge but in $2+1$ dimensions is the instanton number.

In the supersymmetric theory considered here, the instanton contribution to the vacuum to vacuum transition amplitude vanishes. This is because there exist two fermion zero modes*. These modes are unrelated to chiral symmetries but can be obtained by acting with a supersymmetry transformation on the classical solution. The zero modes can be read off from the transformation laws (33), and have the following form:

$$\psi_A = B_{AB}^{\text{cl}}\varepsilon^B,$$

$$\chi_A = D_{AB}\phi^{\text{cl}}\varepsilon^B + \varepsilon_A\lambda\phi^{\text{cl}}\left(\tfrac{1}{2}\phi^{\text{cl}^2} - a^2\right). \tag{34}$$

Here $\boldsymbol{B}_{AB}^{\text{cl}}$ and ϕ^{cl} are the classical instanton solution, and ε_B is a constant.

Although the one-instanton contribution to the ground-state energy vanishes because of the zero modes (34), the instanton gives an expectation value to the supersymmetry-breaking order parameter $\chi_A \cdot \psi^A$, which can be easily calculated. In general, one should expand the χ and ψ fields as a sum of zero and non-zero modes

* For related discussions, see Jackiw and Rebbi [18].

of the Dirac operator. However, a non-zero contribution to the expectation value of $\boldsymbol{\chi}_A \cdot \boldsymbol{\psi}^A$ arises only from the zero-mode terms in χ and ψ; contributions in which χ and ψ do not "absorb" the two zero modes vanish. So we replace $\boldsymbol{\chi}\cdot\boldsymbol{\psi}(x)$ by

$$\sum_{\text{zero modes}} \boldsymbol{\chi}_A(x)^{(\text{zero mode})} \cdot \boldsymbol{\psi}^A(x)^{(\text{zero mode})} = \boldsymbol{B}^{\text{cl}}_{AB}(x) \cdot \boldsymbol{D}^{\text{cl}}_{AB} \phi^{\text{cl}}(x), \tag{35}$$

where eq. (34) has been used to evaluate the right-hand side. After integrating over the position x, which is equivalent to integrating over the location of the instanton, we learn that one can forget about the existence of zero modes if one replaces $\boldsymbol{\chi}\cdot\boldsymbol{\psi}$ by $\int \mathrm{d}^3 x\, \boldsymbol{B}^{\text{cl}}_{AB}(x) \cdot D^{\text{cl}}_{AB} \boldsymbol{\phi}(x)$. Returning to a vector notation and using the Bianchi identity $D_i B_i = 0$, one finds that $\boldsymbol{\chi}\cdot\boldsymbol{\psi}$ should be replaced by $\int \mathrm{d}^3 x\, \partial_i(\boldsymbol{B}_i \cdot \boldsymbol{\phi})$, which is precisely $2\pi a/e$ times the topological charge.

Since this quantity is certainly non-zero both for instantons and anti-instantons, dynamical supersymmetry breaking is within reach. However, instantons and anti-instantons have opposite topological charge so a cancellation may occur. The reason for the cancellation is that $\boldsymbol{\chi}\cdot\boldsymbol{\psi}$ is pseudoscalar and cannot get a vacuum expectation value if the discrete space-time symmetries are conserved. (Also, $\boldsymbol{\chi}\cdot\boldsymbol{\psi}$ is odd under the transformation $\boldsymbol{\phi} \to -\boldsymbol{\phi}$, $\boldsymbol{\chi} \to -\boldsymbol{\chi}$, which, in conjunction with a certain discrete gauge transformation, is a symmetry of the theory.) The cancellation between instantons and anti-instantons can be avoided if the lagrangian contains a term that violates the discrete symmetries.

One might expect that one could stop the cancellation between instantons and anti-instantons by including in the action the topological charge density

$$\Delta L = \frac{\theta e}{4\pi^2 a} \partial_i(\phi \cdot B_i), \tag{36}$$

analogous to $\theta F\tilde{F}$ in four dimensions. This respects supersymmetry because being a total divergence it does not change the equations of motion or the conservation of the supercurrent. This addition to the action would appear to weight instantons and anti-instantons by factors $\mathrm{e}^{\pm i\theta}$, respectively, and so to break the cancellation between them. There are, however, some strong arguments that physical amplitudes are not really θ dependent in $2+1$ dimensions, because the topological charge density is the divergence of a *gauge-invariant* operator.

If this is so, one may still expect to remove the cancellation between instantons and anti-instantons, and so obtain dynamical supersymmetry breaking, by including in the lagrangian other operators that violate the discrete symmetries. An operator with the right quantum numbers is

$$\Delta L = g\boldsymbol{\phi}^2 \partial_i(\boldsymbol{\phi}\cdot\boldsymbol{B}_i), \tag{37}$$

which is not a total divergence and has a supersymmetric generalization. In the classical instanton field, (37) has an expectation value opposite in sign to its value in the anti-instanton field, so the cancellation is lifted. [Although (37) is unrenormalizable, this is presumably only a technicality. Eq. (37) could be obtained as a low energy effective lagrangian in renormalizable theories with additional massive fields that have been integrated out.] Thus, the cancellation between instantons and anti-instantons is not universal, and dynamical supersymmetry breaking can occur in models of this class.

These models have some interesting generalizations, which will not be explored here in detail. The potential theory model can be generalized to systems with an arbitrary finite number of degrees of freedom. The phenomenon just noted that supersymmetry is spontaneously broken only at non-zero θ occurs also in some potential theory models and in some two dimensional gauge theories; it may be very widespread. The most important generalization of the above models would be, of course, their extension to four dimensions.

7. Return to four dimensions

Let us now return to the more or less "realistic" grand unified theories considered in sect. 4 – the theories in which the gauge group is spontaneously broken to $SU(3) \times SU(2) \times U(1)$ at the tree level, without spontaneous breaking of supersymmetry. From sect. 5 (and the previous literature on supersymmetry on which most of sect. 5 is based) we know that in these theories supersymmetry will remain unbroken to all finite orders of perturbation theory. Can these theories develop a natural gauge hierarchy, via a non-perturbative mechanism that would spontaneously break supersymmetry and $SU(2) \times U(1)$ at very low energies?

A non-perturbative mechanism might involve gauge couplings which become strong at low energies as a result of a negative β function. This leads to the question of supersymmetric technicolor, which will be considered in sect. 9. However, in this section and the next one, I wish to consider the possibility that dynamical supersymmetry breaking can occur as a *small* effect due to *weak* couplings, as occurred in the models of sect. 6. Can dynamical sypersymmetry breaking occur in four dimensions when all couplings are weak?

This seems plausible because, as has been discussed, unbroken supersymmetry is linked to the exact vanishing of the vacuum energy. Perhaps supersymmetry can be spontaneously broken by extremely weak quantum corrections, which would give a small but non-zero value to the vacuum energy. Perhaps this occurs in the "realistic" grand unified models of sect. 4.

Even without considering detailed mechanisms, there are two necessary conditions that must be satisfied, if this is to work. One of these has been discussed in the preceding sections, and one has not been discussed explicitly.

First of all, if supersymmetry is to be spontaneously broken, there will have to be a Goldstone fermion, that is, a massless neutral fermion created from the vacuum by the supersymmetry current,

$$\langle 0|S_{\mu\alpha}|\psi_\beta\rangle = f\gamma_{\mu\alpha\beta}, \tag{38}$$

with $f \neq 0$. Since weak quantum corrections will not manufacture a massless fermion as a bound state, dynamical supersymmetry breaking could be induced by weak corrections only in theories in which massless neutral fermions already exist at the tree level. However, as has been mentioned above, all theories of realistic phenomenological interest have massless neutral fermions at the tree level. For instance, the supersymmetric partners of the photon and Z mesons are massless as long as supersymmetry and $SU(2) \times U(1)$ are unbroken. So this constraint is satisfied automatically.

Second, if we are to solve the gauge hierarchy problem via weak corrections that dynamically break supersymmetry, we need a theory in which the low energy physics, in perturbation theory, is unstable, in a sense that will now be described. At the tree level, some particles get masses, and others remain massless. (The massless particles include the gauge mesons, the quarks and leptons, and their supersymmetric partners.) The effective lagrangian for the massless particles is, at the tree level, something of the general form

$$L^{\mathrm{eff}} = \sum C_i O_i^{(4)}, \tag{39}$$

where the C_i are numerical coefficients and the $O_i^{(4)}$ are supersymmetric operators of dimension four, invariant under all unbroken gauge symmetries.

Notice that L^{eff} contains operators of dimension four only. Operators of dimension less than four would give masses to some particles, but L^{eff} is by definition the low energy effective lagrangian for the massless particles only.

Now, a non-perturbative effect which is supposed to trigger supersymmetry breaking can always be interpreted, at low energies, as a correction to the effective lagrangian (39). Even if supersymmetry is going to be spontaneously broken, we can always work with an effective lagrangian that is invariant under supersymmetry (although its minimum is not invariant). Hence, supersymmetry can be spontaneously broken by quantum corrections only if there are supersymmetric, gauge-invariant operators whose addition to (39) would induce supersymmetry breaking.

Actually, we can be far more specific. Non-perturbative effects will modify (39) by terms with coefficients smaller than any power of α, let us say of order $\exp(-1/\alpha)$ to be definite. Effects of order $\exp(-1/\alpha)$ will be far too small to be of realistic interest if they multiply operators of dimension four or more. Even if they triggered supersymmetry breaking, this would give far too *large* a hierarchy – too *large* a ratio of mass scales. Effects of order $\exp(-1/\alpha)$ are of interest only if they involve operators of dimension less than four and are hence *enhanced* by positive powers of the grand unified mass scale. For instance, if non-perturbative effects cause a dimension-two operator $O^{(2)}$ to appear in the effective lagrangian, this would contribute something like

$$\Delta L^{\mathrm{eff}} = M^2 \exp\left(-\frac{1}{\alpha}\right) O^{(2)} \tag{40}$$

to the effective lagrangian, where M is a mass that would be expected to be of order the unified mass scale. If $O^{(2)}$ triggers $\mathrm{SU}(2)\times\mathrm{U}(1)$ breaking and supersymmetry breaking, (40) could be just right to solve the hierarchy problem.

The net effect of this is the following. For dynamical supersymmetry breaking to be possible as a solution of the hierarchy problem, there must exist in the theory an operator of dimension less than four which respects all relevant symmetries (supersymmetry and the unbroken gauge symmetries) yet is not present in the low energy effective lagrangian in perturbation theory. (If it appeared in finite orders of perturbation theory, it would generate excessively large masses and too weak a gauge hierarchy.) If the operator is generated non-perturbatively in the effective lagrangian, we may solve the hierarchy problem.

It sounds like a very tall order to ask that in a theory there should exist operators that are invariant under all relevant symmetries, yet are not generated in perturbation theory. But this criterion is automatically satisfied in our "realistic" grand unified theories! We have seen that there is always at least one operator, the D term of the U(1) subgroup of $\mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1)$, which satisfies all relevant symmetries of the low energy theory but is not generated in perturbation theory. It has dimension two. In addition, depending on details of the model, there may be low dimension operators of type (23) which are absent at the tree level for one reason or another but which satisfy all the low energy symmetries. These operators, too, will not be generated in perturbation theory but might be generated non-perturbatively.

Our more or less "realistic" grand unified theories, with unbroken supersymmetry at the tree level and the gauge group spontaneously broken to $\mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1)$, therefore satisfy all the preliminary conditions that must be satisfied in order that non-perturbative effects might be able to break supersymmetry. It is therefore appropriate to look for particular mechanisms.

Actually, with the present state of knowledge in theoretical physics, there is in the weak coupling domain only one known non-perturbative mechanism that could be relevant, instantons. We are therefore led to ask whether instantons could induce dynamical supersymmetry breaking, as in the models of the previous section. (See also the discussion in ref. [18].)

Suppose that SU(5) is broken to $\mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1)$ at the tree level. The embedding of $\mathrm{SU}(3)\times\mathrm{SU}(2)$ in SU(5) can be visualized by breaking a 5×5 matrix up in blocks:

$$\left(\begin{array}{c|c} 3\times 3 & 3\times 2 \\ \hline 2\times 3 & 2\times 2 \end{array}\right). \tag{41}$$

One may consider instantons lying in one of the unbroken subgroups. Instantons of the Weinberg-Salam SU(2) might dynamically break supersymmetry, but their effects would be unreasonably small, as in 't Hooft's calculation of baryon number violation in the SU(2) × U(1) model [19]. Instantons of color SU(3) might spontaneously break supersymmetry. The mass scale in this case would be a few hundred MeV, which is the scale at which the QCD coupling becomes strong. If instantons of strongly coupled theories induce supersymmetry breaking, this would be of interest in connection with technicolor theories. However, we will postpone the discussion of theories with strong gauge coupling to sect. 9.

Here we wish to consider the possibility that supersymmetry is spontaneously broken by instantons that do *not* lie in the unbroken gauge group. For instance, in the theory of SU(5) broken to SU(3) × SU(2) × U(1), one may consider an instanton in an SU(2) subgroup that lies partly in color SU(3) and partly in weak interaction SU(2). The instanton may lie in the SU(2) subgroup of SU(5) indicated in parenthesis in (42). Instantons of such a subgroup are characterized by a natural

$$\left(\begin{array}{c|c} \begin{pmatrix} X & X \\ \hline X & X \end{pmatrix} \end{array}\right) \tag{42}$$

mass scale, the mass M_X of the heavy bosons of the theory. They are also characterized by a small coupling constant, the coupling α_{GUT} at energies of grand unification. Their effects are therefore explicitly calculable, without infrared divergences or ambiguities.

The action of one of these instantons is $2\pi/\alpha_{GUT}$. Consequently, if they do cause one of the appropriate dimension two operators to appear in the effective lagrangian, then it will appear, as in eq. (40), with a coefficient of order

$$m^2 = M_X^2 \left(\frac{2\pi}{\alpha_{GUT}} \right)^k \exp(-2\pi/\alpha_{GUT}), \tag{43}$$

where k is the number of collective coordinates. Assuming M_X to be in the range of 10^{15} GeV to 10^{19} GeV, this gives a reasonable low energy mass scale, $m \sim 10^3$ GeV, if α_{GUT} is about $\frac{1}{20}$ or $\frac{1}{25}$. This is larger than the conventional value by about a factor of 2. But because supersymmetric theories will contain extra fields not included in the conventional calculation, we should expect that α_{GUT} would be larger. Therefore, eq. (43) seems quantitatively reasonable.

Unfortunately, instanton-induced supersymmetry breaking either does not work in four dimensions or at least does not work as straightforwardly as it works in three dimensions. A simple way to explain this is the following. To obtain supersymmetry breaking we need a non-zero matrix element of the form

$$\langle 0 | S_{\mu\alpha} | \psi_\beta \rangle \,. \tag{44}$$

[In sect. 6 we calculated, instead, an order parameter that is related to (44) by current algebra.] In an instanton field there will always be fermion zero modes, obtained by acting on the classical solution with a supersymmetry transformation. In three dimensions a Majorana fermion has two components. There are therefore two zero modes. Two zero modes is the proper number to give a non-zero matrix element (44). One zero mode is absorbed by the current $S_{\mu\alpha}$. The second couples to the fermion ψ_{β}. However, in four dimensions, a Majorana fermion has four components, so instead of two zero modes one has four zero modes. With four zero modes the instanton gives non-zero values to four fermion matrix elements but not to (44). (Although a classical instanton solution does not quite exist in the cases of interest here, one can count zero modes as if a solution did exist. For a discussion of theories with no classical instanton solution, see recent work by Affleck [20].)

One should not necessarily accept the above reasoning at face value. For instance, why couldn't $S_{\mu\alpha}$ couple to a three-fermion state (and absorb three zero modes)? However, detailed analysis appears to show that, at least in reasonably simple models, $S_{\mu\alpha}$ does not couple to the relevant three fermion state, and the simple argument above – too many zero modes for supersymmetry breaking – is correct. This is disappointing, because on the face of things eq. (43) is an appealing answer to the hierarchy problem.

To have precisely two zero modes in an instanton field would be possible if two of the four supersymmetry generators annihilated the instanton, rather than creating zero modes. This does not occur for instantons that lie in a spontaneously broken part of the gauge group. Instantons of an unbroken gauge symmetry are annihilated by two of the four sypersymmetry charges, but we will postpone discussing dynamical effects of unbroken non-abelian gauge symmetries until sect. 9.

The theoretical situation regarding the "realistic" grand unified models is very unsatisfactory. These theories appear to have the potential for arbitrarily small effects to break supersymmetry. If this does not occur, an argument independent of perturbation theory and independent of detailed instanton analysis should be found. Until it is found, one must suspect that these theories may develop huge gauge hierarchies.

8. Gravitational instantons?

The instanton mechanism considered in the last section would be rather appealing, if it worked, not only because eq. (43) for the mass scale seems reasonable, but also because this mechanism would provide a perhaps unique opportunity to observe non-perturbative effects of grand unification. In eq. (43), the supersymmetry breaking effects are *enhanced*, rather than suppressed, by powers of the enormous mass M_X.

It is intriguing to ask whether, instead, supersymmetry breaking could result from non-perturbative effects of another force usually regarded as weak: gravitation.

Could supersymmetry breaking result from gravitational instantons? Could the mass scale of supersymmetry breaking be of order $M_{\mathrm{p}}^2 \exp(-1/\lambda)$, where M_{p} is the Planck mass and λ is a dimensionless coupling constant – as yet unknown – for gravitation? If λ is small, this could solve the hierarchy problem. If λ is large, this dynamical supersymmetry breaking would not be relevant to the hierarchy problem, but we wish to understand dynamical supersymmetry breaking if it occurs in nature even if it is not relevant to the hierarchy problem.

Let us first discuss in general terms what would be involved in supersymmetry breaking by gravitational instantons. By a gravitational instanton one means, in this context, an asymptotically euclidean solution of the positive signature Einstein equations, perhaps with suitable matter fields present. Actually, several proofs have been given that there are no asymptotically flat gravitational instantons in the absence of matter fields [21]. However, such solutions do exist if a second-rank antisymmetric tensor field is present [22], and probably in other theories. Also, it may not really be necessary to work with exact classical solutions.

In supergravity, there is a spin $\frac{3}{2}$ particle, the Rarita-Schwinger field $\psi_{\mu\alpha}$. It is related by supersymmetry to the graviton. Local supersymmetry is spontaneously broken if and only if the spin $\frac{3}{2}$ particle has non-zero mass. This can be seen as follows. A non-zero mass for the spin $\frac{3}{2}$ particle means that it is no longer degenerate with the graviton (which is presumably still massless), so supersymmetry is broken. On the other hand, S-matrix theory arguments [23] (which have been worked out only in special cases but are presumably valid in general) show that as long as the spin $\frac{3}{2}$ particle is massless, supersymmetry must be unbroken.

If instantons are to break supersymmetry, they must give a mass to the previously massless Rarita-Schwinger field. Instantons must, therefore, contribute to the two-point function $\langle 0|\psi_{\mu\alpha}\psi_{\nu\beta}|0\rangle$. In an instanton field, there will be fermion zero modes. To give a mass to a previously massless Rarita-Schwinger field, two zero modes is the requisite number. Thus, although the rationale is different, we reach the same conclusion as in the previous section. An instanton in the field of which there are precisely two zero modes could very plausibly induce spontaneous supersymmetry breaking.

For illustrative purposes, let us ask whether in pure supergravity theory (only the graviton and Rarita-Schwinger field) there may exist instantons with precisely two zero modes. This theory will be considered for the sake of discussion only. While dynamical supersymmetry breaking may or may not occur in pure supergravity theory, instantons cannot be the mechanism for it. The Rarita-Schwinger field, to get a mass, must have four helicity states instead of two. Instantons will not create the extra two helicity states "out of thin air"; they can induce dynamical supersymmetry breaking only in theories in which there is, at the tree level, a massless spin $\frac{1}{2}$ fermion which can supply the extra two helicity states to make the Rarita-Schwinger field massive. (This is the "super-Higgs" mechanism [24].) What is more, asymptotically euclidean instantons do not exist in pure supergravity, as has already been noted. Nonetheless, the reader is invited to temporarily suspend disbelief and ask whether a hypothetical instanton in pure supergravity could admit precisely two zero modes.

Zero modes will be obtained by making a supersymmetry transformation. For any spinor field $\varepsilon(x)$, the Rarita-Schwinger equations are satisfied by

$$\psi_\mu = D_\mu \varepsilon, \tag{45}$$

since under supersymmetry $\delta\psi_\mu = D_\mu \varepsilon$. For what choices of ε is (45) a "zero mode"?

If ε vanishes at infinity, then ψ_μ should be gauged away, in the fashion of Faddeev-Popov. If ε grows at infinity, ψ_μ is not normalizable and should not be included. However, if ε is asymptotic to a constant at infinity, then ε is not normalizable so (45) is not gauged away in the Faddeev-Popov construction, but ψ_μ is normalizable and is a valid zero mode. Only the asymptotic behavior of ε for large $|x|$ is relevant, because given two choices of ε with the same asymptotic behavior, their difference vanishes at infinity, so the difference between the two ψ fields can be gauged away.

In general, there will be four zero modes, since four linearly independent choices can be made for the asymptotic behavior of ε in (45). The number of zero modes can be less than four only if some of the ψ_μ in (45) vanish (or equivalently, can be gauged away with a normalizable gauge parameter; in this case $\psi_\mu = 0$ if ε is defined properly). Hence, the number of zero modes is less than four if and only if the equation $D_\mu \varepsilon = 0$ has solutions.

That equation is extremely restrictive because it implies the integrability condition $[D_\mu, D_\nu]\varepsilon = 0$ or

$$R_{\mu\nu\alpha\beta}\sigma^{\alpha\beta}\varepsilon = 0. \tag{46}$$

In general, except in flat space there are no solutions. Non-trivial solutions exist only if the Riemann tensor is self-dual or anti-self-dual,

$$R_{\mu\nu\alpha\beta} = \pm\tfrac{1}{2}R_{\mu\nu\gamma\delta}\varepsilon_{\alpha\beta}{}^{\gamma\delta}, \tag{47}$$

in which case there are two solutions, of definite chirality. [Since $\sigma^{\alpha\beta} = \frac{1}{2}\varepsilon^{\alpha\beta\gamma\delta}\sigma^{\gamma\delta}\gamma_5$, (46) follows from (47) if $\gamma_5\varepsilon = \mp\varepsilon$.]

For a self-dual gravitational instanton there are therefore exactly two non-vanishing zero modes in the form (45), corresponding to the two choices of the asymptotic behavior of ε for which $D_\mu\varepsilon = 0$ *cannot* be satisfied.

These two zero modes have the same chirality [positive or negative depending on the sign in (47)]. Two zero modes of the same chirality are just what we would need to give the Rarita-Schwinger field a mass and thus break supersymmetry.

However, while there are many solutions of (47), none of them is asymptotically euclidean [25]. While this can be proved in a variety of ways, the previous comments may be regarded as a "physics proof." An asymptotically euclidean solution of (47) would give the Rarita-Schwinger field a mass, which is impossible, since two helicity states are missing.

When additional fields are included, the formula $\delta\psi_\mu = D_\mu \varepsilon$ is replaced by $\delta\psi_\mu = \bar{D}_\mu \varepsilon$, where $\bar{D}_\mu$ is an operator that includes couplings to the other fields. The integrability condition (46) is replaced by $[\bar{D}_\mu, \bar{D}_\nu]\varepsilon = 0$. In some theories of gravity plus matter, there may be asymptotically euclidean instantons for which the integrability condition has solutions. In this case, dynamical supersymmetry breaking would be likely to follow. For instance, theories with torsion would seem to be logical candidates, since Regge and Hanson [26] have shown that in the presence of torsion, an asymptotically euclidean space may satisfy (46). Supersymmetry breaking would appear to be one of few areas where non-perturbative effects of gravity might plausibly be observable.

9. Supersymmetric technicolor?

In sects. 7 and 8, we have considered the issues that arise if one tries to attribute dynamical supersymmetry breaking to instantons. There is one other place in the current body of knowledge about high energy physics where one could hope to find a mechanism for dynamical supersymmetry breaking. In gauge theories with negative beta function, the gauge coupling becomes strong at low energies. Various non-perturbative effects may occur, including confinement, binding of color singlet hadrons, and dynamical mass generation. In the case of QCD, one of the bound states that forms is the pion – a Goldstone boson for spontaneously broken chiral symmetry. In a supersymmetric theory with a strong gauge coupling, could not the bound states include a massless fermion – a Goldstone fermion of broken supersymmetry?

In the last few years it has been suggested [27] that in addition to the $SU(3) \times SU(2) \times U(1)$ gauge forces, there may exist even stronger gauge forces – "hypercolor" or "technicolor". These forces are suggested to become strong at energies of hundreds of GeV, and to spontaneously break the weak $SU(2) \times U(1)$.

In the absence of elementary scalars, this sort of theory can lead to a natural resolution of the gauge hierarchy problem. The principal difficulty is that because there are no elementary scalars, there are unwanted chiral symmetries which prevent fermion masses. The inclusion of scalars light enough to matter would ruin the solution of the problem of naturalness. The unwanted chiral symmetries can be explicitly broken by means of additional gauge interactions, known as "extended technicolor" [28], but this leads to further difficulties.

These problems could conceivably be solved by a supersymmetric generalization of the usual hypercolor or technicolor scenario. Suppose that supersymmetry is unbroken down to the technicolor mass scale. Suppose that strong technicolor forces bind a Goldstone fermion as well as binding various Goldstone bosons. Then the strong technicolor forces can dynamically break both supersymmetry and $SU(2) \times U(1)$.

On the other hand, because of supersymmetry, the theory will automatically contain elementary scalars with masses no greater than the technicolor mass scale. Consequently, the difficulties of conventional technicolor will not be present in supersymmetric technicolor theories. Yukawa couplings of the elementary scalars can explicitly break the unwanted chiral symmetries and give masses to quarks and leptons.

The most serious obstacle to a successful theory of this sort is probably the question of whether strong gauge forces really do break supersymmetry. On this score there is one weak but encouraging indication. In QCD, the vacuum expectation value of $\bar{\psi}\psi$ serves as an order parameter for spontaneously broken chiral symmetry. It is therefore plausible to suppose that fermion bilinears have vacuum expectation values in supersymmetric gauge theories. However, in supersymmetric gauge theories with chiral matter multiplets, one finds formulas of the sort $\bar{\psi}\psi = \{Q_\alpha, \bar{\psi}^\alpha\phi\}$ so that a non-zero vacuum expectation value of $\bar{\psi}\psi$ would spontaneously break supersymmetry as well as spontaneously breaking chiral symmetry.

Actually, we know that dynamical breaking of various global symmetries occurs in supersymmetric gauge theories, as long as confinement occurs. This follows from recent arguments based on triangle anomalies [29], which are also valid in the supersymmetric context. However, the fact that chiral symmetry is spontaneously broken in these theories does not guarantee that $\bar{\psi}\psi$ has a vacuum expectation value. In fact, an alternative order parameter for chiral symmetry breaking in these theories would be a scalar bilinear of the general type $\bar{\phi}\phi$ ($\bar{\phi}$ and ϕ are two scalar fields that transform differently under global symmetries; they are not complex conjugates of each other). A non-zero $\langle\bar{\phi}\phi\rangle$ does *not* mean that supersymmetry is spontaneously broken. It is perfectly possible that the breaking of global symmetries in these theories leads to non-zero $\langle\bar{\phi}\phi\rangle$, yet does not lead to non-zero $\langle\bar{\psi}\psi\rangle$ – precisely because of unbroken supersymmetry.

A convincing argument pro or con the question of supersymmetry breaking by strong gauge forces would be very welcome.

10. Conclusions

The potential advantages of the scenarios described in this paper should seem obvious. I wish, however, to draw the reader's attention to several difficulties (some of them mentioned previously) that will exist even if one mechanism or another for dynamically broken supersymmetry can be shown to work.

It may be difficult to account for the relative stability of the proton, since the supersymmetric partners of the quarks are color triplet scalars which may mediate proton decay.

One may encounter unwanted relations among quark and lepton masses, similar to those in the simplest SU(5) and O(10) models but perhaps even more severe.

It is hard to unify technicolor, with or without supersymmetry, with $SU(3) \times SU(2) \times U(1)$.

In global supersymmetry, the vacuum energy is positive definite (and equal to f^2) if supersymmetry is spontaneously broken. When gravity is included, this translates into an unacceptably large, positive cosmological constant. In supergravity, but not in global supersymmetry, the Bose-Fermi symmetry can be spontaneously broken while the vacuum energy remains zero. This suggests that, if supersymmetry is relevant to nature, it must be spontaneously broken under circumstances such that gravity cannot be ignored; only if gravity cannot be ignored can supersymmetry be broken without the generation of a positive vacuum energy. This may in turn suggest that supersymmetry should *not* survive to low energies, for if it is broken only at low energies one might expect gravity to be irrelevant. However, this line of reasoning leaves many alternatives open.

On a more optimistic note, let us recall that, since gravity does exist in nature, the "Goldstone fermion" that we have discussed at length is not a true, physical, massless spin $\frac{1}{2}$ particle in the world we live in. If supersymmetry exists and has been spontaneously broken, the Goldstone fermion has been absorbed to become the helicity $\pm\frac{1}{2}$ components of the massive spin $\frac{3}{2}$ Rarita-Schwinger particle [24]. Just as the mass of the W boson is equal, in theories of dynamical symmetry breaking, to eF_π, the mass of the Rarita-Schwinger particle will be f/M_{p}, where M_{p} is the Planck mass (replace F_π by f and e by $1/M_{\mathrm{p}}$). As f ranges from $(10^5\ \mathrm{GeV})^2$ to $(10^6\ \mathrm{GeV})^2$, which is a plausible range at least in the case of supersymmetric technicolor, the spin $\frac{3}{2}$ particle varies in mass from 1 eV to 100 eV. This is the required range to account for the dark matter that is observed to exist [30] in galactic halos and in clusters of galaxies, so it is just barely possible that most of the mass in the universe consists of massive Rarita-Schwinger particles of spin $\frac{3}{2}$!

In conclusion, let us note that while various questions have been addressed in this paper, the most important question has not been answered. Is dynamical supersymmetry breaking in four dimensions a myth or a reality?

I would like to acknowledge comments by K. Bardakci and D.J. Gross, and to thank B. Ovrut and J. Wess for discussions of their work.

Recent discussions of models of the type suggested in sect. 9 have been made by M. Dine, W. Fischler, and M. Srednicki (preprint, Princeton University) and by S. Dimopoulos and S. Raby (preprint, Stanford University).

References

[1] J. Wess and B. Zumino, Nucl. Phys. B70 (1974) 39;
D.V. Volkov and V.P. Akulov, Phys. Lett. 46B (1973) 109;
Y.A. Gol'fand and E.P. Likhtam, JETP Lett. 13 (1971) 323

[2] P. Fayet and J. Iliopoulos, Phys. Lett. 51B (1974) 461;
P. Fayet, Phys. Lett. 58B (1975) 67;
L. O'Raifeartaigh, Nucl. Phys. B96 (1975) 331
[3] P. Fayet, Nucl. Phys. B90 (1975) 104; Phys. Lett. 69B (1977) 489;
84B (1979) 421;
G.R. Farrar and P. Fayet, Phys. Lett. 76B (1978) 575; 79B (1978) 442; 89B (1980) 191
[4] D.Z. Freedman, P. van Nienwenhuizen and S. Ferrara, Phys. Rev. D13 (1976) 3214;
S. Deser and B. Zumino, Phys. Lett. 62B (1976) 335
[5] S. Weinberg, Phys. Lett. 62B (1976) 111;
E. Gildener and S. Weinberg, Phys. Rev. D13 (1976) 333
[6] J. Iliopoulos and B. Zumino, Nucl. Phys. B76 (1974) 310
[7] P. Fayet and S. Ferrara, Phys. Reports 32 (1977) 249
[8] A. Salam and J. Strathdee, Phys. Lett. 49B (1974) 465
[9] H. Georgi and A. Pais, Phys. Rev. D10 (1974) 1246
[10] S. Coleman and E. Weinberg, Phys. Rev. D7 (1973) 1888
[11] B. Zumino, Nucl. Phys. B89 (1975) 535;
P. West, Nucl. Phys. B106 (1976) 219;
D.M. Capper and M. Ramón Medrano, J. Phys. G2 (1976) 269;
W. Lang, Nucl. Phys. B114 (1976) 123;
S. Weinberg, Phys. Lett. 62B (1976) 111
[12] A. Salam and J. Strathdee, Nucl. Phys. B76 (1974) 477; B86 (1975) 142; Phys. Rev. D11 (1975) 1521;
J. Honerkamp, F. Krause, M. Scheunert and M. Schlindwein, Phys. Lett. 53B (1974) 60;
L. Mezencescu and V.I. Ogievetsky, Usp. Fiz. Nauk 117 (1975) 637;
S. Ferrara and O. Piguet, Nucl. Phys. B93 (1975) 261
[13] J. Scherk and J.H. Schwarz, Phys. Lett. 82B (1979) 60; Nucl. Phys. B153 (1979) 61;
E. Cremmer, J. Scherk and J.H. Schwarz, Phys. Lett. 84B (1979) 83
[14] E. Witten, Nucl. Phys. B186 (1981) 412
[15] P. Fayet and J. Iliopoulos, Phys. Lett. 51B (1974) 461
[16] M.T. Grisaru, W. Siegel and M. Rocek, Nucl. Phys. B159 (1979) 429;
R. Jackiw and C. Rebbi, Phys. Rev. D16 (1977) 1053
[17] A. Polyakov, Phys. Lett. 59B (1975) 82
[18] L.F. Abbott, M.T. Grisaru and H.J. Schnitzer, Phys. Rev. D16 (1977) 3002
[19] G. 't Hooft, Phys. Rev. Lett. 37 (1976) 8
[20] I. Affleck, On constrained instantons, Harvard preprint (1980)
[21] P. Schoen and S.T. Yau, Phys. Rev. Lett. 42 (1979) 547;
E. Witten, A simple proof of the positive energy theorem, Princeton preprint, (1980) Comm. Math. Phys., to be published
[22] E. Witten, unpublished
[23] M.T. Grisaru, H.N. Pendleton and P. van Nieuwenhuizen, Phys. Rev. D15 (1977) 996
[24] D.V. Volkov and V.A. Soroka, JETP Lett. 18 (1973) 22;
B. Zumino, Coral Gables Lectures (January, 1977), Lectures at the 1976 Scottish Universities Summer School (Edinburgh Univ. Press, Edinburgh, Scotland, 1976);
E. Cremmer, B. Julia, J. Scherk, P. van Nieuwenhuizen, S. Ferrara and L. Girardello, Phys. Lett. 79B (1978) 231
[25] S. Hawking, Phys. Rev. D14 (1976) 2460;
N. Hitchin, J. Diff. Geom. 9 (1974) 435
[26] T. Regge, private communication
[27] S. Weinberg, Phys. Rev. D13 (1976) 974; D19 (1978) 1277;
L. Susskind, Phys. Rev. D20 (1979) 2619
[28] S. Dimopoulos and L. Susskind, Nucl. Phys. B155 (1979) 237;
E. Eichten and K. Lane, Phys. Lett. 90B (1980) 125
[29] G. 't Hooft, Lecture at Cargèse Summer Inst., 1979, *in* Recent developments in gauge theories, ed. G. 't Hooft et al. (Plenum, 1980);
S. Dimopoulos, S. Raby and L. Susskind, Nucl. Phys. B173 (1980) 208;

S. Coleman and E. Witten, Phys. Rev. Lett. 45 (1980) 100;
Y. Frishman, A. Schwimmer, T. Banks and S. Yankielowicz, Nucl. Phys. B177 (1981) 157;
A. Zee, Pennsylvania preprint (1980);
R. Barbieri, L. Maiani and R. Petronzio, CERN preprint TH. 2900 (1980);
G. Farrar, CERN preprint TH. 2909 (1980);
R. Chanda and P. Roy, CERN preprint TH. 2923 (1980);
J. Preskill and S. Weinberg, Harvard preprint (1981);
Dan-di Wu, Harvard preprint HUTP-81/A001 (1980);
I. Bars and S. Yankielowicz, Yale preprint YTP81-04 (1981)

[30] S.M. Faber and J.S. Gallagher, *in* Annual reviews of astronomy and astrophysics, ed. G. Burbridge D. Layzer and J.G. Philips (1979)