Monopole catalysis of proton decay

1988 Rep. Prog. Phys. 51 189

(http://iopscience.iop.org/0034-4885/51/2/002)

Rep. Prog. Phys. 51 (1988) 189-241. Printed in the UK

Monopole catalysis of proton decay

V A Rubakov

Institute for Nuclear Research of the Academy of Sciences of the USSR, 60th October Anniversary Prospect 7a, Moscow 117312, USSR

Abstract

The theoretical studies of the monopole catalysis of proton decay in grand unified theories are reviewed. Monopole catalysis, i.e. the ability of magnetic monopoles to induce nucleon decay at a strong interaction rate, is a non-perturbative quantum field theory phenomenon that incorporates the effects of the triangle anomaly and complex structure of the gauge theory vacuum and peculiar properties of fermions in the field of a monopole. This effect is of interest for experiment and astrophysics. Study of the properties of the Dirac equation in the background monopole field reveals that quantum mechanics fails in describing the fermion-monopole interactions. The quantum field theoretic treatment is possible within the s-wave approximation where only spherically symmetric fields are considered. We review various techniques and results of the s-wave theory taking a toy SU(2) model and a realistic SU(5) grand unified theory as examples. We also discuss the corrections to the s-wave approximation and the model dependence of the catalysis. In conclusion, we briefly summarise the present status of and future plans for experimental monopole searches.

This review was received in March 1987.

0034-4885/88/020189+53$09.00

Contents

1. Introduction

1.1. Monopole catalysis effect

Interactions of elementary particles, as we know them, are described by gauge theories. The gauge theory of quarks and gluons based on the colour gauge group $SU(3)_c$—quantum chromodynamics (QCD)—is nowadays the ultimate theory of strong interactions. Weak and electromagnetic interactions are successfully unified within the standard electroweak theory with the gauge group $SU(2)_L \times U(1)$. Further development of the unification idea has led to the appealing hypothesis that strong, weak and electromagnetic interactions are low-energy manifestations of a grand unified theory (GUT) whose gauge group breaks down to $SU(3)_c \times SU(2)_L \times U(1)$ at a very high energy scale (for reviews see, e.g., Langacker (1981) and Ross (1981)).

GUT predict superheavy vector and scalar bosons mediating perturbatively baryon-number non-conserving reactions such as proton decay (Pati and Salam 1973, Georgi and Glashow 1974) and neutron–antineutron oscillations (Kuzmin 1970, Glashow 1979, Mohapatra and Marshak 1980, Kazarnovsky *et al* 1980, Chetyrkin *et al* 1981) and presumably relevant to the generation of the observed excess of baryons over antibaryons in the universe (Sakharov 1967, Kuzmin 1970, Ignatiev *et al* 1978, Yoshimura 1978).

Presently, it is understood that the physical content of non-Abelian gauge theories is by no means exhausted by perturbation theory. An important non-perturbative property is the non-trivial structure of a gauge theory vacuum and related non-conservation of fermion quantum numbers. In four-dimensional gauge theories the complex vacuum structure has been established within the functional integral approach by 't Hooft (1976a, b), Callan *et al* (1976) and Jackiw and Rebbi (1976a); it is closely connected to the instanton solutions of Belavin *et al* (1975). In exactly solvable two-dimensional gauge models, the non-trivial vacua were constructed explicitly (Krasnikov *et al* 1980). In QCD, the vacuum structure effects lead to the non-perturbative non-conservation of chirality and provide a possible solution to the U(1) problem ('t Hooft 1976a, b). They also raise a problem of strong *CP* conservation ('t Hooft 1976a, b) which can be solved by the invention of an extra classical symmetry and axion (Peccei and Quinn 1977, Weinberg 1978, Wilczek 1978). In the electroweak theory, the non-trivial vacuum structure results in the non-conservation of baryon and lepton numbers ('t Hooft 1976a, b). In usual circumstances, it occurs as a tunnelling effect, so the corresponding amplitudes are exponentially suppressed and thus practically negligible. However, the non-perturbative electroweak baryon-number non-conservation is enhanced at high fermionic densities (Rubakov and Tavkhelidze 1985, Matveev *et al* 1986a, b), at high temperatures (Kuzmin *et al* 1985) and also in decays of heavy particles (Rubakov 1985a, b, Ambjørn and Rubakov 1985, Rubakov *et al* 1985), so it is of potential importance for cosmology and experiment.

Another non-perturbative aspect of gauge theories is the existence of solitons appearing already at the classical level of field theory as finite-energy solutions to the field equations. Of particular interest are the magnetic monopoles of 't Hooft (1974b) and Polyakov (1974) and the dyons of Julia and Zee (1975) (for reviews see Goddard and Olive (1978) and Coleman (1981)). As found by Tyupkin *et al* (1975) and

Monastyrski and Perelomov (1975), the existence of magnetic monopoles is characteristic to all GUT based on a (semi)simple gauge group. Schematically, a monopole is an extended object with a tiny core containing superheavy vector and scalar bosons (mediating baryon-number non-conservation); far outside this core, only magnetic (and, possibly, colour magnetic) fields are present.

Several properties of monopoles can be evaluated at the classical level; these are mass ($M_M \sim M_{GU}/\alpha \sim 10^{16}$ GeV where M_{GU} is the grand unification scale and α is the fine structure constant; the numerical value given is for the SU(5) model of Georgi and Glashow (1974)), radius of the core ($r_M \sim M_{GU}^{-1} \sim 10^{-29}$ cm), magnetic and colour magnetic charges, etc. At the quantum level, monopole solutions correspond to topologically stable particles with form factors determined by the classical field configuration (Jackiw 1977, Faddeev and Korepin 1978, Rajaraman 1982 and references therein).

This review is devoted to the discussion of a peculiar property of magnetic monopoles in GUT which appears at the quantum level and is closely related to the structure of the ground state in gauge theories. Namely, we consider the ability of monopoles to induce baryon-number non-conserving reactions of the following type:

$$\mathrm{p} + \mathrm{mon} \rightarrow \mathrm{e}^+ + \mathrm{mon}\ (+\mathrm{pions})$$

with large cross sections typical to strong interactions, roughly of order of 1 fm^2 (monopole catalysis of proton decay; for an alternative review see Craigie (1986a)). The idea of monopole catalysis was put forward by this author (Rubakov 1981a, b, 1982) and later and independently by Callan (1982a, b, 1983). We mention here also the paper by Wilczek (1982) who claimed independently that the baryon-number-violating reactions in the presence of a monopole proceed at weak interaction rates; although this claim is not correct in most GUT, it is close to the results of Rubakov and Callan in the sense that the rate suggested is much higher than that typical to GUT (cf Dokos and Tomaras 1980).

The interest in monopole catalysis is for several reasons. First, the theoretical studies of this effect shed some light on non-perturbative aspects of quantum gauge theories. Second, magnetic monopoles and monopole catalysis of proton decay are among the few predictions of grand unified theories which could be, in principle, tested experimentally; several large scale underground and underwater detectors are searching for monopoles and proton decays induced by them and some new installations are under preparation. Third, the properties of monopoles, including monopole catalysis of proton decay, are of importance for cosmology and astrophysics. Finally, monopole catalysis of proton decay serves as an inspiring example of an exothermal catalytic reaction predicted by particle theory.

There are two basic observations behind the idea of monopole catalysis. The first one is connected to the properties of the Dirac equation in the presence of a monopole, i.e. it emerges at the level of quantum mechanics of fermions in the background monopole field. As found by Tamm (1931), the conserved angular momentum of a charged particle in the presence of a monopole contains an extra part, so that the angular momentum of a fermion is integer, rather than half-integer†, $J = 0, 1, 2, \ldots$. A crucial property of the lowest angular momentum fermions (s waves) is the absence

† This may not be the case for non-minimal magnetic charge monopoles. However, the discussion below applies to non-minimal cases as well.

of a centrifugal barrier (Dereli *et al* 1975, Kazama *et al* 1977). These fermions fall on the monopole centre and easily feel the structure of the monopole core, however tiny it is (Kazama *et al* 1977, Goldhaber 1977, Callias 1977). The monopole core contains superheavy vector bosons of a GUT; interactions with these bosons may naturally lead to rapid baryon-number non-conservation. In fact, this picture is oversimplified: we shall see in § 2 that quantum mechanical analysis is not appropriate for the study of monopole catalysis since it contradicts the conservation of electric charge. A consistent description is possible only within quantum field theory.

The second observation (Rubakov 1981a, 1982) is related to the structure of the ground-state and triangle-anomaly effects in the presence of a monopole. We shall see in § 3 that these effects are unsuppressed in the monopole sector (unlike the vacuum sector), and give rise to the rapid non-conservation of fermion quantum numbers. In some models this provides the only resolution of the electric-charge non-conservation puzzle; in any case, the vacuum structure and anomaly play an important role in the dynamics of the monopole catalysis. We note here that the relevance of the vacuum structure effects to monopole and dyon physics was recognised by Witten (1979) who obtained a relation between the dyon electric charge and the vacuum θ angle; the anomaly effects in the background field of a dyon were studied by Blaer *et al* (1981, 1982) (see also Marciano and Pagels 1976), independently of Rubakov and Callan.

Both these observations imply that monopole catalysis is a non-perturbative quantum field theory phenomenon. It is not a tunnelling effect either, so the standard techniques of quantum field theory (perturbation theory and semiclassical approximation) are of no use. Fortunately, there is a simplification making the field theoretic treatment of the catalysis possible and, in fact, rather straightforward. Namely, one can restrict onself to the s-wave sector of the theory (s-wave approximation). Indeed, the absence of a centrifugal barrier is the property of s-wave fermions only; the enhancement of the vacuum structure and anomaly effects is due to the s-wave fluctuations of the bosonic fields around a monopole. Thus it is natural to switch off all higher partial waves (both fermionic and bosonic); the angular dependence of remaining s-wave fields is fixed, so they are described by an effective two-dimensional field theory on a half-plane ($0 \leq r < \infty$, $-\infty < t < \infty$). This field theory can be treated non-perturbatively by making use of various techniques developed for two-dimensional models. In particular, the s-wave dynamics in the simplest SU(2) example with massless fermions can be solved exactly (Rubakov 1981a, 1982, Callan 1982a, b) in complete analogy to the two-dimensional massless quantum electrodynamics (Schwinger 1962, Lowenstein and Swieca 1971). The model with massive fermions is conveniently analysed within the bosonisation approach of Callan (1982b) which is a modification of the bosonisation technique developed for two-dimensional theories by Coleman (1975), Mandelstam (1975) and Pogrebkov and Sushko (1975) (see also Dell'Antonio *et al* 1972). The effects of strong (and weak) interactions in the s-wave sector of realistic GUT can be studied either by the bosonisation technique (Rubakov and Serebryakov 1983) or within the $1/N_c$-expansion method (Craigie *et al* 1984), the latter being analogous to that invented by 't Hooft (1974a) in the context of two-dimensional QCD. Some exact results concerning the s-wave strong interactions were obtained also on rather general grounds by Craigie and Nahm (1984a, b).

Numerous studies performed during the last few years have led to a reasonable understanding of interactions of fermions (quarks and leptons) with monopoles that give rise to the monopole catalysis of proton decay. However, several important points are still unclear. The most serious problem is the calculation of the catalysis cross section and branching ratios. Needless to say, these quantities are central for experimental and astrophysical implications. However, their evaluation meets severe difficulties in matching large distances ($r \gg 1$ fm, where a proton interacts with a monopole due to electric charge and magnetic moment), intermediate ones ($r \sim 1$ fm, where the quark structure of a proton and confinement come into play) and short distances ($r \ll 1$ fm, where monopole-quark interactions result in the baryon-number non-conservation). There are a few results and suggestions in this direction (Rubakov and Serebryakov 1983, Arafune and Fukugita 1983a, Schmid 1983, Callan and Witten 1984, Bernreuther and Craigie 1985); however, much further work is needed to get firm estimates.

Another important aspect is the model dependence of the monopole catalysis effect. It has been found that catalysis does occur in many GUT (Dawson and Schellekens 1983a, Goldhaber 1983, Nair 1983, Schellekens 1984a, b) and for many monopole types (Rubakov and Serebryakov 1983a, Schellekens 1984a, b). However, the catalysis is absent in some GUT (Dawson and Schellekens 1983a, Schellekens 1984a, b, Rubakov and Serebryakov 1984), so the effect is model dependent (unlike the very existence of monopoles). A related question is whether or not the proton decay is catalysed by the Kaluza–Klein monopoles found by Sorkin (1983) and Gross and Perry (1983) in five dimensions and generalised to higher dimensions by Bais and Batenburg (1985).

This review is organised as follows. In the rest of § 1 we summarise some well known results concerning vacuum structure and anomalous violation of fermion quantum numbers (§ 1.2), SU(5) GUT (§ 1.3), the structure of magnetic monopoles in a toy SU(2) model (§ 1.4) and in the realistic SU(5) GUT (§ 1.5). These provide the necessary background for the discussion of fermion–monopole interactions.

Before coming to the realistic GUT monopoles, we consider monopole–fermion interactions in the simplified SU(2) model of Georgi and Glashow (1972) in §§ 2–4. Section 2 is devoted to the massless Dirac equation in the background field of the 't Hooft–Polyakov monopole in the SU(2) model. We identify the conserved angular momentum and electric charge of fermions, determine the asymptotic states of the s-wave fermions and conclude that either fermion number or electric charge must be violated near the monopole core (§ 2.1). At the quantum mechanical level it is the electric charge that is not conserved (there is simply no source for the fermion number violation at this level); we check this point explicitly in § 2.2. Section 2.2 also contains a preliminary discussion of the processes allowed by the conservation laws in the s-wave monopole–fermion interactions.

Thus, quantum mechanics in the background monopole field fails in describing fermion–monopole interactions in the s-wave sector since it contradicts the electric charge conservation. Furthermore, there should exist an efficient source for the fermion number violation in quantum field theory. The only known source is the triangle anomaly and complex structure of the ground state. In § 3.1, we give a simple argument implying that the anomaly effects are not suppressed in the presence of a monopole. We also stress the importance of the s-wave fluctuations of the gauge field. In § 3.2 we formulate and solve an approximate s-wave quantum field theory while in § 3.3 we calculate some fermion-number-violating Green functions and find that they indeed contain no suppression factors. This indicates strongly that the fermion-number-non-conserving reactions proceed at high rates in the presence of a monopole. Interactions of fermions with the monopole core induce some other flavour-mixing Green functions not related to the triangle anomaly; these are discussed in § 3.4. In realistic GUT they also violate the baryon number. Section 3.5 summarises known results concerning potentially dangerous corrections to the s-wave approximation; the conclusion is that they do not make the s-wave approximation unreliable.

In § 4 we describe the bosonisation approach to s-wave fermion–monopole dynamics. In this approach, fermions far outside the core are represented by solitons of the effective bosonic theory (§ 4.1). In § 4.2 the Hamiltonian and ground states of the equivalent bosonic theory are presented, while in § 4.3 the soliton scattering off the monopole is considered, which leads to the non-conservation of fermion quantum numbers.

Section 5 is devoted to the catalysis of proton decay by the simplest (fundamental) monopole in a realistic SU(5) GUT. In § 5.1 we present an oversimplified treatment where SU(5) is reduced to the $SU(2)_M$ of §§ 2–4. We translate the results concerning the fermion Green functions and soliton scattering in the bosonised theory; at this point the monopole catalysis of proton decay emerges. In § 5.2 we turn on the s-wave fields absent in the $SU(2)_M$ model but present in the SU(5) theory. We find that the full s-wave theory has the gauge group $SU(2)_M \times SU(2)_c \times U(1)'$. We outline an approach to the extra $SU(2)_c \times U(1)'$ interactions based on a $1/N_c$ expansion; this approach is especially convenient for calculating the anomalous Green functions. An alternative bosonised treatment is considered in § 5.3. In this formalism, the $SU(2)_c$ interactions decouple, so the s-wave bosonised dynamics is still tractable. As an example, we present a simple picture of proton decay in the presence of a monopole in a bag-like model of a proton.

The main conclusion of §§ 5.2 and 5.3 is that the complications arising in the SU(5) theory do not lead to the suppression of the monopole catalysis. In § 5.4 we discuss the problems arising in estimating the catalysis cross section.

In § 6 we outline some results on the dependence of the catalysis effect on the choice of a GUT and on the type of monopole.

In conclusion (§ 7) we briefly summarise the present status and future plans for the experimental searches for monopoles and monopole catalysis of proton decay.

1.2. Gauge theory vacua and anomalous non-conservation of fermion quantum numbers

The non-trivial structure of vacuum is a common property of non-Abelian gauge theories (see reviews by Jackiw (1977), Coleman (1978) and Rajaraman (1982)). We consider here for simplicity an SU(2) gauge theory with or without the Higgs fields. We use matrix notations for the gauge field,

$$A_\mu = -ig\tfrac{1}{2}\tau^a A_\mu^a$$

where g is the gauge coupling constant, A_μ^a $(a = 1, 2, 3)$ are real vector fields and τ^a are the Pauli matrices. Under the gauge transformations $\omega(x) \in SU(2)$, the gauge field transforms as follows:

$$A_\mu \to A_\mu^{(\omega)} = \omega A_\mu \omega^{-1} + \omega \partial_\mu \omega^{-1}.$$

It is convenient to use the temporal gauge,

$$A_0 = 0. \tag{1.1}$$

In this gauge, the residual gauge transformations are determined by time-independent gauge functions, $\omega = \omega(\boldsymbol{x})$. We restrict ourselves to gauge functions constant at spatial infinity†,

$$\lim_{|\boldsymbol{x}|\to\infty} \omega(\boldsymbol{x}) = \text{constant}.$$

Under this condition, gauge functions can be characterised by the topological number

$$n[\omega] = \frac{1}{24\pi^2}\int \mathrm{d}^3x\varepsilon^{ijk}\,\mathrm{Tr}(\partial_i\omega\omega^{-1}\partial_j\omega\omega^{-1}\partial_k\omega\omega^{-1})$$

which is integer valued and invariant under continuous deformations of $\omega(\boldsymbol{x})$. An example of a gauge function with topological number n is

$$\omega_n(\boldsymbol{x}) = \exp(\mathrm{i}\tfrac{1}{2}\boldsymbol{\tau n}\Omega_n(r)) \tag{1.2}$$

where $\boldsymbol{n} = \boldsymbol{x}/r$ and $\Omega_n(r)$ is an arbitrary smooth function obeying

$$\Omega_n(r=0) = 0 \qquad \Omega_n(r=\infty) = 2\pi n. \tag{1.3}$$

The classical vacua are pure gauge configurations

$$\boldsymbol{A} = \omega\boldsymbol{\partial}\omega^{-1}.$$

If the model contains the Higgs field φ, then also $\varphi = \varphi_0^{(\omega)}$ where φ_0 is the Higgs expectation value at $\boldsymbol{A} = 0$ and $\varphi_0^{(\omega)}$ is the gauge transform of φ_0. Topologically distinct vacua correspond to gauge functions with different topological numbers. Thus there exists an infinite discrete set of gauge-equivalent classical vacua labelled by an integer n (see figure 1). This structure survives after the quantisation: there exists an infinite

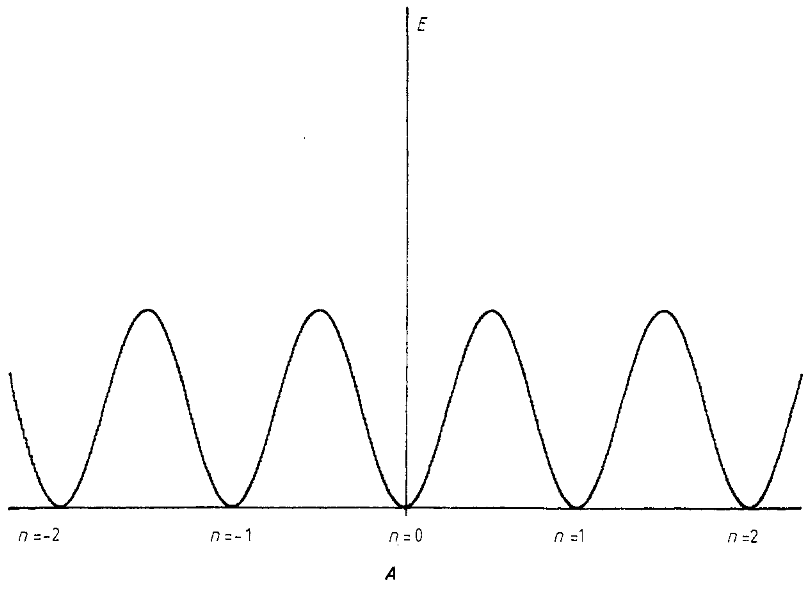

Figure 1. A schematic plot of the classical static energy against the gauge field. Different minima correspond to topologically distinct vacua. The wavefunction of a state $|n\rangle$ is concentrated near the nth minimum.

† This restriction is sometimes too strong in the monopole sector; an example is given by Sathiapalan and Tomaras (1983).

series of states $|n\rangle$, $n=0, \pm 1, \pm 2, \ldots$, gauge equivalent to each other and thus degenerate. Namely, the trivial classical vacuum, $\boldsymbol{A}=0$, corresponds, in quantum theory, to the perturbation theory vacuum, $|0\rangle$, while the state $|n\rangle$ is

$$|n\rangle = U[\omega_n]|0\rangle \tag{1.4}$$

where $U[\omega_n]$ is a unitary operator implementing a gauge transformation ω_n with $n[\omega_n]=n$. We can take $\omega_n=(\omega_1)^n$ so that

$$U[\omega_n]=(U[\omega_1])^n \tag{1.5}$$

where $U[\omega_1]$ implements the gauge transformation with unit topological number; all other choices differ from (1.5) by topologically trivial gauge transformations which are taken care of by the Gauss' law constraint imposed on all physical states.

In quantum theory, tunnelling transitions between distinct vacua (say, $n=0$ and $n\neq 0$) are possible. These are described by Euclidean field configurations obeying the following boundary conditions:

$$\begin{aligned} &\boldsymbol{A}\to 0 \qquad \varphi\to\varphi_0 \qquad \text{at } t\to-\infty \\ &\boldsymbol{A}\to \omega_n \boldsymbol{\partial}\omega_n^{-1} \qquad \varphi\to\varphi_0^{(\omega_n)} \qquad \text{at } t\to+\infty. \end{aligned} \tag{1.6}$$

For these configurations one finds

$$Q[A]=n \tag{1.7}$$

where

$$Q[A]=-\frac{1}{32\pi^2}\int \mathrm{d}^3x\,\mathrm{d}t\,\mathrm{Tr}\,F_{\mu\nu}\tilde{F}_{\mu\nu} \tag{1.8}$$

is the winding number of the gauge field, $\tilde{F}_{\mu\nu}=\frac{1}{2}\varepsilon_{\mu\nu\lambda\rho}F_{\lambda\rho}$.

The semiclassical tunnelling amplitude is determined by the field configurations obeying the condition (1.6) and minimising the Euclidean action. These are the (multi-) instantons of Belavin *et al* (1975) transformed into the temporal gauge (complications in models with the Higgs fields were discussed by 't Hooft (1976b)). The Euclidean action for the n-instanton solution is equal to $8\pi^2 n/g^2$, so the tunnelling amplitude is proportional to $\exp(-8\pi^2 n/g^2)$. The fact that it is exponentially suppressed justifies the semiclassical approximation in weakly coupled theories.

Since topologically distinct vacua tunnel into each other, none of them is an eigenstate of the Hamiltonian. The eigenstates are constructed by noticing that the Hamiltonian is gauge invariant, i.e. it commutes with unitary operators implementing gauge transformations. Thus the Hamiltonian and, say, $U[\omega_1]$ are diagonalised simultaneously, the eigenstates (θ vacua) being

$$|\theta\rangle=\sum_{n=-\infty}^{\infty}\exp(-\mathrm{i}n\theta)|n\rangle \tag{1.9}$$

where $0\leqslant\theta\leqslant 2\pi$, θ is an arbitrary parameter. Different θ sectors of the theory (consisting of physical states obtained by the action of gauge-invariant operators on different θ vacua) are orthogonal to each other. Therefore, one can fix the value of θ which becomes a new parameter of the theory along with coupling constants, masses, etc. As found by 't Hooft (1976a, b) in theories with massless fermions, the transitions between the regions of the configuration space surrounding topologically distinct classical vacua give rise to the non-conservation of fermion number† (or chirality, depending on the model). Consider, for example, the SU(2) model with a left-handed fermionic doublet ψ_L (as shown by Witten (1982), the SU(2) models with an even number of left-handed doublets make sense only; however, the arguments below are valid for each doublet separately). The gauge-invariant fermionic current

$$J_\mu = \bar{\psi}_L \gamma_\mu \psi_L$$

is not conserved because of the triangle anomaly of Adler (1969), Bell and Jackiw (1969) and Bardeen (1969):

$$\partial_\mu J_\mu = \frac{1}{32\pi^2} \operatorname{Tr} F_{\mu\nu} \tilde{F}_{\mu\nu}. \tag{1.10}$$

Equations (1.8) and (1.10) imply that the fermion number is violated in the presence of gauge fields with the asymptotics (1.6) which describe the transitions between n vacua with different topological numbers:

$$\Delta N_F = \int d^3x \, dt \, \partial_\mu J_\mu = -n. \tag{1.11}$$

This non-conservation emerges in the following way (Callan *et al* 1978, Christ 1980, Nielsen and Ninomiya 1983, Ambjørn *et al* 1983 and references therein). Let the gauge field $\boldsymbol{A}(\boldsymbol{x}, t)$ adiabatically interpolate between the asymptotic values (1.6) (at the moment we treat t as the Minkowski time). At each t one can evaluate the fermionic spectrum (the eigenvalues of the Dirac Hamiltonian in the background field $\boldsymbol{A}(\boldsymbol{x}, t)$, t being fixed) which also changes adiabatically. The initial and final spectra are the same (and are free, since $\boldsymbol{A}$ is zero or pure gauge at $t = \pm\infty$); however, the motion of levels is non-trivial (see figure 2). Namely, some positive energy levels cross zero from

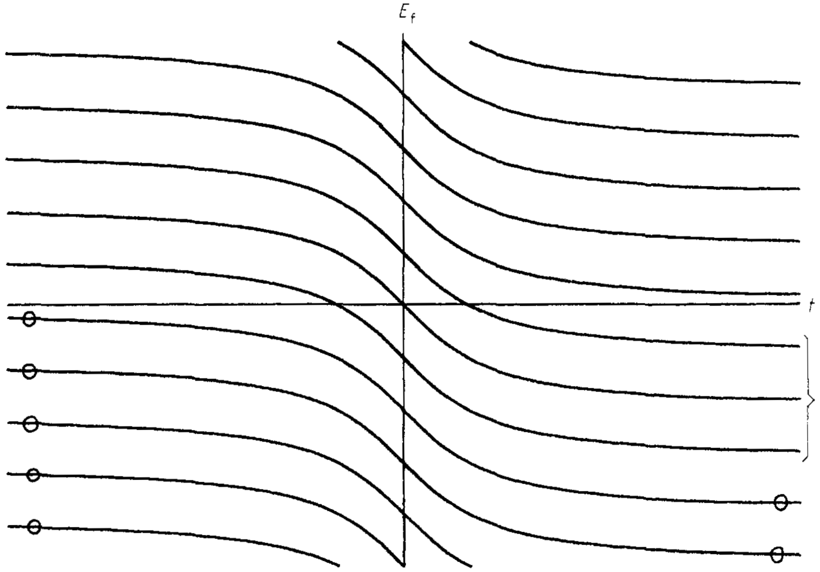

Figure 2. Level crossing phenomenon.

† This is also true in some models where fermions acquire masses via interactions with the Higgs field; an example is provided by the standard electroweak theory (Krasnikov *et al* 1979a).

above and become negative energy ones and vice versa. It turns out that the difference between the numbers of level crossings from above and from below is exactly equal to n (figure 2 corresponds to the case where no levels cross zero from below, and it is assumed that $n>0$). Thus, if an initial state were, for example, the fermionic vacuum (all negative energy levels and no positive energy ones were occupied), then the final state would contain n antifermions (empty negative energy levels). This picture of the fermion-number non-conservation is valid also for non-adiabatic fields $\boldsymbol{A}$ (Christ 1980); it can be also translated into the Euclidean formalism (Lavrelashvili 1986) appropriate for describing the tunnelling (instanton) transitions.

It is worth noting that the same picture is valid for every left-handed fermionic flavour present in the model. For right-handed fermions one has $\Delta N_{\mathrm{F}}=+n$ instead of equation (1.11). So, the selection rule has in general the following form

$$\Delta N_{\mathrm{F}}^{(i,\mathrm{L})}=-\Delta N_{\mathrm{F}}^{(j,\mathrm{R})}=-n \tag{1.12}$$

where (i, L) and (j, R) label left-handed and right-handed fermions respectively. Thus, in vector-like theories like QCD it is chirality that is violated while fermion number is conserved; in $(V-A)$ theories like the standard electroweak theory fermion (baryon and lepton) number is not conserved.

In the Green function language, the fermion number (or chirality) violation is manifested by non-zero matrix elements of operators carrying fermion number, between the states $|0\rangle$ and $|n\rangle$,

$$\langle n|\hat{O}_n|0\rangle \neq 0$$

where $\hat{O}_n$ must obey the selection rule (1.12), i.e. $\hat{O}_n$ must carry n units of each left-handed flavour and $(-n)$ units of each right-handed one. Clearly, these matrix elements are proportional to the amplitude of the transition between the topologically distinct classical vacua. For gauge-invariant operators $\hat{O}_n$ one finds

$$\langle\theta|\hat{O}_n|\theta\rangle=\exp(\mathrm{i}n\theta)\langle n|\hat{O}_n|0\rangle$$

so the values of $\langle n|\hat{O}_n|0\rangle$ measure the magnitude of the fermion-number violation. The fact that these matrix elements are non-zero was established by 't Hooft (1976a, b) who also performed an explicit calculation of the instanton contribution to the fermion-number-violating Green functions.

1.3. Grand unification: SU(5)

To fix the notation, we present here the basic formulae of the simplest grand unified theory based on the gauge group SU(5) (see reviews by Langacker (1981) and Ross (1981)). Colour and weak groups are embedded into SU(5) in the following way:

$$\mathrm{SU}(3)_{\mathrm{c}}=\begin{pmatrix}\mathrm{SU}(3) & 0\\ 0 & \mathbb{1}_2\end{pmatrix}$$

$$\mathrm{SU}(2)_L=\begin{pmatrix}\mathbb{1}_3 & 0\\ 0 & \mathrm{SU}(2)\end{pmatrix}$$

and the electromagnetic charge matrix is

$$Q_{\mathrm{EM}}=\mathrm{diag}(-\tfrac{1}{3},-\tfrac{1}{3},-\tfrac{1}{3},1,0).$$

Hereafter $\mathbb{1}_k$ is a k-dimensional unit matrix.

Left-handed fermions and antifermions of each generation form $\bar{5}$- and 10-plets (the latter is the second rank antisymmetric tensor). For example, the first generation is

$$\bar{5} = (\bar{\mathrm{d}}_1, \bar{\mathrm{d}}_2, \bar{\mathrm{d}}_3, \mathrm{e}^-, \nu_e)_L$$

$$10 = \begin{pmatrix} \bar{\mathrm{u}}_3 & -\bar{\mathrm{u}}_2 & \mathrm{u}^1 & \mathrm{d}^1 \\ & \bar{\mathrm{u}}_1 & \mathrm{u}^2 & \mathrm{d}^2 \\ & & \mathrm{u}^3 & \mathrm{d}^3 \\ & & & \mathrm{e}^+ \end{pmatrix}_L \tag{1.13}$$

where a bar denotes an antiparticle and the indices 1, 2, 3 refer to colour. In equation (1.13) we neglect the Kobayashi–Maskawa mixing.

SU(5) breaks down to $SU(3)_c \times U(1)_{EM}$ in two steps. At the first step, SU(5) breaks down to $SU(3)_c \times SU(2)_L \times U(1)$ due to the vacuum condensation of the Higgs 24-plet (adjoint representation),

$$\langle \Phi_{24} \rangle = V \,\mathrm{diag}(2, 2, 2, -3, -3) \qquad V \sim 10^{15}\ \mathrm{GeV}.$$

At the second step, $SU(3)_c \times SU(2) \times U(1)$ breaks down to $SU(3)_c \times U(1)_{EM}$ due to either the Higgs 5-plet or 45-plet or both. The former case is the simplest one; the vacuum expectation value is

$$\langle \Phi_5 \rangle = v \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \qquad v \approx 250\ \mathrm{GeV}.$$

Although this possibility (minimal SU(5) model) is ruled out by the proton decay experiments (see, e.g., Totsuka 1985), we shall concentrate on it for simplicity; our discussion will be valid for other variants as well.

The Higgs 24-plet is not directly coupled to fermions; the fermions acquire masses due to their Yukawa couplings to the Higgs 5-plet.

1.4. SU(2) monopole

The simplest model where the 't Hooft–Polyakov monopoles are present is the SU(2) gauge model of Georgi and Glashow (1972). Except the gauge field, the model contains a Higgs triplet $\Phi = \Phi^a \tau^a$ where Φ^a $(a = 1, 2, 3)$ are real scalar fields. The vacuum expectation value of Φ is non-zero; in the unitary gauge

$$\langle \Phi \rangle = v\tau^3$$

and vector fields $W^{\pm}_\mu = (1/\sqrt{2})(A^1_\mu \pm \mathrm{i}A^2_\mu)$ acquire masses $M_V = gv$ while A^3_μ remains massless ('photon'). The unbroken group $U(1)_{EM}$ is generated by τ^3 (in the unitary gauge). The fact that $U(1)_{EM}$ is unbroken follows from the transformation law of the Higgs field under the gauge transformations $\omega(x) \in SU(2)$,

$$\Phi \to \Phi^{(\omega)} = \omega \Phi \omega^{-1}$$

which implies that $\langle \Phi \rangle$ is invariant under $U(1)_{EM}$. We note also that the electric charges of $W^{\pm}$ are equal to $\pm g$, while the minimal possible electric charge would be carried by SU(2) doublets,

$$q_{\min} = \tfrac{1}{2}g. \tag{1.14}$$

The 't Hooft-Polyakov monopole is a static smooth finite energy solution to the classical field equations. It has the following form:

$$A_0^{\mathrm{cl}} = 0$$
$$A_i^{\mathrm{cl}} = \frac{\varepsilon_{aij}\tau^a n_j}{2\mathrm{i}r}[1 - K(r)] \tag{1.15}$$
$$\Phi^{\mathrm{cl}} = \tau^a n_a v[1 - H(r)]$$

where $K(r)$ and $H(r)$ obey

$$K(0) = H(0) = 1 \qquad K(\infty) = H(\infty) = 0.$$

K and H can be obtained, in principle, by solving the classical field equations. Both $K(r)$ and $H(r)$ exponentially tend to zero as $r \to \infty$, the rates being M_V and M_H, respectively, where M_H is the Higgs boson mass. Thus, a monopole has a core of a radius $r_M \sim M_V^{-1}$ (if $M_H \sim M_V$). Outside the core, the fields take their asymptotic values

$$A_i^{\mathrm{cl,as}} = \frac{\varepsilon_{aij}\tau^a n_j}{2\mathrm{i}r}$$
$$\Phi^{\mathrm{cl,as}} = \tau^a n_a v. \tag{1.16}$$

It is worth noting that the configuration (1.15) is symmetric under spatial rotations supplemented by global SU(2) transformations.

The monopole mass can be evaluated (up to $O(g^2)$ corrections) at the classical level: it is equal to the classical energy of the configuration (1.15). An order of magnitude estimate is

$$M_M \sim M_V/\alpha$$

where $\alpha = g^2/4\pi$.

In the regular gauge (1.15), the unbroken subgroup $U(1)_{EM}$ is generated by

$$t_{EM} = \tau^a n_a \tag{1.17}$$

which commutes with Φ^{cl}. It is sometimes convenient to perform a gauge transformation

$$\omega = \exp(-\mathrm{i}\tau^3\varphi/2)\exp(\mathrm{i}\tau^2\vartheta/2)\exp(\mathrm{i}\tau^3\varphi/2) \tag{1.18}$$

where φ and ϑ are the polar angles. This transformation takes the regular configuration (1.15) into the unitary gauge; at large r

$$A_i^{\mathrm{cl}(\omega)} = (g/2\mathrm{i})\tau^3 A_i^{\mathrm{D}}$$
$$\Phi^{\mathrm{cl}(\omega)} = v\tau^3 \tag{1.19}$$

where $\boldsymbol{A}^{\mathrm{D}}$ is the standard expression for the vector potential of the Dirac monopole (Dirac 1931) with magnetic charge $g_M = 4\pi g^{-1}$,

$$\boldsymbol{A}^{\mathrm{D}} = \frac{1 - \cos\vartheta}{rg\sin\vartheta}\hat{\boldsymbol{\varphi}}. \tag{1.20}$$

The gauge transformation (1.18) is singular at the semi-axis $\vartheta = \pi$; this is the reason for the appearance of a pure gauge singularity (Dirac string) in the unitary gauge potential (1.20). It is worth noting that the monopole just described carries minimal magnetic charge obeying the quantisation condition of Dirac (1931)

$$qg_M = 2\pi \times \text{integer}.$$

Indeed, the minimal electric charge in the model is given by equation (1.14). This quantisation rule (or its modifications due to colour) is valid for any GUT monopole.

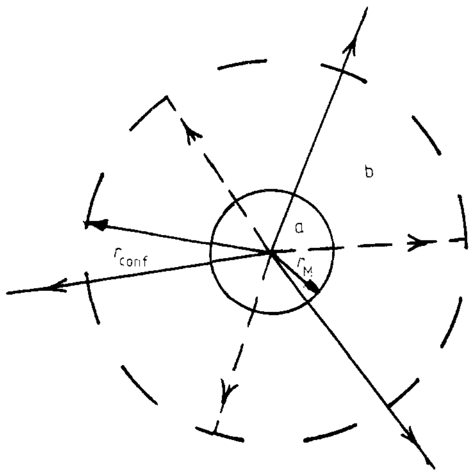

Figure 3. The structure of the fields of the fundamental SU(5) monopole. In the central region a of size $r_{\rm M} \sim 10^{-29}$ cm, the expectation value of Φ_{24} is zero, and the SU(5) symmetry is partially restored. In region b both colour and ordinary magnetic fields (represented by broken and full lines, respectively) are present. The size of this region is determined by the confinement scale, $r_{\rm conf} \sim 10^{-13}$ cm. Outside region b, the colour magnetic field is presumably screened, so only the ordinary magnetic field extends to infinity.

1.5. Fundamental monopole in SU(5)

The construction of § 1.4 can be generalised to more complicated models including GUT (for reviews see Goddard and Olive (1978), Nahm (1986)). In particular, the simplest (fundamental) SU(5) monopole (Bais 1978, Daniel *et al* 1980, Dokos and Tomaras 1980) is basically the $\mathrm{SU}(2)_M$ monopole where $\mathrm{SU}(2)_M$ is embedded into SU(5) as follows:

$$\mathrm{SU}(2)_M = \begin{pmatrix} \mathbb{1}_2 & & 0 \\ & \mathrm{SU}(2) & \\ 0 & & \mathbb{1}_1 \end{pmatrix}. \tag{1.21}$$

The only unbroken generator of $\mathrm{SU}(2)_M$ is

$$t^3_{\rm M} = \tfrac{1}{2}\,\mathrm{diag}(0, 0, 1, -1, 0) = \tfrac{1}{2}(Q_{\rm EM} + \tfrac{1}{3} Y_{\rm c})$$

where $Y_{\rm c}$ is colour hypercharge. Accordingly, the fundamental SU(5) monopole carries both the usual magnetic charge $g_M = 2\pi/e$ and colour magnetic charge. The massive vector bosons, X, of the $\mathrm{SU}(2)_M$ (analogues of $W^{\pm}$ of § 1.4) are superheavy ($M_x \sim M_{\rm GU} \sim 10^{15}$ GeV) so the monopole mass and core size are determined by the grand unification scale,

$$M_{\rm M} \sim M_{\rm GU}/\alpha \qquad r_{\rm M} \sim M_{\rm GU}^{-1}.$$

The structure of the monopole fields is illustrated in figure 3. In any realistic situation, available energies are much smaller than $r_{\rm M}^{-1}$. Nevertheless, as we shall see in § 2, s-wave fermions are able to probe the structure of the monopole core.

2. Quantum mechanics of fermions around a monopole and how it fails

2.1. Dirac equation in the presence of an SU(2) monopole: s-wave fermions and their asymptotic states

In this section we consider solutions of the massless Dirac equation in the background field of a monopole in the SU(2) model of § 1.4. Clearly, equations for left-handed and right-handed fermions decouple, so we concentrate here on the left-handed part of the Dirac equation. In the unitary gauge, a left-handed fermionic doublet contains fermions a_+ and b_- with electric charges $(+\frac{1}{2}g)$ and $(-\frac{1}{2}g)$ respectively,

$$\psi_{\mathrm{L}} = \begin{pmatrix} a_+ \\ b_- \end{pmatrix}_{\mathrm{L}}. \tag{2.1}$$

The electric charge matrix in this gauge is proportional to τ^3.

In the regular gauge, the Dirac equation in the background monopole field is

$$\mathrm{i}\sigma^\mu(\partial_\mu + A_\mu^{\mathrm{cl}})\psi_{\mathrm{L}} = 0$$

where A_μ^{cl} is given by (1.15), $\sigma^\mu = (\mathbb{1}_2, \boldsymbol{\sigma})$ and ψ_{L} are Lorentz doublets.

As pointed out in § 1.4, the monopole field is invariant under spatial rotations supplemented by global SU(2) transformations. Therefore, there exists a conserved angular momentum

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{s} + \tfrac{1}{2}\boldsymbol{\tau} \tag{2.2}$$

where $\boldsymbol{L}$ and $\boldsymbol{s}$ are the standard orbital and spin momenta, respectively, and $\boldsymbol{\tau}$ acts on the SU(2) indices. Because of the extra term $\frac{1}{2}\boldsymbol{\tau}$ ('spin from isospin', see Jackiw and Rebbi 1976b, Hasenfratz and 't Hooft 1976), the angular momentum is integer valued. This fact is in accord with the old observation by Tamm (1931) that the angular momentum in the presence of a monopole contains an additional term making the angular momentum of fermions integer; the correspondence is established by the gauge transformation† (1.18).

We first discuss the solutions outside the core ($r \gg r_{\mathrm{M}}$) where $K(r) = 0$ and the monopole configuration (1.16) coincides with the field of the Dirac monopole, equation (1.19), up to the gauge transformation (1.18). In this region, the operator $\boldsymbol{\tau n}$ is conserved. In fact, this operator is the electric charge of a fermion (up to a factor $\frac{1}{2}g$): the gauge transformation (1.18) takes it into τ^3.

Let us consider explicitly the lowest angular momentum fermions (s waves, $J = 0$). The most general spherically symmetric fermionic wavefunction has the following form (Jackiw and Rebbi 1976b):

$$\psi_{\alpha i} = \frac{1}{\sqrt{8\pi}\, r}[\varepsilon_{\alpha i}\chi_1(r, t) - \mathrm{i}\tau^a_{\alpha\beta}\varepsilon_{\beta i}n_a\chi_2(r, t)] \tag{2.3}$$

where $\alpha = 1, 2$ and $i = 1, 2$ are Lorentz and SU(2) indices, respectively, and $\chi_{1,2}$ are radial functions. We introduce a column

$$\chi = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}. \tag{2.4}$$

In these notations, the radial part of the Dirac equation outside the core is

$$(\mathrm{i}\tilde{\gamma}^0\partial_t + \mathrm{i}\tilde{\gamma}^1\partial_r)\chi = 0 \tag{2.5}$$

† This correspondence is lacking for some other spherically symmetric monopoles; examples are given by Rubakov and Serebryakov (1983) and Schellekens (1984b).

where $\tilde{\gamma}^0 = \sigma^1$, $\tilde{\gamma}^1 = -\mathrm{i}\sigma^3$ are two-dimensional γ matrices. A crucial property of (2.5) is the absence of a centrifugal barrier. This property is not shared by higher partial waves. In our notation, the electric charge operator coincides with the two-dimensional γ^5 matrix, $\tilde{\gamma}^5 = \sigma^2$. Its conservation is explicit in (2.5).

The two independent solutions outside the core are

$$\begin{aligned} \chi_+ &= (1/\sqrt{2}) \exp[-\mathrm{i}E(t+r)]\begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} \\ \chi_- &= (1/\sqrt{2}) \exp[-\mathrm{i}E(t-r)]\begin{pmatrix} 1 \\ -\mathrm{i} \end{pmatrix}. \end{aligned} \tag{2.6}$$

At $E > 0$, they describe positive charge fermions and negative charge ones, a_+ and b_-, respectively. Clearly, χ_+ is an incoming wave, while χ_- is an outgoing one; there are no transitions of an incoming s wave into an outgoing s wave outside the core. Furthermore, left-handed positive charge s-wave fermions can only fall on the monopole, while left-handed negative charge ones can only be emitted. For the right-handed s-wave fermions the situation is reversed: positive (negative) charge fermions can only be emitted (fall). These facts are realised by Dereli *et al* (1975) and Kazama *et al* (1977). The situation is summarised in table 1. An analogous situation takes place in the case of massive Dirac fermions (Dereli *et al* 1975, Kazama *et al* 1977), the role of chirality being played by helicity, as indicated in table 1.

Table 1. Asymptotic states of s-wave fermions.

Electric charge	Chirality (helicity at $m \neq 0$)	Direction of motion
+	L(+)	in
−	L(+)	out
+	R(−)	out
−	R(−)	in

These results imply that the s-wave fermions easily feel the structure of the monopole core. Indeed, at the quantum mechanical level, the fate of, say, an incoming positive charge left-handed fermion is determined by the interactions with the core, i.e. by the properties of the Dirac equation in the core region. This is the unique property of the s-wave fermions: higher partial waves are expelled by the centrifugal barrier long before they reach the core.

Let us deviate at this point from our main line of reasoning to discuss briefly the case of the Dirac (pointlike) monopole. The above results imply that the Dirac operator in that case is ill defined: the unitarity requires that each solution should contain both incoming and outgoing waves at equal weights. On the other hand, equation (2.5) is valid, in the case of the Dirac monopole, at any $r > 0$. So, an *a priori* boundary condition should be imposed at $r = 0$ to mix up incoming and outgoing waves (Goldhaber 1977, Callias 1977, Yamagishi 1983, Grossman 1983). One can imagine various types of boundary conditions, for example,

$$\psi_{\mathrm{L},+}^{J=0} = \mathrm{e}^{\mathrm{i}\theta}\psi_{\mathrm{R},+}^{J=0} \tag{2.7}$$

$$\psi_{\mathrm{L},+}^{J=0} = \mathrm{e}^{\mathrm{i}\theta}\psi_{\mathrm{L},-}^{J=0} \tag{2.8}$$

where θ is an arbitrary phase. The former conserves electric charge but violates chirality (if a model contains only left-handed fermions, ψ_R is to be understood as the wavefunction of an antifermion; condition (2.7) then violates fermion number). The condition (2.8) conserves chirality and fermion number but violates electric charge.

In the Dirac monopole case, the choice of the boundary condition is arbitrary. If one insists on electric charge conservation, one has to impose the condition (2.7). The phase θ plays a role analogous to the vacuum angle of § 1.2: if fermions are massive, then the fermionic vacuum rearranges in such a way that *CP* breaks down, the phase θ being the *CP*-violating parameter (Grossman 1983, Yamagishi 1983).

2.2. *Charge exchange at the core*

Coming back to the case of the 't Hooft–Polyakov monopole, we recall that the monopole configuration is non-singular (in the regular gauge). So the Dirac operator is well defined, and the amplitude of the transition of an incoming wave into an outgoing one can be calculated within quantum mechanics. In the limit $r_M \to 0$ (more precisely, at $r_M \ll \lambda$ where λ is the wavelength of a fermion) one expects that the core effects are reduced to a boundary condition like (2.7) or (2.8). Now, fermion number and chirality are conserved in quantum mechanics, so the only possibility is the boundary condition that violates electric charge, equation (2.8). Let us see this explicitly.

Making use of the notations (2.3) and (2.4) and recalling the explicit form of the monopole field, equation (1.15), we write the full massless left-handed radial equation,

$$[\mathrm{i}\tilde{\gamma}^0\partial_t + \mathrm{i}\tilde{\gamma}^1\partial_r - (K/r)(1-\sigma^3)]\chi = 0. \tag{2.9}$$

Since $K(0)=1$, the regular solution to (2.9) obeys

$$(1-\sigma^3)\chi = 0 \qquad \text{at } r=0. \tag{2.10}$$

Clearly, this relation is valid also at $r \sim r_M$ (since the fermion wavelength is much smaller than r_M). At $r \gg r_M$ one comes back to equation (2.5); the core effects can be summarised by the boundary condition (2.10). Making use of equation (2.6), one finds that the boundary condition (2.10) coincides with (2.8) for $\theta = 0$. These arguments are supported by the explicit solution of the Dirac equation obtained by Biebl (1982) and Marciano and Muzinich (1983) (see also Isler *et al* 1986) who used the explicit monopole configuration found by Bogomolny (1976) and Prasad and Sommerfield (1975) in the limit of vanishing Higgs self-coupling.

Thus, at the quantum mechanical level, the electric charge of fermions is not conserved in the presence of a monopole. This fact is not too surprising, since the monopole core contains charge vector bosons; interactions with background charged fields do, in general, violate charge conservation. The property that the charge non-conservation in the presence of a monopole is unsuppressed (at the level of the Dirac equation), in spite of a very small size of the core, is due to the absence of the centrifugal barrier for s-wave fermions which penetrate the core at unit probability.

However, the electric charge non-conservation is not physically acceptable. Rather, this result signalises the failure of the quantum mechanical description of the fermion–monopole interactions. One is inevitably led to the quantum field theoretic treatment where the charge conservation is expected to be restored. The quantum mechanical analysis provides, however, some insight into the field theory phenomena to be involved: table 1 implies that there should exist some effects leading to the violation of chirality and/or fermion number. These effects should be strong enough to convert incoming left-handed s-wave fermions into right-handed ones. Another possibility would be the charge exchange between fermions and monopole core in the processes like

$$a_{+,\mathrm{L}}+\text{monopole}\to b_{-,\mathrm{L}}+\text{dyon}. \tag{2.11}$$

However, the Coulomb energy of the dyon (mass difference between the dyon and the monopole) is of order g^2/r_{M}, so the scattering process (2.11) is forbidden by energy conservation unless the energy of the incoming fermion is extremely large ($E \gg g^2/r_{\mathrm{M}}$).

In quantum field theory, there does exist a source of the fermion number or chirality non-conservation: it is the triangle anomaly and vacuum structure. So, one expects this non-perturbative phenomenon to be unsuppressed in the presence of a monopole. We discuss this point in some detail in § 3.

To conclude this section, we point out that the properties of the asymptotic states summarised in table 1, together with the conservation laws, provide some insight into the allowed processes (Sen 1983, Callan and Das 1983). Consider, for example, a model with two massless left-handed doublets $(a_+^{(1,2)}, b_-^{(1,2)})_{\mathrm{L}}$. We find from table 1 that $a_{\mathrm{L},+}^{(1,2)}$ and $\bar{a}_{\mathrm{R},-}^{(1,2)}$ can only fall on the monopole while $b_{\mathrm{L},-}^{(1,2)}$ and $\bar{b}_{\mathrm{R},+}^{(1,2)}$ can only be emitted. Electric charge conservation implies that an incoming $a_{\mathrm{L},+}^{(1)}$ evolves into a final state which must contain $\bar{b}_{\mathrm{R},+}^{(1)}$ or $\bar{b}_{\mathrm{R},+}^{(2)}$ plus $(b\bar{b})$ pairs. Clearly, this process violates fermion number so it proceeds due to the anomaly. The selection rule (1.12) means that the changes in the fermion numbers of each doublet are the same, so the only allowed reaction for an incoming $a^{(1)}$ is

$$a^{(1)}+\text{mon}\to \bar{b}^{(2)}+\text{mon}\ (+\bar{b}b\ \text{pairs}). \tag{2.12}$$

On the contrary, if $a^{(1)}$ and $\bar{a}^{(2)}$ or $a^{(2)}$ fall on the monopole together, then two reactions are allowed,

$$a^{(1)}+a^{(2)}+\text{mon}\to \bar{b}^{(2)}+\bar{b}^{(1)}+\text{mon}\ (+b\bar{b}\ \text{pairs}) \tag{2.13}$$

$$a^{(1)}+\bar{a}^{(2)}+\text{mon}\to b^{(1)}+\bar{b}^{(2)}+\text{mon}\ (+b\bar{b}\ \text{pairs}). \tag{2.14}$$

The former violates the fermion number of every doublet by two, so it proceeds via anomaly (it is merely a combination of two reactions of the type (2.12)). The latter conserves both fermion numbers, so the anomaly is not expected to be too relevant for it. Processes of type (2.14) also occur without monopoles (they are due to the heavy vector boson exchange), but in the vacuum sector their low-energy cross sections are suppressed by the inverse powers of the vector boson mass. The peculiarity of the monopole sector is that fermions interact strongly with the core (which contains heavy vector bosons), so the reaction (2.14) is expected to be unsuppressed in the presence of a monopole.

The situation becomes more subtle if the number of fermionic doublets is larger than two. In a model with four left-handed doublets no final state seems to be consistent with the conservation laws (which include the selection rule (1.12)) for an initial $a^{(1)}$ (Callan 1984). We shall come back to this point in § 4. For the initial states containing more than one fermion, the final states can still be guessed. An example of the fermion-number-violating (i.e. anomalous) process is

$$a^{(1)}+a^{(2)}+\text{mon}\to \bar{b}^{(3)}+\bar{b}^{(4)}+\text{mon}\ (+b\bar{b}\ \text{pairs}) \tag{2.15}$$

while the reaction (2.13) is not allowed by the selection rule (1.12). The non-anomalous reaction (2.14) is still possible.

3. Fermion Green functions in the SU(2) model

3.1. Large anomaly effects: a simple argument

In this section we continue to discuss the SU(2) model of § 2, but now at the quantum field theory level. As argued in § 2.2, the effects related to the structure of the ground state and triangle anomaly are expected to be unsuppressed in the presence of a monopole. Here we present a simple argument implying that this is indeed the case (Rubakov 1981a, 1982).

As discussed in § 1.2, the manifestations of the anomalous fermion-number non-conservation in the vacuum sector are non-zero matrix elements of operators $\hat{O}_n$ carrying certain fermion numbers, between the perturbation theory vacuum and topologically distinct ones. The discussion of § 1.2 is straightforwardly generalised to the monopole sector. One can construct different monopole configurations by gauge transforming the classical monopole field, the gauge functions $\omega(x)$ having non-zero topological numbers,

$$\begin{aligned} \boldsymbol{A}^{\mathrm{cl}(n)} &= \omega_n \boldsymbol{A}^{\mathrm{cl}} \omega_n^{-1} + \omega_n \boldsymbol{\partial} \omega_n^{-1} \\ \Phi^{\mathrm{cl}(n)} &= \omega_n \Phi^{\mathrm{cl}} \omega_n^{-1} \\ n[\omega_n] &= n \end{aligned}$$

(we still use the temporal gauge $A_0 = 0$). These configurations correspond to degenerate monopole states,

$$|M, n\rangle = U[\omega_n]|M, 0\rangle$$

(the notation is the same as in § 1.2). The gauge-invariant monopole states $|M, \theta\rangle$ are constructed in the same way as the θ vacua, see equation (1.9).

One expects the following matrix elements to be non-zero:

$$\langle M, n|\hat{O}_n|M, 0\rangle \tag{3.1}$$

where $\hat{O}_n$ carries n units of fermion number of each left-handed flavour, and $(-n)$ units of each right-handed flavour (see the discussion of the selection rules in § 1.2). An example of $\hat{O}_1$ in a model with two left-handed doublets is

$$\psi^{(1)}\psi^{(2)} \equiv \varepsilon^{ij}\psi^{(1)}_{\mathrm{L},i}\psi^{(2)}_{\mathrm{L},j} \tag{3.2}$$

where $i, j = 1, 2$ are the SU(2) indices; notice that $\psi^{(1)}\psi^{(2)}$ is gauge invariant. The matrix elements (3.1) measure the magnitude of the anomalous fermion-number non-conservation.

The (normalised) matrix element of the form (3.1) is given by the Euclidean functional integral,

$$\frac{\langle M, n|\hat{O}_n|M, 0\rangle}{\langle M, 0|M, 0\rangle} = \int \mathrm{d}A\, \mathrm{d}\varphi \prod_s \mathrm{d}\psi^{(s)}\, \mathrm{d}\bar{\psi}^{(s)} \exp[-(S - S_{\mathrm{M}})] O_n \tag{3.3}$$

where S_{M} is the action for the monopole configuration, $S_{\mathrm{M}} = M_{\mathrm{M}} T$ (T is the normalisation time), and the integration is performed over the fields obeying the following boundary conditions:

$$\begin{aligned} &\boldsymbol{A}, \Phi \to \boldsymbol{A}^{\mathrm{cl}}, \Phi^{\mathrm{cl}} && t \to -\infty \\ &\boldsymbol{A}, \Phi \to \boldsymbol{A}^{\mathrm{cl}(n)}, \Phi^{\mathrm{cl}(n)} && t \to +\infty \end{aligned} \tag{3.4}$$

and

$$\psi, \bar{\psi} \to 0 \qquad t \to \pm\infty.$$

An important difference between the vacuum sector and the monopole sector is that the relevant gauge functions, ω_n, can be chosen in such a way that they do not change Φ^{cl},

$$\Phi^{\mathrm{cl}(n)} = \Phi^{\mathrm{cl}}. \tag{3.5}$$

This property is shared, for example, by the gauge functions (1.2). Equation (3.5) implies that ω_n belong to the unbroken subgroup $\mathrm{U}(1)_{\mathrm{EM}}$ so that the transition (3.4) is of a purely electromagnetic nature.

Let us phrase the last remark in a slightly different way. The bosonic configurations describing the transitions between $|M, 0\rangle$ and $|M, n\rangle$ and thus obeying equation (3.4) are characterised by $Q[\boldsymbol{A}] = n$: as shown by Christ and Jackiw (1980), equation (1.8) is valid also in the monopole sector. The classical field strength of the monopole outside the core is

$$\begin{aligned} &F_{0i}^{\mathrm{cl}} = 0 \qquad F_{ij}^{\mathrm{cl}} = \varepsilon_{ijk} H_k^{\mathrm{cl}} \\ &H_k^{\mathrm{cl}} = \frac{1}{2\mathrm{i}r} (\tau^a n_a) n_k. \end{aligned} \tag{3.6}$$

(Equation (3.6) just means that at large r, the monopole magnetic field contains only a $\mathrm{U}(1)_{\mathrm{EM}}$ part: $\boldsymbol{H}^{\mathrm{cl}}$ is directed in the group space along the Higgs field. Equation (3.6) implies also that the magnetic field only has a radial component: $H_k^{\mathrm{cl}} \propto n_k$.) The gauge field fluctuations with non-vanishing $Q[\boldsymbol{A}]$ may have only electric components of the field strength non-zero, $F_{0i} \neq 0$, $\delta F_{ij} = 0$. For these fluctuations $Q \propto \int \mathrm{d}^3x \, \mathrm{d}t \, \mathrm{Tr}\, F_{0i} H_i$. Furthermore, far away from the core, only the ordinary electric field, $\boldsymbol{\mathscr{E}}$, contributes to Q, $F_{0i} \propto (\tau^a n_a) \mathscr{E}_i$, and only its radial component is relevant,

$$Q \propto \int \mathrm{d}r \, \mathrm{d}t \, \mathscr{E}_r \qquad \mathscr{E}_r = \boldsymbol{\mathscr{E}} \boldsymbol{n}.$$

Thus the radial fluctuations of the ordinary electric field can have winding number non-zero in the presence of a monopole and lead to transitions between distinct monopole states.

It is worth noting that the winding number density, $q \propto \mathrm{Tr}\, F_{\mu\nu} \tilde{F}_{\mu\nu}$, of a *dyon* field is non-zero already at the classical level. The anomaly equation (1.10) implies that the fermion number should grow continuously in time in the dyon background field, i.e. the dyon field should create fermions. This process was analysed by Blaer *et al* (1981, 1982) who neglected the back reaction of the created fermions on the dyon (the latter assumption is valid for dyons with large electric charges).

Let us come back to the transitions (3.4). To see that they are not exponentially suppressed, we construct Euclidean configurations of the bosonic fields obeying (3.4) and having small action $(S - S_{\mathrm{M}})$. The contributions of these configurations into the functional integral (3.3) are not exponentially small.

Consider the following configuration:

$$\begin{aligned} &A_0 = 0 \\ &\boldsymbol{A} = h \boldsymbol{A}^{\mathrm{cl}} h^{-1} + h \boldsymbol{\partial} h^{-1} \\ &\Phi = \Phi^{\mathrm{cl}} \end{aligned} \tag{3.7}$$

where

$$h(x, t) = \exp(\mathrm{i}\tfrac{1}{2}\tau^a n_a S(r, t)).$$

For equation (3.4) to be satisfied, $S(r, t)$ should obey

$$S(r, t) \to 0 \qquad t \to -\infty$$

$$S(r, t) \to \Omega_n(r) \qquad t \to +\infty$$

where $\Omega(r)$ is an arbitrary function obeying (1.3). If we take $S(r, t)$ to be non-vanishing only outside the core, then the configuration (3.7) describes a radial fluctuation of the ordinary electric field,

$$F_{ij} = F_{ij}^{\mathrm{cl}} \qquad F_{0i} = (1/2\mathrm{i})\tau^a n_a n_i \partial_r \partial_t S.$$

We choose S in such a way that $\partial_r S$ and $\partial_t S$ rapidly tend to zero as $r \to \infty$ and $t \to \pm\infty$, respectively. The action for the configuration (3.7) is

$$S - S_{\mathrm{M}} = \frac{2\pi}{g^2} \int_{-\infty}^{\infty} \mathrm{d}t \int_0^{\infty} \mathrm{d}r[r^2(\partial_r \partial_t S)^2 + K^2(\partial_t S)^2]. \tag{3.8}$$

Since $K(r) \to 0$ as $r \to \infty$, this action is finite (this would not be the case in the vacuum sector where $K \equiv 1$: since $\partial_t S(r = \infty) \neq 0$, the action would be infinite and the configurations of the form (3.7) would not contribute to the functional integral). To see that the value of the action (3.8) can be arbitrarily small, we make a time rescaling, $t \to \lambda t$. Under this rescaling $(S - S_{\mathrm{M}}) \to \lambda^{-1}(S - S_{\mathrm{M}})$; choosing λ large enough, we can make the action to be as small as we wish. This completes the argument.

Thus, we have argued that the matrix elements (3.3) are not exponentially suppressed. The size of the configuration (3.7) can be large, so one can expect no suppression by the inverse powers of the vector boson mass either (this argument can be made more precise, see Rubakov (1981a, 1982)). We now turn to a more quantitative discussion of the fermion-number-violating matrix elements in the presence of a monopole.

3.2. *The s-wave approximation and its functional integral solution*

As discussed above, spherically symmetric fields play a central role in the monopole–fermion interactions. Indeed, only spherically symmetric fermions probe the short-distance structure of the monopole; spherically symmetric electric field fluctuations are responsible for the unsuppressed transitions between distinct monopole states. Therefore, a natural zeroth-order approximation to the monopole–fermion dynamics is the s-wave approximation where all higher partial wave fields are frozen out, while the s-wave fields are treated non-perturbatively at the quantum field theory level. Another simplification occurs when the energies of fermions are much smaller than other parameters like vector boson mass (the precise criterion will be given later on): one can disregard heavy vector and scalar bosons and take the limit $r_{\mathrm{M}} \to 0$ wherever possible.

In the SU(2) model, the s-wave massless bosonic fluctuations are determined by two functions $a_0(r, t)$ and $a_1(r, t)$,

$$\begin{aligned} A_0 &= -\mathrm{i}(\boldsymbol{\tau}\boldsymbol{n})a_0 \\ \boldsymbol{A} &= -\mathrm{i}(\boldsymbol{\tau}\boldsymbol{n})a_1 + \boldsymbol{A}^{\mathrm{cl}}. \end{aligned} \tag{3.9}$$

The s-wave fermionic fields have the form (2.3) (we consider here only left-handed fermions). Clearly, the s-wave sector of the theory is described by an effective two-dimensional model on a half-plane $0 \leq r < \infty$; $-\infty < t < \infty$. The action for this model is obtained by inserting (3.9) and (2.3) into the original four-dimensional action. One finds

$$S = S_a + S_\chi \tag{3.10}$$

$$S_a = -\frac{2\pi}{g^2} \int \mathrm{d}r\,\mathrm{d}t\, r^2 f_{\alpha\beta}^2 \tag{3.11}$$

$$S_\chi = \sum_s \int \mathrm{d}r\,\mathrm{d}t\, \bar{\chi}^{(s)} \not{D}_{(2)} \chi^{(s)} \tag{3.12}$$

where $\alpha, \beta = 0, 1$, $f_{\alpha\beta} = \partial_\alpha a_\beta - \partial_\beta a_\alpha$, $\partial_0 \equiv \partial_t$, $\partial_1 \equiv \partial_r$,

$$\not{D}_{(2)} = \mathrm{i}\tilde{\gamma}^\alpha(\partial_\alpha + \mathrm{i}\tilde{\gamma}^5 a_\alpha) \tag{3.13}$$

$\tilde{\gamma}^\alpha$, $\tilde{\gamma}^5$ are two-dimensional γ matrices introduced in § 2.1 and s is the flavour index labelling fermionic doublets.

The action (3.10) is invariant under Abelian gauge transformations,

$$\begin{aligned} &a_\alpha \to a_\alpha + \partial_\alpha \beta(r, t) \\ &\chi \to \exp(-\mathrm{i}\beta\tilde{\gamma}^5)\chi. \end{aligned} \tag{3.14}$$

This gauge invariance is of course a remnant of the original SU(2) gauge symmetry, the corresponding SU(2) transformations being

$$\omega(\boldsymbol{x}) = \exp(\mathrm{i}\boldsymbol{\tau}\boldsymbol{n}\beta).$$

The two-dimensional field strength $f_{\alpha\beta}$ has only one independent component, f_{10}. It is related to the radial electric field as follows:

$$\begin{aligned} &F_{i0} = (g/2\mathrm{i})(\tau^a n_a) n_i \mathscr{E}_r \\ &\mathscr{E}_r = 2f_{10}/g. \end{aligned} \tag{3.15}$$

The fermion fields obey the boundary conditions (see equation (2.10))

$$(1 - \sigma^3)\chi^{(s)} = 0 \qquad \text{at } r = 0. \tag{3.16}$$

This model is analogous to the two-dimensional massless axial electrodynamics (γ^5 analogue of the Schwinger model, see Krasnikov *et al* (1979b)), the only peculiarities are the r-dependent kinetic term in equation (3.11) and the boundary condition (3.16). Like the Schwinger model, the s-wave approximation admits an exact solution (Rubakov 1981a, 1982, Callan 1982a). We present this solution in the Euclidean functional integral framework.

Any gauge field configuration can be represented as follows:

$$a_\alpha = \varepsilon_{\alpha\beta}\partial_\beta\sigma + \partial_\alpha\gamma. \tag{3.17}$$

Here $\sigma(r, t)$ is related to the radial electric field,

$$\mathscr{E}_r = (2/g)\partial^2\sigma \tag{3.18}$$

where $\partial^2 = \partial_t^2 + \partial_r^2$ is the two-dimensional Laplacian. $\gamma(r, t)$ is a pure gauge part; it does not enter gauge-invariant quantities and should be expressed in terms of σ after gauge fixing. Neither γ nor σ are defined uniquely: without changing a_α one can add to γ any harmonic function $\alpha(r, t)$ and simultaneously add to σ a function $\tilde{\alpha}$ such that $\partial_t \tilde{\alpha} = \partial_r \alpha$. We shall make use of this freedom later on.

To evaluate the integral over fermions, we make a change of variables,

$$\chi^{(s)} = \exp(-\sigma - \mathrm{i}\tilde{\gamma}^5 \gamma)\chi_0^{(s)}. \tag{3.19}$$

Then the fermionic part of the action takes the free form

$$S_\chi = \sum_s \int \mathrm{d}r\, \mathrm{d}t\, \bar{\chi}_0^{(s)} i\tilde{\gamma}^\alpha \partial_\alpha \chi_0^{(s)}. \tag{3.20}$$

Thus, the transformation (3.19) makes the integration over fermions trivial, provided that the field χ_0 obeys the boundary condition independent of γ and σ. To meet the latter requirement, we make use of the above freedom in the definition of γ and σ and impose the following boundary condition:

$$\gamma(r=0) = 0. \tag{3.21}$$

Then the boundary condition (3.16) becomes

$$(1-\sigma^3)\chi_0 = 0. \tag{3.22}$$

Note that the field A_0 is regular at the origin only if $a_0(r=0)=0$, which, together with equation (3.21), implies

$$\partial_r \sigma(r=0) = 0. \tag{3.23}$$

The Jacobian of the transformation (3.19) can be evaluated by the method of Fujikawa (1980). One finds

$$\begin{aligned} J &\equiv [\det \not{D}_{(2)}]^{n_f} = \exp(-\Gamma_1) \\ \Gamma_1 &= -\frac{n_f}{2\pi} \int \mathrm{d}r\, \mathrm{d}t\, \sigma \partial^2 \sigma \end{aligned} \tag{3.24}$$

where n_f is the number of fermionic flavours. Thus the effective bosonic action is

$$\begin{aligned} S_{\text{eff}}^{\text{B}} &= S_a + \Gamma_1 \\ &= \tfrac{1}{2} \int \sigma(r, t) L_{r,t} \sigma(r, t)\, \mathrm{d}r\, \mathrm{d}t \end{aligned} \tag{3.25}$$

where

$$L_{r,t} = -n_f \partial^2/\pi + 8\pi \partial^2 r^2 \partial^2/g^2.$$

In terms of new variables, the model is described by the quadratic action

$$S = S_{\text{eff}}^{\text{B}} + S_{\chi_0}$$

which just means that the model is solved. In particular, the fermion $2k$-point functions over the perturbation theory monopole state are

$$\begin{aligned} &\langle M, 0|\chi^{(i_1)}(r_1, t_1) \ldots \bar{\chi}^{(j_k)}(r'_k, t'_k)|M, 0\rangle \\ &\quad = \langle \exp(-\sigma(r_1, t_1) - \mathrm{i}\tilde{\gamma}^5_{(i_1)} \gamma(r_1, t_1)) \ldots \exp(\sigma(r'_k, t'_k) - \mathrm{i}\tilde{\gamma}^5_{(j_k)} \gamma(r'_k, t'_k))\rangle_{\text{eff}} \\ &\quad \times \langle \chi_0^{(i_1)}(r_1, t_1) \ldots \bar{\chi}_0^{(j_k)}(r'_k, t'_k)\rangle_0 \end{aligned} \tag{3.26}$$

where $\langle\ \rangle_{\text{eff}}$ denotes the vacuum expectation value in the bosonic theory with the action (3.25) and $\langle\ \rangle_0$ refers to the free fermionic Green function in the theory (3.20) and (3.22).

To calculate the Green functions, one only needs the propagators of σ and χ_0; making use of the boundary conditions (3.23) and (3.22) we find the propagator of σ,

$$\mathscr{P}(r, t; r', t') = -\frac{\pi}{n_f}(\mathscr{D}(r-r', t-t') + \mathscr{D}(r+r', t-t') - \mathscr{R}(r, t; r', t')) \tag{3.27}$$

and the propagator of χ_0

$$G_0(r, t; r', t') = (-i\tilde{\gamma}^0\partial_t + i\tilde{\gamma}^1\partial_r)[\mathscr{D}(r-r', t-t') - \mathscr{D}(r+r', t-t')\tilde{\gamma}^1]. \tag{3.28}$$

Here $\mathscr{D}$ is the propagator of the two-dimensional free massless scalar field (Klaiber 1968)

$$\mathscr{D}(r, t) = (1/4\pi)\ln[\mu^2(r^2+t^2)]$$

(μ is an arbitrary mass scale which does not enter the gauge invariant Green functions), and

$$\begin{aligned} \mathscr{R} &= \frac{1}{2\pi} Q_d\left(1 + \frac{(r-r')^2 + (t-t')^2}{2rr'}\right) \\ d &= (\tfrac{1}{4} + n_f g^2/8\pi)^{1/2} - \tfrac{1}{2} \end{aligned} \tag{3.29}$$

Q_d is the Legendre function.

Equation (3.27) tells us that the field σ can be represented as follows (Rubakov 1982, Ezawa and Iwazaki 1983):

$$\sigma = (\pi/n_f)^{1/2}(\Sigma + \eta) \tag{3.30}$$

where η is a two-dimensional massless scalar field quantised with negative metrics and Σ is a two-dimensional field obeying (in Minkowski spacetime)

$$(\partial^2 + g^2 n_f/8\pi^2 r^2)\Sigma = 0. \tag{3.31}$$

Equations (3.18) and (3.30) imply that Σ is directly related to the radial electric field,

$$\mathscr{E}_r = \frac{g(n_f)^{1/2}}{4\pi^{3/2} r}\Sigma. \tag{3.32}$$

The field η has no physical significance; an analogous field appears also in the exact solution of the Schwinger model (Lowenstein and Swieca 1971).

3.3. *Anomalous Green functions by the cluster argument*

Now that we know the exact solution of the s-wave approximation, we can study anomalous fermion-number non-conservation in a quantitative way. Consider a model with two left-handed doublets. The simplest matrix element with fermion-number non-conservation is

$$\langle M, \theta | f(r, t) | M, \theta \rangle \tag{3.33}$$

where

$$f = \chi_1^{(1)}\chi_1^{(2)} + \chi_2^{(1)}\chi_2^{(2)}.$$

Note that f is invariant under the gauge transformations (3.14). It is related to the operator $\psi^{(1)}\psi^{(2)}$ introduced in (3.2) as follows:

$$\psi^{(1)}\psi^{(2)} = \frac{1}{4\pi r^2} f + \text{contributions from higher partial waves.} \tag{3.34}$$

We call the matrix element (3.33) the s-wave fermion condensate. In the perturbation theory, the condensate is zero. However, the anomaly tells us that it is non-zero non-perturbatively; moreover, the arguments of § 3.1 imply that it is in fact large. Here we show this explicitly (Rubakov 1981a, 1982, Callan 1982a).

It is worth noting that in the unitary gauge

$$f \propto a^{(1)}b^{(2)} + a^{(2)}b^{(1)}$$

where $a^{(s)}$, $b^{(s)}$ are the physical fields introduced in (2.1). The non-zero value of the condensate would imply the existence of the process (2.12). The large value of the condensate would signify the large probability of the process (2.12) which would open up the possibility of solving the charge conservation puzzle existing at the quantum mechanical level as discussed in § 2.2. In fact, we know already that the process (2.12) is the only one consistent with the conservation laws.

The direct calculation (Craigie *et al* 1984) of the condensate (3.33) is slightly involved as it is based on the evaluation of the functional integrals over fermions in arbitrary topologically non-trivial gauge fields. We describe here a somewhat indirect technique (Rubakov 1981a, 1982) analogous to that suggested by Nielsen and Schroer (1977a, b) in the context of the Schwinger model. Namely, we calculate the Euclidean two-point function

$$\mathscr{F}(r_1, t_1; r_2, t_2) = \langle M, \theta | f(r_1, t_1) f^{+}(r_2, t_2) | M, \theta\rangle. \tag{3.35}$$

Since the operator $f(r_1, t_1)f^{+}(r_2, t_2)$ is gauge invariant and carries zero fermion number, its expectation value over $|M, \theta\rangle$ coincides with the expectation value over the perturbation theory monopole state,

$$\mathscr{F} = \langle M, 0 | f(r_1, t_1) f^{+}(r_2, t_2) | M, 0\rangle. \tag{3.36}$$

On the other hand, different θ sectors decouple, so the function (3.35) obeys the cluster property

$$\lim_{|t_1-t_2|\to\infty} \mathscr{F}(r_1, t_1; r_2, t_2) = \langle M, \theta | f(r_1, t_1) | M, \theta\rangle \langle M, \theta | f^{+}(r_2, t_2) | M, \theta\rangle.$$

Thus the asymptotic behaviour of $\mathscr{F}$ determines the condensate up to an unimportant phase.

The calculation of $\mathscr{F}$ is straightforward. Making use of (3.26) we find

$$\mathscr{F} = \langle \exp[-2\sigma(r_1, t_1) + 2\sigma(r_2, t_2)]\rangle_{\text{eff}} \operatorname{Tr}[G_0(r_1, t_1; r_2, t_2) G_0(r_2, t_2; r_1, t_1)]. \tag{3.37}$$

The functional integral over $d\sigma$ giving the bosonic factor in (3.37) is Gaussian and can be easily evaluated,

$$\begin{aligned}&\langle \exp[-2\sigma(r_1, t_1) + 2\sigma(r_2, t_2)]\rangle_{\text{eff}}\\ &\quad = \exp\{-4[2\mathscr{P}(r_1, t_1; r_2, t_2) - \mathscr{P}(r_1, t_1; r_1, t_1) - \mathscr{P}(r_2, t_2; r_2, t_2)]\}.\end{aligned}$$

The main contribution comes from the neighbourhood of the saddle-point field,

$$\sigma_c(r, t) = \sigma^{-}(r, t | r_1, t_1) + \sigma^{+}(r, t | r_2, t_2) \tag{3.38}$$

where

$$\sigma^{\pm}(r, t | r', t') = \mp 2\mathscr{P}(r, t; r', t').$$

Making use of the explicit forms of $\mathscr{P}$ and G_0, equations (3.27) and (3.28), we find

$$\lim_{|t_1-t_2|\to\infty} \mathscr{F}(r_1, t_1; r_2, t_2) = \frac{1}{16\pi^2 r_1 r_2}(1 + O(g^2)) \tag{3.39}$$

so that

$$\langle M, \theta | f(r) | M, \theta \rangle = \frac{1}{4\pi r}(1 + O(g^2)). \tag{3.40}$$

As expected, the fermion condensate is non-zero and contains no suppression factors. The fermion number is strongly non-conserved.

Let us make two remarks concerning the above calculation. First, the non-zero limit (3.39) comes from the cancellation of the fermionic factor in (3.37) decreasing like $(t_1 - t_2)^{-2}$ and the bosonic factor growing like $(t_1 - t_2)^2$. In models with a larger number of fermionic flavours, the analogous cancellation occurs only when the selection rule (1.12) is satisfied. Indeed, consider an SU(2) model with an even number (for the consistency of the model) of left-handed doublets. It is straightforward to calculate the two-point functions of gauge-invariant operators containing an even number, k, of fermionic fields ($k \leqslant n_f$)

$$\begin{aligned} \mathscr{F}_k(r_1, t_1; r_2, t_2) &= \langle J_k(r_1, t_1) J_k^+(r_2, t_2) \rangle \\ J_k(r, t) &= (\chi^{(1)} \ldots \chi^{(k)})(r, t). \end{aligned} \tag{3.41}$$

At large time, the corresponding bosonic factor increases like $(t_1 - t_2)^{k^2/n_f}$ while the fermionic factor decreases like $(t_1 - t_2)^k$. Therefore, $\mathscr{F}_k$ tends to zero as $|t_1 - t_2| \to \infty$ unless $k = n_f$, i.e. unless all fermionic fields enter the product in (3.41). By the cluster argument one finds that the minimal non-zero condensate violating the fermion number is

$$\langle M, \theta | \chi^{(1)} \ldots \chi^{(n_f)} | M, \theta \rangle. \tag{3.42}$$

This is the way the selection rule (1.11) comes about in the Green function approach.

Second, (3.38) implies that the main contributions to $\langle f \rangle$ and $\langle f^+ \rangle$ comes from the vicinities of the configurations determined by σ^- and σ^+ respectively. Making use of (3.18) one can calculate the winding numbers, $Q[A]$, of these configurations,

$$Q[A(\sigma_\pm)] = \pm 1.$$

Thus these configurations indeed describe the transitions between the distinct monopole states, $|M, 0\rangle$ and $|M, 1\rangle$. The winding number density of these configurations is concentrated at exponentially small distances from the monopole. So our approximation of vanishing r_M is valid, strictly speaking, only for large $r_{1,2}$, namely $n_f g^2 \ln r/r_M \gg 1$ (Craigie *et al* 1984, Schmid and Trugenberger 1984). Applied to fermion-monopole scattering, this observation means that the dynamics in the core region can be disregarded (as here) only if $n_f g^2 \ln M_v/E_f \gg 1$ where E_f is the incident fermion energy. The opposite situation, $n_f g^2 \ln M_V/E_f \ll 1$ has been studied by Schmid and Trugenberger (1984, 1986); their analysis is more complicated, but the main conclusions seem to remain intact. We think that further studies are required to make the dynamics in this case clear.

3.4. Other flavour-mixing Green functions

Apart from the fermion-number-violating Green functions, other flavour-changing Green functions are of interest for the monopole catalysis of proton decay. These appear entirely due to the interactions with the core and are not directly related to the triangle anomaly. Since the s-wave fermions interact with the core strongly, these flavour-mixing Green functions are unsuppressed, and the corresponding processes occur at high rates.

In the SU(2) model, the simplest flavour-mixing gauge-invariant condensate not related to the anomaly is (Kazama 1983)

$$\langle \bar{\chi}^{(1)}(1+\tilde{\gamma}^5)\chi^{(1)}\bar{\chi}^{(2)}(1-\tilde{\gamma}^5)\chi^{(2)}\rangle. \tag{3.43}$$

It can be rewritten in terms of the physical fields as follows:

$$\langle \bar{b}^{(1)}a^{(1)}\bar{a}^{(2)}b^{(2)}\rangle \tag{3.44}$$

and corresponds to the process (2.14).

The matrix element (3.43) conserves the fermion numbers, so it can be calculated over the perturbation theory state $|M, 0\rangle$. Making use of the general formula (3.26) we find that the bosonic factor is equal to 1, while the fermionic factor gives

$$\langle (\bar{b}^{(1)}a^{(1)}\bar{a}^{(2)}b^{(2)})(r)\rangle = \text{constant}/r^6 \tag{3.45}$$

where the numerical constant is irrelevant for our purposes. Equation (3.45) is valid in models with any number of fermionic flavours. We see that the flavour-mixing non-anomalous Green functions are indeed unsuppressed, contrary to the vacuum sector.

Let us now discuss briefly the problem of electric charge conservation in the Green function framework. Violation of the electric charge would be manifested by non-zero expectation values of charged operators. These matrix elements were analysed by Ezawa and Iwazaki (1983), Yoneya (1984), and Kazama and Sen (1984) who found that they are all equal to zero in the s-wave approximation. Thus, electric charge conservation is restored in quantum field theory, as expected. A more refined analysis of this problem was carried out by Balachandran and Schechter (1983, 1984), Yan (1983), Kazama and Sen (1984), and Polchinski (1984) who took into account the excitations of heavy bosonic fields inside the core. They also demonstrated electric charge conservation. The novel feature appearing naturally in the latter analysis is the possibility of virtual transitions of type (2.11) in which fermions deposit electric charge on the core for a short period of time (see also Schmid and Trugenberger 1984).

3.5. Beyond the s-wave approximation

The above analysis was based upon the s-wave approximation which was argued to be a reasonable description of the monopole–fermion interactions leading to the non-conservation of fermion quantum numbers. However, there are potentially dangerous corrections which might in principle give rise to the suppression of this effect.

There exist several types of corrections to the s-wave approximation. First, fermions with $J \neq 0$ change the effective action of the bosonic fields and might in principle suppress the contributions of relevant configurations into the functional integral. This effect was estimated for the saddle-point field determined by $\sigma^{\pm}$ (Rubakov 1982). It has been found that the correction to the effective action is finite and proportional to g^4. So this correction is negligible at small g, at least for the calculation of the fermion condensate.

Second, the interactions of s-wave fermions with higher partial wave fields change the s-wave fermion propagator. The corresponding lowest-order graph is shown in figure 4. In theories with massive fermions, this correction might change the helicity of the s-wave fermion and thus lead to its reflection from the monopole at relatively large distances. This effect would suppress the non-conservation of fermion quantum

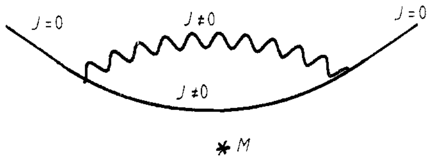

Figure 4. The lowest-order correction to the s-wave fermion propagator.

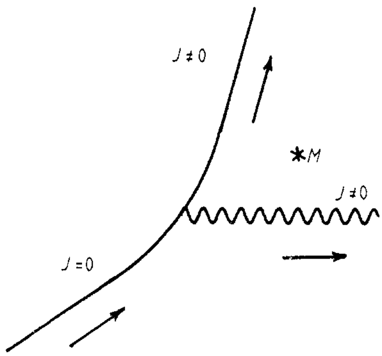

Figure 5. Radiative transition of an s-wave fermion into the $J \neq 0$ state.

numbers, since the latter occurs at short distances. The relevance of this correction was realised by this author (Rubakov 1982) and by Pak *et al* (1982) who pointed out that the anomalous magnetic moment of a fermion can have such an effect (see also Kazama *et al* 1977). However, it was understood by these authors that the correction of figure 4 cannot be approximated by the anomalous magnetic moment only, because the latter is merely the first term in the expansion of the fermion self-energy in powers of H/m_F^2, which is a large parameter at relevant distances. The graph of figure 4 was analysed by Kozakai and Minakata (1985) and Serebryakov (1985) who found that this correction is in fact negligible and does not suppress the processes of interest.

Third, an s-wave fermion can transform into a higher partial wave by emitting a real photon; see figure 5. This process, if it occurred at a high rate, would also lead to a suppression, since higher partial waves are expelled from the monopole by the centrifugal barrier. The analysis of this process was performed by Itoh and Kazama (1985, 1986) who showed that the cross section is finite and of order g^2. This means that this process is in fact irrelevant in weakly coupled theories. However, Itoh and Kazama (1985, 1986) claimed that the cross section contains large numerical factors, so this process could be of some importance in realistic theories. This point is not yet clear.

Thus, the existing results indicate that the corrections to the s-wave approximation are indeed small. However, the study of the whole set of the corrections has not been completed yet.

4. Bosonised SU(2) fermions

4.1. *Bosonised 'free' s-wave fermions*

In § 3 we have considered the fermion Green functions violating fermion quantum numbers. We have seen that these Green functions are large in the presence of a monopole, which indicates strongly that the corresponding processes occur at high rates. However, the Green function approach neither provides an explicit dynamical picture of the fermion–monopole scattering, nor is it convenient for calculating cross sections and branching ratios. Furthermore, it is hard to extend this approach to the case of massive fermions.

An alternative to the Green function approach to the s-wave fermion–monopole dynamics is the bosonisation technique suggested and developed by Callan (1982a, b, 1983). We begin the presentation of this technique with the simplest case of massless left-handed fermions in the s-wave state; furthermore, we disregard at the moment the bosonic fields a_α. In what follows we use the representation of the two-dimensional γ matrices where $\tilde{\gamma}^5$ is diagonal. We do not change the notation for χ since it will not lead to any confusion. We work in the Minkowski spacetime.

The key point of the bosonisation method is that the 'free' fermionic field $\chi(r, t)$ can be expressed in terms of the bosonic field $\varphi(r, t)$ as follows:

$$\chi(r, t) = :\exp\left[\mathrm{i}\sqrt{\pi}\left(\varphi(r, t) - \tilde{\gamma}^5 \int_0^r \mathrm{d}r'\, \partial_t \varphi(r', t)\right)\right]: \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} \tag{4.1}$$

where we have omitted some algebraic factors (Halpern 1975, Serebryakov 1986) not essential for our purposes. The action for the field φ is free,

$$S_\varphi^{\text{free}} = \int \mathrm{d}r\, \mathrm{d}t[\tfrac{1}{2}(\partial_t\varphi)^2 - \tfrac{1}{2}(\partial_r\varphi)^2].$$

The boundary condition at the origin is

$$\partial_r \varphi = 0 \qquad \text{at } r = 0. \tag{4.2}$$

One can check that the expression on the right-hand side of (4.1) has indeed the same Green functions as free fermions.

The fermionic currents are expressed locally in terms of the scalar field. In particular, the axial current which interacts with a_α in the full s-wave theory has the following representation:

$$j^{\alpha(5)} \equiv \bar{\chi}\tilde{\gamma}^\alpha\tilde{\gamma}^5\chi = (1/\sqrt{\pi})\varepsilon^{\alpha\beta}\partial_\beta\varphi. \tag{4.3}$$

We also recall that $\tilde{\gamma}^5$ coincides with the electric charge matrix, so that

$$n_\pm = \bar{\chi}\tilde{\gamma}^0\tfrac{1}{2}(1 \pm \tilde{\gamma}^5)\chi$$

are the number densities of the positive and negative charge s-wave fermions (a and b). They can be also expressed through the bosonic field,

$$n_\pm = (1/2\sqrt{\pi})(\pm\partial_r\varphi + \pi_\varphi) \tag{4.4}$$

where $\pi_\varphi \equiv \partial_t\varphi$ is the canonical momentum conjugate to φ. The numbers of positive and negative charge fermions inside a spherical shell $r_1 \leq r \leq r_2$ are

$$N_\pm(r_1, r_2) = \int_{r_1}^{r_2} n_\pm(r)\, \mathrm{d}r = \frac{1}{2\sqrt{\pi}} \int_{r_1}^{r_2} \mathrm{d}r(\pm\partial_r\varphi + \pi_\varphi).$$

We can now identify excitations carrying non-zero fermion numbers. A general solution to the free-field equation

$$\partial^2\varphi = 0 \tag{4.5}$$

is

$$\varphi(r,t)=\varphi_+(r+t)+\varphi_-(r-t)$$

where $\varphi_\pm$ are arbitrary functions. The fermion number densities for this solution are

$$n_\pm=\pm(1/\sqrt{\pi})\partial_r\varphi_\pm.$$

Thus, the fermionic and anti-fermionic excitations are described by the scalar field configurations as shown in figure 6. Note that adding a constant to φ changes neither the field equation (4.5) nor the expressions for fermion number densities (4.4). Therefore, any configuration shown in figure 6 can be shifted in the vertical direction.

These assignments are perfectly consistent with table 1 of § 2. The boundary condition (4.2) transforms an incoming wave a into an outgoing wave b; see figure 7. This is the electric-charge-violating process that occurs in quantum mechanics as discussed in § 2. In bosonic language, the charge non-conservation can be understood as follows. The total electric charge of fermions is obtained from (4.4),

$$Q=\int_0^\infty (n_+-n_-)\,\mathrm{d}r=(1/\sqrt{\pi})[\varphi(r=\infty)-\varphi(r=0)]. \tag{4.6}$$

Now, $\varphi(r=\infty)$ is trivially conserved, while $\varphi(r=0)$ changes in the scattering process of figure 7. One is tempted to interpret $(1/\sqrt{\pi})\varphi(r=0)$ as the electric charge at the monopole core; we find that if one switches off the gauge interactions (as we do here),

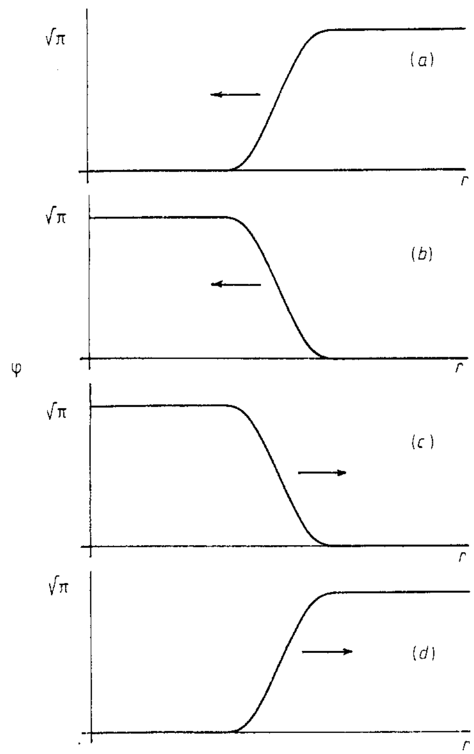

Figure 6. Fermionic excitations in the free bosonised theory ((a) a; (b) $\bar{a}$; (c) b; (d) $\bar{b}$). As before, a and b are positive and negative charge left-handed fermions, $\bar{a}$ and $\bar{b}$ are their antiparticles. The arrows indicate the direction of motion.

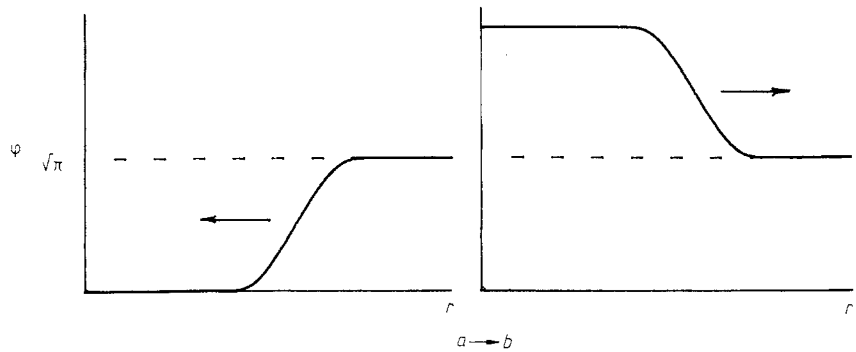

Figure 7. A charge-violating process in the free bosonised theory.

then charge exchange between fermions and the core is possible. Clearly, turning on the gauge interactions will lead to the complete suppression of this charge exchange, since the Coulomb energy of the electric charge at the core is very large (infinite in the limit $r_M \to 0$ we are working with). This point will be made explicit in § 4.2.

The bosonised version of the model contains asymptotic states with fractional fermion numbers absent in the original model. These states are scalar field configurations with

$$\varphi(r=\infty)-\varphi(r=0)\neq \text{integer}\times\sqrt{\pi}. \tag{4.7}$$

These states are required for the consistent description of the s-wave fermion–monopole scattering in the interacting massless theory: as pointed out in § 2, the conservation laws in some models allow no final states with integer fermion number for certain initial states.

These additional states are absent when fermions are massive. The case of Dirac masses is most important for our purposes. A massive fermionic doublet has both left-handed and right-handed components,

$$\psi_{L,R}=\begin{pmatrix} a \\ b \end{pmatrix}_{L,R}.$$

Accordingly, the s-wave fermions can be bosonised by introducing two bosonic fields, $\varphi_L(r,t)$ and $\varphi_R(r,t)$. However, it is inconvenient to express the mass term through these fields. Instead, it is useful to make a canonical transformation,

$$\begin{aligned}
\phi_a &= \tfrac{1}{2}(\varphi_L-\varphi_R)+\tfrac{1}{2}\int_0^r dr'(\pi_L+\pi_R)\\
\phi_b &= -\tfrac{1}{2}(\varphi_L-\varphi_R)+\tfrac{1}{2}\int_0^r dr'(\pi_L+\pi_R)\\
\Pi_a &= \tfrac{1}{2}(\pi_L-\pi_R)+\tfrac{1}{2}\partial_r(\varphi_L+\varphi_R)\\
\Pi_b &= -\tfrac{1}{2}(\pi_L-\pi_R)+\tfrac{1}{2}\partial_r(\varphi_L+\varphi_R).
\end{aligned} \tag{4.8}$$

The massless part of the free Hamiltonian does not change under this transformation,

$$H^{\text{free}}_{m=0}=\int dr\,\tfrac{1}{2}[\Pi_a^2+\Pi_b^2+(\partial_r\phi_a)^2+(\partial_r\phi_b)^2] \tag{4.9}$$

while the mass term takes a simple form,

$$H_m = \int \mathrm{d}r\, \tfrac{1}{2}\mu m[(1-\cos 2\sqrt{\pi}\phi_a)+(1-\cos 2\sqrt{\pi}\phi_b)]. \tag{4.10}$$

Here μ is an arbitrary normalisation point and the cosine terms are assumed to be normal ordered with respect to μ.

The definition (4.8) and the boundary condition (4.2) imply the following boundary conditions:

$$\begin{aligned} &\phi_a = -\phi_b \\ &\partial_r\phi_a = \partial_r\phi_b \qquad \text{at } r=0. \end{aligned} \tag{4.11}$$

The number densities of fermions a and b are conveniently expressed through $\phi_{a,b}$,

$$n_{a,b} = n_{\pm,\mathrm{L}} + n_{\pm,\mathrm{R}} = (1/\sqrt{\pi})\partial_r\phi_{a,b}$$

where we have made use of (4.4) and its analogue for right-handed fermions.

The resulting bosonic theory with the Hamiltonian $(H^{\mathrm{free}}_{m=0}+H_m)$ coincides with the direct sum of two sine-Gordon theories (up to the boundary condition (4.11)). The original fermions are represented by the sine-Gordon solitons. These solitons are the only asymptotic states of the massive theory, except mesons (zero fermion number states which can be interpreted as fermion–antifermion pairs moving together). The fractional fermion number states are absent in the massive theory as they have infinite energies.

It is worth noting that the fields $\phi_{a,b}$ can be shifted by $[\pm\sqrt{\pi}(\text{integer})]$ without changing the physical situation and the boundary condition, so the 'free' massive theory has an infinite series of vacua,

$$\phi_a = \sqrt{\pi}\,\nu \qquad \phi_b = -\sqrt{\pi}\,\nu$$

where ν is an arbitrary integer.

4.2. Interacting theory: the Hamiltonian and ground states

We now turn on the interaction of fermions with the gauge field a_α. It is convenient to use the Coulomb gauge,

$$a_1 = 0.$$

In this gauge, the free gauge field action, equation (3.11), is

$$S_a = \frac{4\pi}{g^2}\int r^2(\partial_r a_0)^2\,\mathrm{d}r\,\mathrm{d}t$$

so the field a_0 is not a dynamical variable and should be eliminated by using Gauss' law.

As discussed in § 3.2, s-wave fermion fields χ interact with a_α in an axial vector manner. The corresponding current in the bosonised formalism is given by (4.3) for left-handed fermions. For right-handed fermions one has

$$j^{\alpha(5)} = -(1/\sqrt{\pi})\varepsilon^{\alpha\beta}\partial_\beta\varphi_{\mathrm{R}}.$$

Thus, the interaction term in the SU(2) model with n_{L} left-handed and n_{R} right-handed doublets is

$$S_{\mathrm{int}} = \frac{1}{\sqrt{\pi}}\int \mathrm{d}r\,\mathrm{d}t\,\partial_r\left(\sum_{s=1}^{n_{\mathrm{L}}}\varphi_{\mathrm{L}}^{(s)} - \sum_{v=1}^{n_{\mathrm{R}}}\varphi_{\mathrm{R}}^{(v)}\right)a_0.$$

Gauss' law is

$$\frac{\delta}{\delta a_0}(S_a+S_{\rm int})\equiv\frac{8\pi^2}{g^2}r^2\partial_r a_0-\frac{1}{\sqrt{\pi}}\Big(\sum\varphi_{\rm L}-\sum\varphi_{\rm R}\Big)=0. \tag{4.12}$$

Expressing a_0 via $\varphi_{\rm L,R}$ we find the following Coulomb term in the scalar field action:

$$S_{\rm Coul}=-\frac{g^2}{16\pi^2}\int {\rm d}r\,{\rm d}t\frac{1}{r^2}\Big(\sum\varphi_{\rm L}-\sum\varphi_{\rm R}\Big)^2. \tag{4.13}$$

In the massless case, the total bosonised action,

$$S_{m=0}=\sum_s S^{\rm free}_{\varphi_{\rm L}^{(s)}}+\sum_v S^{\rm free}_{\varphi_{\rm R}^{(v)}}+S_{\rm Coul} \tag{4.14}$$

is quadratic in the scalar fields, so the model is exactly solvable. In fact, this solution is the same as the functional integral one of § 3. Indeed, the fields $\varphi_{\rm L,R}$ are expressed through massless free scalar fields and a field Σ defined by

$$\Sigma=\frac{1}{(n_{\rm L}+n_{\rm R})^{1/2}}\Big(\sum\varphi_{\rm L}-\sum\varphi_{\rm R}\Big). \tag{4.15}$$

The latter is just the same field as that defined by (3.30): it is straightforward to see that the field equation for Σ coincides exactly with (3.31), while (4.12) tells us that the radial electric field, $\mathscr{E}_r\equiv(2/g)\partial_r a_0$ is related to Σ precisely by (3.32). One can also show that the gauge-invariant Green functions coincide for the bosonic and functional integral solutions.

The Coulomb term in the action, (4.14), corresponds to the following term in the Hamiltonian:

$$H_{\rm Coul}=\frac{g^2}{16\pi^2}\int {\rm d}r\frac{1}{r^2}\Big(\sum\varphi_{\rm L}-\sum\varphi_{\rm R}\Big)^2. \tag{4.16}$$

This term can be interpreted as the Coulomb energy,

$$H_{\rm Coul}=\tfrac{1}{2}\int {\rm d}^3x\,\mathscr{E}_r^2. \tag{4.17}$$

The equivalence between equations (4.16) and (4.17) is established by (3.31) and (4.15). We note also that equations (3.31) and (4.15) imply that

$$Q(r)=(1/\sqrt{\pi})\Big(\sum\varphi_{\rm L}-\sum\varphi_{\rm R}\Big)(r) \tag{4.18}$$

should be interpreted as the electric charge (in units $\frac{1}{2}g$) contained inside a sphere of radius r. Equation (4.18) is a proper generalisation of equation (4.6). The important difference between the interacting and free theories is that in the former one has

$$\sum\varphi_{\rm L}-\sum\varphi_{\rm R}=0\qquad \text{at } r=0$$

otherwise the Coulomb energy would be infinite. Thus, in the interaction theory, the charge exchange between fermions and the monopole core is totally suppressed, as expected. The electric charge of fermions is conserved in the full theory.

We now turn to the case of massive fermions. Performing the canonical transformation, equation (4.8), we find that the Coulomb term becomes

$$H_{\rm Coul}=\int {\rm d}r\frac{g^2}{16\pi^2r^2}\Big(\sum(\phi_a-\phi_b)\Big)^2 \tag{4.19}$$

where the sum runs over fermion doublets. Thus, we arrive at the following prescription for constructing the s-wave bosonised Hamiltonian. For each Dirac fermion (say, f) one introduces a scalar field ϕ_f and its conjugate momentum Π_f. The Hamiltonian is

$$H=\sum_f (H^{\text{free}}_{m=0}+H_m)+H_{\text{Coul}}$$

where

$$H^{\text{free}}_{m=0}=\int \mathrm{d}r\,\tfrac{1}{2}[\Pi_f^2+(\partial_r\phi_f)^2] \tag{4.20}$$

$$H_m=\int \mathrm{d}r\,\tfrac{1}{2}\mu m_f[1-\cos(2\sqrt{\pi}\phi_f)] \tag{4.21}$$

$$H_{\text{Coul}}=\int \mathrm{d}r\frac{g^2}{16\pi^2r^2}\left(\sum_f q_f\phi_f\right)^2 \tag{4.22}$$

where q_f is the electric charge of fermion f in units $\frac{1}{2}g$. The scalar fields corresponding to different members of the SU(2) doublet are related at $r=0$ by the boundary conditions (4.11). These boundary conditions reflect the structure of the monopole core.

The massive s-wave model is not exactly solvable. However, in the bosonised theory, a semiclassical description of the dynamics is possible, which provides a qualitative understanding of the s-wave fermion–monopole interactions. We first determine the semiclassical ground states of the bosonised theory, i.e. the configurations having zero energies. It is clear from (4.20) that the ground-state configurations are constant in space and time. Equations (4.11), (4.19) and (4.21) tell us that these constants are

$$\phi^{\text{vac}}_{a_i}=-\phi^{\text{vac}}_{b_i}=\sqrt{\pi}\nu_i \tag{4.23}$$

where $\{\nu_i\}$ is any set of integers obeying

$$\sum_i \nu_i=0$$

where the sum runs over all fermionic doublets. Thus there exists an infinite set of degenerate ground states labelled by $(n_{\mathrm{f}}-1)$ integers where n_{f} is the total number of the Dirac fermionic doublets. Asymptotic fermion states in the bosonised theory are the sine-Gordon solitons discussed in § 4.1 (the Coulomb term is negligible far away from the monopole). These can be constructed above any of the vacua (4.23).

We have assumed in the above discussion of the ground states that the θ angle of the theory is zero. The theory at non-zero θ was studied by Callan (1982b) who found that the ground-state configurations are inhomogeneous and that their electric charge is non-zero and is given by the formula of Witten (1979), $Q=-g\theta/2\pi$. Callan (1982b) considered also a model where fermions acquire masses due to the interactions with the triplet Higgs field. He observed that the ground states carry half-integer fermion numbers, in agreement with the earlier analysis by Jackiw and Rebbi (1976b).

In the bosonised formalism, one easily understands the decoupling of heavy fermionic flavours: heavy modes are merely not excited. On the other hand, masses of light fermions are expected to be irrelevant at short distances from the monopole, since the Coulomb terms dominate over the mass terms at small r. However, the latter fact is not at all trivial if quantum effects are properly taken into account (Virasoro 1983a), especially in models where fermionic masses are due to the Yukawa coupling to some Higgs field (as in the SU(5) GUT). This problem was analysed by Bennett (1985) who found that in realistic theories, the expectation value of the Higgs field (and, correspondingly, the fermion masses) grows at short distances, but this fact has little effect on the fermion scattering off the monopole because of the small values of the Yukawa coupling constants.

4.3. Flavour non-conservation

Let us discuss, within the bosonised formalism, the scattering processes of the s-wave fermions off the monopole. We consider for simplicity the massless theory in which the bosonic action is quadratic in the scalar fields. The model is exactly solvable; furthermore, a semiclassical analysis gives exact results.

We begin with the simplest model containing two left-handed fermionic doublets. In the bosonised formalism, the s-wave sector is described by two scalar fields $\varphi^{(1)}$ and $\varphi^{(2)}$ (we omit the subscript L). Far away from the monopole, the Coulomb term is negligible, so the model coincides with the 'free' theory of § 4.1. In particular, the asymptotic fermion states are described by lumps as shown in figure 6.

It is convenient to perform a canonical transformation to the fields

$$\Sigma = (1/\sqrt{2})(\varphi^{(1)} + \varphi^{(2)})$$

$$f = (1/\sqrt{2})(\varphi^{(1)} - \varphi^{(2)}).$$

The field Σ was introduced earlier (see (4.15)); it obeys (3.31) with $n_f = 2$. The field f obeys the free equation

$$\partial^2 f = 0 \tag{4.24}$$

and the boundary condition

$$\partial_r f = 0 \qquad \text{at } r = 0 \tag{4.25}$$

(the latter is a consequence of (4.2)). We can simplify the analysis further. The non-singular solution to (3.31) with fixed energy is

$$\Sigma_E(r, t) = \exp(-\mathrm{i}Et)(\tfrac{1}{2}\pi Er)^{1/2} J_\nu(Er) \tag{4.26}$$

where J_ν is the Bessel function, $\nu = (\frac{1}{4} + g^2/4\pi)^{1/2}$. At large r equation (4.26) is

$$\Sigma_E(r, t) = \exp(-\mathrm{i}Et) \sin(Er - \tfrac{1}{2}d\pi)$$

where d is given by (3.29), $d = \mathrm{O}(g^2)$. Thus, as far as large distances are concerned and $\mathrm{O}(g^2)$ effects are neglected, the field Σ can be considered effectively as the free field obeying

$$\partial^2 \Sigma = 0 \tag{4.27}$$

$$\Sigma(r = 0) = 0 \tag{4.28}$$

i.e. the only effect of the Coulomb term on scattering is the appearance of the boundary condition (4.28) instead of (4.2). Note, however, that the complete solution, equation (4.26), obeys both the boundary condition (4.28) *and* (4.2).

The solution to (4.24) and (4.25) with the boundary conditions (4.27) and (4.28) is easily found for any initial data (incoming states). For example, for the initial state containing $a^{(1)}$ one finds the final state $\bar{b}^{(2)}$; see figure 8. This is just the fermion-number-violating process (2.12).

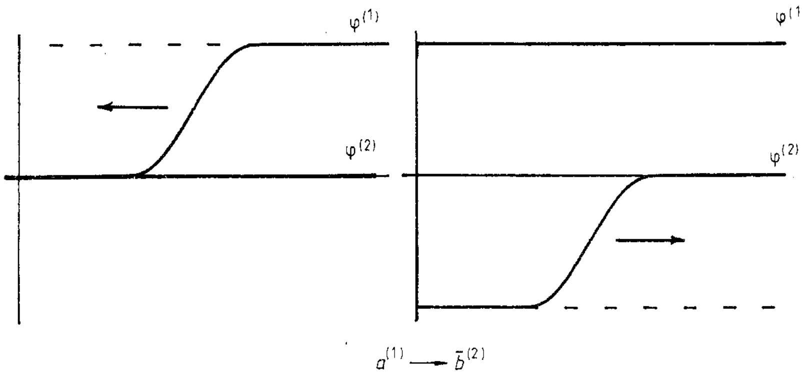

Figure 8. The reaction $a^{(1)}+\text{mon}\rightarrow \bar{b}^{(2)}+\text{mon}$ in the bosonised theory. As before, arrows indicate the direction of motion.

A new phenomenon occurs in a model with four left-handed fermionic doublets. The bosonised version can be studied in the same way as above. For an incoming $a^{(1)}$ one finds the final state as shown in figure 9. In the final state, the scalar fields carry half-integer fermion numbers. The schematic expression for this reaction is (Sen 1983, Callan 1984)

$$a^{(1)}+\text{mon}\rightarrow \tfrac{1}{2}b^{(1)}+\tfrac{1}{2}(\bar{b}^{(2)}+\bar{b}^{(3)}+\bar{b}^{(4)})+\text{mon}. \tag{4.29}$$

This is the way the puzzle mentioned at the end of § 2.2 (no outgoing state with integer fermion numbers for an incoming $a^{(1)}$) is solved; note that no conservation laws are violated in this reaction. As discussed in § 4.1, the asymptotic states with fractional fermion numbers are peculiar to the massless theory; they disappear when the mass terms are included. In the massive theory, the reaction (4.29) is an intermediate step at $r \ll m_{\mathrm{F}}^{-1}$; at larger radii, the half-fermions transform either into solitons (fermions) or mesons (zero-fermion-number excitations). The final states in the classical bosonic theory for various mass patterns were calculated numerically by Dawson and Schellekens (1983b). In the quantum theory, equation (4.29) must be given a probabilistic interpretation; we return to this point in § 5.1.

In the massless theory, both the field equations and boundary conditions are linear. Therefore, all reactions can be found by joining together a few basic ones. The latter

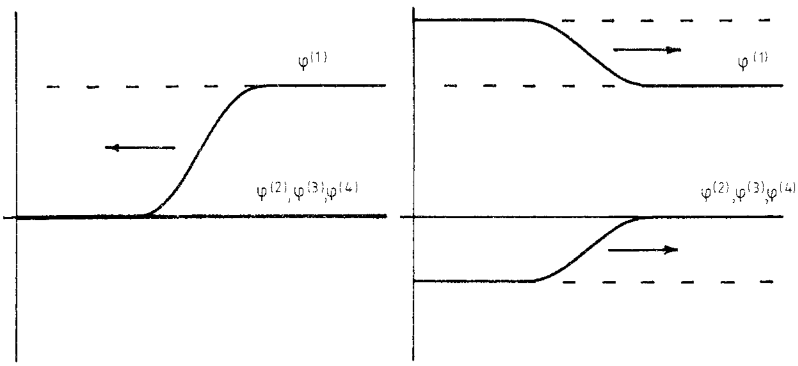

Figure 9. The fate of $a^{(1)}$ in the model with four massless left-handed doublets.

can be chosen as the reaction (4.29), its conjugate,

$$\bar{a}^{(1)}+\text{mon}\rightarrow\tfrac{1}{2}\bar{b}^{(1)}+\tfrac{1}{2}(b^{(2)}+b^{(3)}+b^{(4)})+\text{mon} \tag{4.30}$$

and those processes which can be obtained from (4.29) and (4.30) by permuting the flavour indices. In this way one recovers the reactions (2.14) and (2.15).

Thus the bosonised version of the theory is convenient for studying the scattering processes in the s-wave sector of the fermion-monopole theory. It incorporates naturally all conservation laws, including the selection rule (1.11), and is perfectly consistent with the Green function approach. Properly generalised, the bosonised formalism is useful for investigating the monopole catalysis of proton decay in realistic GUT.

5. Catalysis of proton decay by a fundamental SU(5) monopole

5.1. $SU(2)_M$ approximation

We are now in a position to consider fermion-monopole interactions in realistic GUT. As the simplest example we discuss the fundamental SU(5) monopole whose classical structure is described in § 1.5. To simplify our discussion as much as possible, we first switch off all grand unified interactions except those of the $SU(2)_M$ subgroup defined by (1.20). This $SU(2)_M$ approximation is, in fact, an oversimplification even in the s-wave sector; however, we shall see in §§ 5.2 and 5.3 that the basic results will not be altered when other relevant interactions are turned on.

In the $SU(2)_M$ approximation, the original GUT is reduced to the SU(2) model of §§ 2-4, and the fundamental monopole is just the minimal monopole of SU(2). With respect to the $SU(2)_M$, the fermions of the first generation form four left-handed doublets,

$$\begin{pmatrix}-\bar{\mathrm{u}}_2\\ \mathrm{u}^1\end{pmatrix}_{\mathrm{L}} \qquad \begin{pmatrix}\bar{\mathrm{u}}_1\\ \mathrm{u}^2\end{pmatrix}_{\mathrm{L}} \tag{5.1}$$

$$\begin{pmatrix}\mathrm{d}^3\\ \mathrm{e}^+\end{pmatrix}_{\mathrm{L}} \qquad \begin{pmatrix}\mathrm{e}^-\\ -\bar{\mathrm{d}}_3\end{pmatrix}_{\mathrm{L}} \tag{5.2}$$

where 1, 2, 3 are colour indices and a bar denotes an antiparticle. Other fermions of the first generation are $SU(2)_M$ singlets; they do not interact with the monopole magnetic and chromomagnetic fields and are of no interest for our purposes.

If we neglect higher generations, we can immediately translate the results of the SU(2) model with four left-handed doublets summarised in §§ 3 and 4. In particular, the fermion condensates appearing due to the vacuum structure and anomaly are

$$\langle(\bar{\mathrm{u}}_{1\mathrm{L}}\bar{\mathrm{u}}_{2\mathrm{L}}\bar{\mathrm{d}}_{3\mathrm{L}}\mathrm{e}^{+}_{\mathrm{L}})(r)\rangle \tag{5.3}$$

$$\langle(\bar{\mathrm{u}}_{1\mathrm{L}}\mathrm{u}^{1}_{\mathrm{L}}+\bar{\mathrm{u}}_{2\mathrm{L}}\mathrm{u}^{2}_{\mathrm{L}})\bar{\mathrm{d}}^{3}_{\mathrm{L}}\mathrm{d}_{3\mathrm{L}}(r)\rangle \tag{5.4}$$

etc. In the s-wave approximation, these are calculated as described in § 3.3. One finds that all of them are equal to constant/r^6, where the constant is of the order of one (cf (3.34) and (3.40)). It is worth noting that these condensates are not invariant under the colour group $SU(3)_c$; this is a reflection of the fact that the global colour is broken in the presence of a monopole (Abouelsaood 1983a, b, Nelson and Manohar 1983, Balachandran *et al* 1983, Nelson and Coleman 1984).

The quantum numbers of the condensate (5.3) correspond to the nucleon decay process,

$$\mathrm{p} + \mathrm{mon} \to \mathrm{e}^{+} + \mathrm{mon}\ (+\mathrm{pions}) \tag{5.5}$$

$$\mathrm{n} + \mathrm{mon} \to \mathrm{e}^{+} + \pi^{-} + \mathrm{mon}\ (+\mathrm{pions}). \tag{5.6}$$

The fact that this condensate is unsuppressed implies that these processes occur at high rates; since the only relevant energy scale is that of strong interactions (~1 GeV), the cross sections of the monopole catalysis reactions (5.5) and (5.6) are those typical to strong interactions (Rubakov 1981a, b, 1982, Callan 1982b, 1983).

The condensates not directly related to the anomaly violate baryon number as well. These are the translations of (3.45) and similar equations obtained from (3.45) by the permutations of flavour indices. For example, one has

$$\langle(\bar{\mathrm{u}}_{1\mathrm{L}}\bar{\mathrm{u}}_{2\mathrm{R}}\bar{\mathrm{d}}_{3\mathrm{R}}\mathrm{e}_{\mathrm{L}}^{+})(r)\rangle = \text{constant}/r^{6} \tag{5.7}$$

where the constant is again of the order of one. This condensate also corresponds to the processes (5.5) and (5.6).

In the $\mathrm{SU}(2)_M$ approximation, the baryon-number non-conservation at the quark level is seen also in the bosonised formalism of § 4. In particular, among the reactions of the type (4.29) one finds the following process:

$$\mathrm{d}_{\mathrm{L}}^{3} + \mathrm{mon} \to \tfrac{1}{2}\mathrm{e}_{\mathrm{L}}^{+} + \tfrac{1}{2}(\bar{\mathrm{u}}_{1\mathrm{R}} + \bar{\mathrm{u}}_{2\mathrm{R}}) + \tfrac{1}{2}\mathrm{d}_{3\mathrm{R}} + \mathrm{mon}. \tag{5.8}$$

The quantum numbers of this process correspond to the reaction (Sen 1983, Callan 1984)

$$\mathrm{p} + \mathrm{mon} \to \tfrac{1}{2}\mathrm{p} + \tfrac{1}{2}\mathrm{e}^{+} + \mathrm{mon}. \tag{5.9}$$

The wavefunction of the final state can be interpreted as a linear superposition of a proton and a positron, so that equation (5.9) describes effectively two types of processes occurring, roughly speaking, at equal rates, namely, the baryon-number-violating reaction (5.5) and a baryon-number-conserving one.

Among the reactions of the type (4.30) there are also baryon-number-violating ones, e.g.,

$$\mathrm{u}_{\mathrm{R}}^{1} + \mathrm{mon} \to \tfrac{1}{2}\mathrm{u}_{\mathrm{L}}^{1} + \tfrac{1}{2}\bar{\mathrm{u}}_{2\mathrm{R}} + \tfrac{1}{2}\bar{\mathrm{d}}_{3\mathrm{L}} + \tfrac{1}{2}\mathrm{e}_{\mathrm{L}}^{+} + \mathrm{mon}. \tag{5.10}$$

Its quantum numbers coincide with those of (5.9). Other examples of baryon-number-violating processes are translations of (2.14) and (2.15),

$$\mathrm{d}_{\mathrm{L}}^{3} + \mathrm{u}_{\mathrm{R}}^{1} + \mathrm{mon} \to \mathrm{e}_{\mathrm{L}}^{+} + \bar{\mathrm{u}}_{2\mathrm{R}} + \mathrm{mon} \tag{5.11}$$

$$\mathrm{u}_{\mathrm{R}}^{1} + \mathrm{u}_{\mathrm{R}}^{2} + \mathrm{mon} \to \mathrm{e}_{\mathrm{L}}^{+} + \bar{\mathrm{d}}_{3\mathrm{L}} + \mathrm{mon}. \tag{5.12}$$

The former reaction is not directly related to the anomaly (this type of process was first discussed, within the bosonised formalism, by Seo (1983)), while the latter is anomalous. In the scattering picture, it is even more clear that the baryon-number-violating processes occur at high rates: recall that the reaction (5.8) is the only one allowed by the conservation laws for an initial $\mathrm{d}_{\mathrm{L}}^{3}$ (in the limit of massless fermions).

The realistic SU(5) theory also contains fermions of higher generations. Among these, muons and s-quarks have masses comparable to the typical hadronic energy scale, so they should be taken into account in the discussion of the monopole–fermion interactions at distances of order of 1 fm or less. In an analogy to equation (5.2), muons and s-quarks form two left-handed $\mathrm{SU}(2)_M$ doublets,

$$\begin{pmatrix} \mathrm{s}^{3} \\ \mu^{+} \end{pmatrix}_{\mathrm{L}} \qquad \begin{pmatrix} \mu^{-} \\ -\bar{\mathrm{s}}_{3} \end{pmatrix}_{\mathrm{L}}$$

while $s^{1,2}$ are $SU(2)_M$ singlets. If these doublets were massless, the previous results would be modified, since the $SU(2)_M$ model would contain six doublets instead of four. The modifications would be as follows.

(i) Instead of the anomalous condensates (5.3) and (5.4), six-fermion condensates would emerge (cf (3.42)), which would include

$$\langle \bar{u}_{1L}\bar{u}_{2L}\bar{d}_{3L}e^{+}_{L}s^{3}_{L}\bar{s}_{3L}\rangle \tag{5.13}$$

$$\langle \bar{u}_{1L}\bar{u}_{2L}\bar{d}_{3L}e^{+}_{L}\mu^{+}_{L}\mu^{-}_{L}\rangle \tag{5.14}$$

$$\langle \bar{u}_{1L}\bar{u}_{2L}\bar{d}_{3L}d^{3}_{L}\mu^{+}_{L}\bar{s}_{3L}\rangle. \tag{5.15}$$

(ii) Together with the condensate (5.7), other non-anomalous four-fermion condensates would be formed, e.g.,

$$\langle \bar{u}_{1L}\bar{u}_{2R}\bar{s}_{3L}\mu^{+}_{R}\rangle. \tag{5.16}$$

The condensate (5.14) corresponds to the process (Rubakov 1981a, b)

$$p + \text{mon} \to e^{+} + \mu^{+}\mu^{-} + \text{mon} \tag{5.17}$$

while the condensates (5.15) and (5.16) have quantum numbers of the reactions (Bais *et al* 1983, Rubakov and Serebryakov 1983)

$$\begin{aligned} &p + \text{mon} \to \mu^{+} + K^{0} + \text{mon} \\ &p + \text{mon} \to \mu^{+} + K^{+} + \pi^{-} + \text{mon}. \end{aligned} \tag{5.18}$$

One expects the amplitudes of the processes (5.17) and (5.18) to be roughly of the same order of magnitude as the amplitude of the reaction (5.5).

In reality, muons and s-quarks have non-negligible masses. Furthermore, $\langle \bar{s}_L s_L\rangle$ is non-zero in the QCD vacuum (Gell-Mann *et al* 1968, Scadron 1981). This implies, in particular, that both the condensate (5.13) and (5.3) are non-zero. On the other hand, the muon and s-quark masses are not too large, so one expects the condensates (5.13)–(5.16) not to be suppressed considerably, i.e. the processes (5.17) and (5.18) to occur at roughly the same rates as the reaction (5.5). In particular, the process (5.17) is not a radiative correction to (5.5).

Thus, in the realistic theory, a large number of channels for the monopole catalysis of proton decay is expected, at least in the $SU(2)_M$ approximation. An important problem is the calculation of branching ratios. Presently, this problem has not been solved even at the quark level.

5.2. From $SU(2)_M$ to $SU(2)_M \times SU(2)_c \times U(1)$: $1/N_c$ expansion

Even in the s-wave sector, the $SU(2)_M$ fields do not exhaust all relevant bosonic degrees of freedom. If we neglect, as before, heavy gauge and Higgs excitations, we still have to count all massless s-wave fields. Except the gauge field corresponding to $t^3_M = \frac{1}{2}(0, 0, 1, -1, 0)$, there are gluon and photon fields which form singlets and doublets under $SU(2)_M$. Among these, only $SU(2)_M$ singlets contain s-wave components, according to equation (2.2). The singlet gauge fields correspond to the generator

$$Q' = (1/2\sqrt{2})\,\mathrm{diag}(1, 1, -1, -1, 0) \tag{5.19}$$

(a linear combination of the electric charge and colour hypercharge) and to the generators of an $SU(2)_c$ subgroup of $SU(3)_c$,

$$SU(2)_c = \begin{pmatrix} SU(2) & 0 \\ 0 & \mathbb{1}_1 \end{pmatrix} \subset SU(3)_c.$$

Thus the subgroup of SU(5) playing a role in the s-wave fermion-monopole dynamics is

$$\mathrm{SU}(2)_M \times \mathrm{SU}(2)_\mathrm{c} \times \mathrm{U}(1)'. \tag{5.20}$$

Since the generators of $\mathrm{SU}(2)_\mathrm{c} \times \mathrm{U}(1)'$ commute with $\mathrm{SU}(2)_M$, the s-wave parts of the corresponding gauge fields are just their radial components,

$$\boldsymbol{A}_{\mathrm{SU}(2)_\mathrm{c}} = -\mathrm{i}\boldsymbol{n}t_\mathrm{c}^a a_1^a(r,t) \qquad A^0_{\mathrm{SU}(2)_\mathrm{c}} = -\mathrm{i}t_\mathrm{c}^a a_0^a(r,t)$$
$$\boldsymbol{A}_{Q'} = -\mathrm{i}\boldsymbol{n}Q'a_1'(r,t) \qquad A^0_{Q'} = -\mathrm{i}Q'a_0'(r,t)$$

where $a=1,2,3$ refers to $\mathrm{SU}(2)_\mathrm{c}$, t^a are generators of $\mathrm{SU}(2)_\mathrm{c}$, and a_α^a and a_α' are radial gauge fields.

In analogy to § 3.2, one can write down the s-wave action in the limit of massless fermions,

$$S = \int \mathrm{d}r\,\mathrm{d}t \sum_{\text{gauge fields}} \left(-\frac{2\pi}{g^2}r^2\right) f_{\alpha\beta}^2 + \int \mathrm{d}r\,\mathrm{d}t \sum_{\text{fermions}} \mathrm{i}\bar{\chi}\tilde{\gamma}^\alpha D_\alpha \chi$$

where $f_{\alpha\beta}$ are the two-dimensional field strengths corresponding to a_α^a, a_α' and the radial gauge field a_α of t_M^3,

$$D_\alpha = \partial_\alpha + \mathrm{i}\tilde{\gamma}^5 a_\alpha - \mathrm{i}q'a_\alpha' - \mathrm{i}t_\mathrm{c}^a a_\alpha^a.$$

Making use of (5.19) one finds that $q'=0$ for the doublets (5.1) and $q'=-1/2\sqrt{2}$ and $q'=+1/2\sqrt{2}$ for the first and second doublets of (5.2), respectively. The action of t_c^a is determined by the fact that $\mathrm{SU}(2)_\mathrm{c}$ acts on the doublets (5.1) as the horizontal group, while d^3 and e are $\mathrm{SU}(2)_\mathrm{c}$ singlets.

To study the Euclidean Green functions, we notice first that the Abelian gauge fields a_α and a_α' factor out. Indeed, in analogy to (3.19) we can write

$$\chi = \exp(-\sigma - \mathrm{i}\tilde{\gamma}^5\gamma)\exp[q'(\tilde{\gamma}^5\sigma' + \gamma')]\chi_0 \tag{5.21}$$

where σ' and γ' are related to a_α' by a formula analogous to (3.17). An important difference between $\mathrm{U}(1)_M$ and $\mathrm{U}(1)'$ is that the boundary condition for σ' is

$$\sigma' = 0 \qquad \text{at } r=0. \tag{5.22}$$

Indeed, under this condition, the field χ_0 obeys the σ-independent boundary condition, equation (3.22). In terms of the field χ_0, the fermionic part of the action becomes (cf (3.20))

$$S_{\chi_0} = \int \mathrm{d}r\,\mathrm{d}t \sum_{\text{fermions}} \bar{\chi}_0 \mathrm{i}\tilde{\gamma}^\alpha(\partial_\alpha - \mathrm{i}t_\mathrm{c}^a a_\alpha^a)\chi_0.$$

Thus the model is reduced to a two-dimensional $\mathrm{SU}(2)_\mathrm{c}$ gauge theory with certain boundary conditions.

The Jacobian of the transformation is obtained as outlined in § 3.2,

$$J = \exp(-\Gamma_1)$$
$$\Gamma_1 = \sum_{\text{fermions}} \frac{1}{2\pi}\int \mathrm{d}r\,\mathrm{d}t\,(\sigma\partial^2\sigma + q'^2\sigma'\partial^2\sigma').$$

So the effective action for σ' has almost the same form as that for σ, equation (3.25). However, because of the different boundary condition, the σ' propagator

$$\mathcal{P}_{\sigma'} \propto \mathcal{D}(r-r', t-t') - \mathcal{D}(r+r', t-t') + \mathcal{R}$$

tends to zero at large $|t-t'|$, unlike the σ propagator which grows logarithmically in this limit. An important consequence is that the second exponential factor in (5.21) which contains σ' is irrelevant for the calculation of the gauge-invariant anomalous Green functions by the cluster argument: the two-point functions of these exponents tend to one at large time. Thus the U(1)$'$ interactions can be disregarded when evaluating the anomalous Green functions.

Formally, the difference between a_α and a'_α is that the former interacts with fermions χ in an axial-vector manner, while the latter interacts vectorially. A deeper reason for the different roles played by U(1) and U(1)$'$ is that the 'electric' field of the former is directed, in the group space, along the monopole 'magnetic' field, so it contributes to the winding number of the gauge field (see the discussion in § 3.1). On the other hand, the U(1)$'$ radial 'electric' field does not contribute to the anomaly.

Thus, in the model under discussion, the fermionic Green functions can be represented in a form analogous to (3.26). However, χ_0 are no longer free fields; rather, they interact with the non-Abelian fields a^a_α. This part of the theory is not exactly solvable. We sketch here a $1/N_c$ expansion approach to the radial $\mathrm{SU}(2)_c$ problem which was developed by Craigie *et al* (1984) in analogy with the analysis by 't Hooft (1974a) of the two-dimensional QCD. Clearly, the $1/N_c$ expansion is hard to justify at $N_c=2$; however, one hopes that the results are qualitatively correct. Furthermore, some results obtained in the leading order in $1/N_c$ turn out to be exact (Craigie and Nahm 1984a, b, Craigie 1986a, b). Another approach to the radial fermion–monopole interactions described by the gauge group (5.20) is based on the bosonisation technique; we present it in § 5.3.

To develop the $1/N_c$ expansion, one makes use of a bilocal field formalism suggested by Ebert and Pervushin (1978) in the context of the two-dimensional QCD. It is convenient to choose the temporal gauge, $a^a_0=0$, and integrate out the field a^a_1. In this gauge, the bosonic part of the action is quadratic in a^a_1, so the integration over a^a_1 leads to a non-local four-fermion interaction. The latter can be transformed into a non-local interaction of the Yukawa type at the expense of introducing bilocal fields $\Xi(r,t;r',t')$ and $\Theta^a(r,t;r',t')$ which transform as a singlet and triplet, respectively, under the global $\mathrm{SU}(2)_c$. Then one integrates out fermions and arrives at the following action for the bilocal fields:

$$S[\Xi,\Theta]=-\tfrac{1}{2}(\Xi,R^{-1}\Xi)-\tfrac{1}{2}(\Theta^a,R^{-1}\Theta^a)+\tfrac{1}{2}\mathrm{Tr}\ln[(G_0^{-1}+\Xi)^2+\Theta^a\Theta^a]$$

where

$$R=-\tilde{\gamma}^\alpha\otimes\tilde{\gamma}_\alpha t^a_c\otimes t^a_c\left(\frac{4\pi r^2}{g^2}\partial_t^2\right)^{-1}$$

is the free propagator of the bilocal fields and G_0 is the free-fermion propagator.

It is natural to split Ξ and Θ^a into 'classical' and 'quantum' parts,

$$\Xi=\Xi_0+\xi \qquad \Theta^a=\Theta^a_0+\vartheta.$$

The classical parts obey the following equations:

$$\frac{\delta S}{\delta\Xi}(\Xi_0,\Theta_0)=0 \qquad \frac{\delta S}{\delta\Theta^a}(\Xi_0,\Theta_0)=0.$$

These equations give $\Theta^a_0=0$ and

$$\begin{aligned}&\Xi_0=RG\\&G^{-1}=G_0^{-1}+\Xi_0.\end{aligned}\tag{5.23}$$

The function $G(r, t; r', t')$ is just the exact fermion propagator to leading order in $1/N_c$, while (5.23) is the Schwinger–Dyson equation for G.

In an analogy to the two-dimensional QCD, this equation can be solved explicitly. In terms of the Fourier transform of G in the variable $(t-t')$, the solution is

$$G(r, r'; \omega) = \exp\left[-\frac{g^2}{4\pi^2 rr'}|r-r'|\left(-\frac{1}{\omega}+\frac{1}{\lambda}\right)\right] G_0(r, r'; \omega)$$

where λ is an infrared cutoff. In the limit $\lambda \to 0$, the u-quark propagator vanishes, so u quarks do not propagate. The propagating degrees of freedom are described by the colour-singlet field ξ (for details see Craigie *et al* 1984). Its propagator is found from the Bethe–Salpeter equation which determines the propagating bound states. It is important that there exist massless quark–antiquark bound states $(\bar{u}_{Li}u^i_R)$ and $(\bar{u}_{Ri}u^i_L)$ always moving to and from the monopole, respectively. There are also two massless diquark states singlet under $SU(2)_c$, $(\varepsilon_{ij}u^i_R u^j_R)$ and $(\varepsilon_{ij}u^i_L u^j_L)$ which also move only towards the monopole and from it, respectively, and two antidiquark states conjugate to these. The direction of motion of these states corresponds to that of free fermions summarised in table 1.

The key property of the radial QCD is that the 'meson' and quark pictures are dual to each other: the correlation functions of locally singlet, gauge-invariant operators calculated within the $SU(2)_c$ gauge theory are equal to those in the free theory. Although this result was obtained first within the $1/N_c$ expansion (see Craigie *et al* 1984), it turns out to be exact (Craigie and Nahm 1984a, b, Craigie 1986a, b). We shall see this fact explicitly in the bosonised formalism in § 5.3. The duality means, in particular, that the calculation of the anomalous condensates is not affected by the $SU(2)_c$ interactions, so that the condensates (5.3) and (5.4) are still non-zero and of order of $1/r^6$ (as discussed above, the $U(1)'$ interactions have no effect on the condensates either). Thus the monopole catalysis of proton decay occurs in the full s-wave theory as well.

On the other hand, some elementary processes involving quarks are totally suppressed by the $SU(2)_c$ interactions (Craigie *et al* 1984), at least to leading order in $1/N_c$. These are the reactions which would involve incoming or outgoing states, non-singlet under $SU(2)_c$: the non-singlet states do not propagate. Examples of the suppressed reactions are provided by the processes (5.10) and (5.11). On the other hand, the reactions of type (5.8) and (5.12) are allowed (if u-quark pairs are understood as colour singlet bound states).

Thus the extra interactions absent in the $SU(2)_M$ approximation change the details of the fermion–monopole dynamics. However, they do not greatly suppress the monopole catalysis of proton decay. This result is supported by the bosonised treatment to be discussed next.

5.3. Bosonised treatment

To describe the s-wave monopole–fermion dynamics in SU(5) theory in the bosonised formalism, it is convenient to use the variables defined by (4.8), where each Dirac fermion corresponds to a scalar field $\phi(r, t)$. For the fermions (5.1) and (5.2) we have the boson fields ϕ_{u^1}, ϕ_{u^2}, ϕ_d and ϕ_e which obey the following boundary conditions at $r=0$:

$$\phi_{u^1} = \phi_{u^2} \qquad \partial_r\phi_{u^1} = -\partial_r\phi_{u^2} \tag{5.24}$$

$$\phi_d = \phi_e \qquad \partial_r\phi_d = -\partial_r\phi_e . \tag{5.25}$$

Making use of the representation (4.1), one finds the bosonised Hamiltonian of the $SU(2)_M \times SU(2)_c \times U(1)'$ model

$$H = \int dr(\mathcal{H}^{\text{free}}_{m=0} + \mathcal{H}_{Y_c} + \mathcal{H}_{EM} + \mathcal{H}_{SU(2)_c}) + \text{mass terms}$$

where $\mathcal{H}^{\text{free}}_{m=0}$ is the free bosonic Hamiltonian density,

$$\mathcal{H}_{Y_c} = \frac{\alpha_s}{6\pi r^2}(\phi_d - \tfrac{1}{2}\phi_{u^1} - \tfrac{1}{2}\phi_{u^2})^2 \tag{5.26}$$

$$\mathcal{H}_{EM} = \frac{\alpha}{2\pi r^2}(\phi_e + \tfrac{1}{3}\phi_d - \tfrac{2}{3}\phi_{u^1} - \tfrac{2}{3}\phi_{u^2})^2 \tag{5.27}$$

are the Coulomb terms related to the colour hypercharge and electromagnetic charge (it is convenient to take $U(1)_{Y_c} \times U(1)_{EM}$ instead of $U(1)_{t^3_M} \times U(1)'$ as an Abelian subgroup of $SU(2)_M \times U(1)'$), and $\mathcal{H}_{SU(2)_c}$ describes the $SU(2)_c$ interactions. The expressions for the Coulomb terms (5.26) and (5.27) were found by Callan (1983), while $\mathcal{H}_{SU(2)_c}$ was obtained by Rubakov and Serebryakov (1984). The expression for $\mathcal{H}_{SU(2)_c}$ is rather messy; the only property of $\mathcal{H}_{SU(2)_c}$ important for our purposes is that it depends on the only combination $(\phi_{u^1} - \phi_{u^2})$ which is non-singlet under $SU(2)_c$. Since $\mathcal{H}_{Y_c}$ and $\mathcal{H}_{EM}$ contain only the singlet combination,

$$\phi_u = \tfrac{1}{2}(\phi_{u^1} + \phi_{u^2})$$

the $SU(2)_c$ interactions decouple in the limit of massless quarks.

Thus the massless theory is analysed straightforwardly. In an analogy to § 4.3 one finds that the combination

$$f = \tfrac{1}{2}(\phi_e + \phi_d + 2\phi_u)$$

does not enter the Coulomb terms, so it obeys the free equation, $\partial^2 f = 0$, and the boundary condition

$$\partial_r f = 0 \qquad \text{at } r = 0 \tag{5.28}$$

which follows from (5.24) and (5.25). On the other hand, the Coulomb terms effectively change the boundary conditions for two other independent combinations, say

$$\Sigma_1 = \phi_d - \phi_u$$

$$\Sigma_2 = \phi_e + \tfrac{1}{3}\phi_d - \tfrac{4}{3}\phi_u.$$

In fact, Σ_1 and Σ_2 are neither properly normalised nor orthogonal to each other; however, these facts are not important to us. The only relevant property of $\Sigma_{1,2}$ is that they are orthogonal to f.

In the weak coupling limit, Σ_1 and Σ_2 obey the free equations, $\partial^2\Sigma_{1,2} = 0$, far away from the monopole. The effective boundary conditions are

$$\Sigma_1 = \Sigma_2 = 0 \qquad \text{at } r = 0. \tag{5.29}$$

A solution to the free equations with the boundary conditions (5.28) and (5.29) is easily constructed for any given initial data. For example, one finds that an incoming configuration corresponding to a d quark evolves into outgoing half-solitons as shown in figure 10. This is just the process (5.8) (the correspondence between solitons and fermions is described in § 4.2).

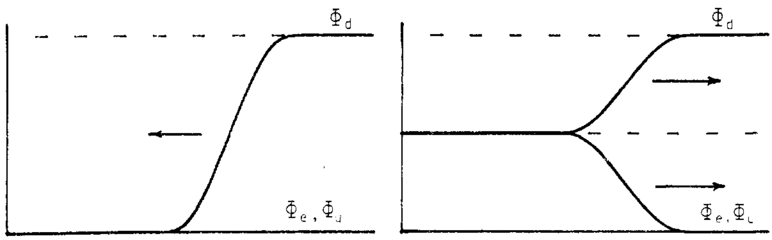

Figure 10. The process $d+\text{mon} \rightarrow \frac{1}{2}(d+\bar{u}_1+\bar{u}_2+e^+)+\text{mon}$ in the bosonised formalism.

Yet another way to see the proton decay catalysed by the monopole is to make use of the bag-like model of a proton suggested by Virasoro (1983b). One assumes that the monopole is placed in the centre of the bag and quarks inside the bag are in the s-wave state. Then the initial proton is described by the configuration of the boson fields shown in figure 11(a). To ensure that quarks are confined inside the bag, one imposes the boundary condition valid at any time,

$$\phi_u = \phi_d = \sqrt{\pi} \qquad \text{at } r = r_p$$

where r_p is the radius of the bag. The time evolution is then obtained straightforwardly (Rubakov and Serebryakov 1984). One finds that at $t > 2r_p$, the configuration is as shown in figure 11(b). It describes an outgoing positron and an excitation of the bag with zero baryon number (above the vacuum $\phi_{u^1} = \phi_{u^2} = \phi_d = \sqrt{\pi}$). Thus the lifetime of a proton in the presence of a monopole is, in this model, $\tau_p \sim 2r_p$ ($\sim 10^{-23}$ s for $r_p \sim 1$ fm).

Thus, both the bosonised formalism and $1/N_c$-expansion technique lead to the conclusion that the catalysis is not suppressed by the s-wave interactions absent in the $SU(2)_M$ model. Both these methods reproduce the decoupling of the $SU(2)_c$ interactions in the massless limit. The advantage of the $1/N_c$ expansion is that it provides a more detailed picture of the dynamics of the u quarks near the monopole. On the other hand, the bosonised technique is convenient for the qualitative description of the catalysis.

It is worth noting that the bosonised treatment is convenient also for analysing the effects of weak interactions at $r \lesssim M_W^{-1}$. The possible relevance of these interactions

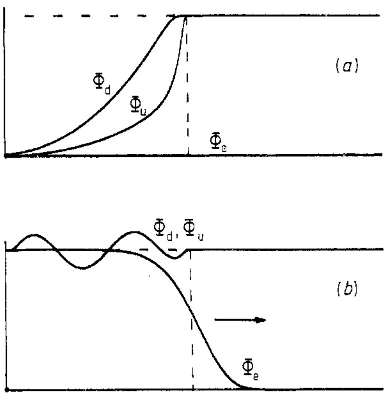

Figure 11. (a) An initial state of a proton in the bag-like model. (b) The final state: an outgoing positron and zero-baryon-number excitation of the bag.

for the catalysis was pointed out by Grossman *et al* (1983). However, the weak effects turned out not to suppress the monopole catalysis in the SU(5) theory (Rubakov and Serebryakov 1984, Goldhaber 1983, Sen 1983, Schellekens 1984a, b).

5.4. *Estimating the cross section*

The most important quantities for phenomenological applications are the catalysis cross section and branching ratios at low relative velocities of a nucleon and monopole. These quantities for various nuclei are required as well. Unfortunately, the calculation of the cross section and branching ratios meets with severe difficulties, so only a few results are available at the moment.

At large distances, the electromagnetic interactions of nucleons (and nuclei) and monopoles play a major role. These give rise to a peculiar dependence of the cross sections on the relative velocity, $\beta = v/c$, at low β. We recall that a pointlike charged Dirac fermion falls on the monopole if it is in the s wave (see § 2). The cross section of this process is (Kazama *et al* 1977)

$$\sigma_{\text{fall}} = \pi/2m_F^2\beta^2.$$

The origin of this β^{-2} behaviour is somewhat analogous to the enhancement of inelastic cross sections by an attractive Coulomb potential (Landau and Lifshitz 1964). So one can expect the catalysis cross section for a proton to be (Rubakov and Serebryakov 1983)

$$\sigma_p = \sigma_0/\beta^2 \qquad \text{at } \beta \ll 1 \tag{5.30}$$

where $\sigma_0 \sim 1\ \text{fm}^2$ on dimensional grounds. A more sophisticated analysis by Arafune and Fukugita (1983a) confirmed this expectation of the β^{-2} behaviour. Arafune and Fukugita (1983a) also studied the dependence on β of the catalysis cross sections for various nuclei and found that in many interesting cases suppression factors appear due to the interactions at large distances. They also estimated the validity regions of the estimates like (5.30).

Another effect of the interactions at relatively large distances is the possibility for a monopole to capture a proton or nucleus into a trapped orbit. The estimates due to Goebel (1983), Bracci and Fiorentini (1983) and Olaussen *et al* (1983, 1984) indicate that the capture cross section is of order 1–0.1 mb at $\beta \sim 10^{-3}$. This effect would lead to the enhancement of the catalysis rate for small catalysis cross sections (Bernreuther and Craigie 1985): if σ_0 were small, the catalysis would occur as a two-step process; at the first step the proton (or nucleus) would be captured and only then would it decay.

The calculation of the absolute value of the cross section involves a difficult problem of matching large distances ($r \gg 1$ fm) to intermediate ($r \sim 1$ fm) and small ($r \ll 1$ fm) ones. The only calculation reported so far is due to Bernreuther and Craigie (1985) (see also Craigie 1986a) who made use of the non-relativistic quark model. Their results seem to imply that σ_0 is of order 0.1 mb. However, much further work is needed to reach a firm estimate. An interesting suggestion in this direction was made by Callan and Witten (1984) who studied the catalysis within the soliton model of a proton (Skyrme 1961) and argued that it might be possible to calculate the cross section by solving classical field equations of a sigma model coupled to a monopole.

Thus the existing results indicate that the catalysis cross section is at least in the millibarn region; it is presumably much larger at low monopole–proton velocities.

Even a cross section of the order of $10^{-20}\,\mathrm{cm}^2$ at $\beta = 10^{-3}$–10^{-4} is not unrealistic. A large uncertainty in the theoretical expectations reflects the poor status of the subject.

6. Model dependence

The dependence of the catalysis on the monopole type and on the choice of GUT has been discussed rather extensively during the last few years. It has been found that in the SU(5) theory, the proton decay is induced not only by a fundamental monopole, but also by monopoles of many other types (Rubakov and Serebryakov 1983, Dawson and Schellekens 1983a, Schellekens 1984a, b). Even introducing supersymmetry does not change the properties of the catalysis (Dawson 1983): the fermion-monopole interactions near the core remain the same in the supersymmetric SU(5) GUT. It is worth noting that the dominant modes of the spontaneous proton decay in the supersymmetric SU(5) theory are $p \to \mu^+ K^0$ and $p \to \bar{\nu}_\mu K^+$ while the proton decay induced by a monopole still gives the modes (5.5), (5.17) and (5.18).

Monopole catalysis is inherent also in the SO(10) GUT (Dawson and Schellekens 1983a, b, Schellekens 1984a, b) and $SU(4) \times SU(2)_L \times SU(2)_R$ theory (Goldhaber 1983, Schellekens 1984a, b, Craigie 1986a). Schellekens (1984b) developed a convenient formalism for studying the s-wave fermion-monopole interactions for arbitrary spherically symmetric monopoles, which is an appropriate generalisation of the techniques presented in §§ 2 and 3. He found that catalysis occurs in a variety of models and for many monopole types. Goldhaber (1983), Schellekens (1984a, b) and Sen (1984, 1985) showed that in some models, the monopole catalysis is driven by the electroweak anomaly in the baryon number, and the catalysis still proceeds at strong interaction rates.

However, the catalysis effect is model dependent, unlike the very existence of monopoles in GUT. The simplest example of a GUT without catalysis is provided by the SU(5) theory with the doubling of generations. This model was suggested by Kuzmin and Shaposhnikov (1983) from a different point of view. The first generation fermions form two $\bar{5}$-plets, two 10-plets and a singlet,

$$\bar{5}_1 = (\bar{d}, E^-, E^0)_L \qquad \bar{5}_2 = (\bar{D}, e^-, \nu_e)_L$$

$$10_1 = \begin{pmatrix} \bar{u}UD \\ e^+ \end{pmatrix}_L \qquad 10_2 = \begin{pmatrix} \bar{U}ud \\ E^+ \end{pmatrix}_L$$

$$1 = E^0_R$$

where E^0, E^-, D and U are heavy ($m \gtrsim 100\,\mathrm{GeV}$) leptons and quarks. Spontaneous proton decay is forbidden in this model, but other baryon-number-violating processes, like n–$\bar{n}$ oscillations, are allowed. As discussed by Rubakov and Serebryakov (1984), instead of the monopole catalysis reactions (5.5), (5.17) and (5.18), the processes involving heavy fermions would be catalysed, for example

$$p + \mathrm{mon} \to E^+ + \mathrm{mon} + \ldots.$$

These reactions, however, are forbidden by energy conservation at low monopole-proton velocities, so monopole catalysis is totally suppressed at small β.

Other examples of GUT without catalysis were found by Dawson and Schellekens (1983a) and Schellekens (1984a, b). Although these models (as well as the Kuzmin-Shaposhnikov one) look rather unnatural, their existence ensures that the catalysis is not a general property of GUT.

Another aspect of the model dependence is the related problem of the catalysis of proton decay by the Kaluza-Klein monopoles. Monopole solutions in the five-dimensional Kaluza-Klein theory were found by Sorkin (1983) and Gross and Perry (1983), and they were generalised to higher dimensions by Bais and Batenburg (1985). Bais and Batenburg (1984), Kobayashi and Sugamoto (1984), Nelson (1984), and Ezawa and Iwazaki (1984) studied the interactions of fermions with the Kaluza-Klein monopoles at the quantum mechanical level. They found that the helicity changing effective boundary conditions emerge in the s-wave Dirac equation instead of the charge exchange typical to the s-wave fermions in the presence of the 't Hooft-Polyakov monopole. This fact was considered as an indication of the absence of the monopole catalysis in the Kaluza-Klein theories. However, we think that a decisive conclusion cannot yet be drawn, because the simplest Kaluza-Klein model discussed so far has several undesirable features which make it unrealistic. These features, however, seem to be important for the monopole-fermion interactions. Namely, the model contains a long-ranged scalar field (which should acquire a mass by some mechanism in realistic theories) interacting with fermions and preventing them from approaching the core; the fermion-gauge-field interaction is vectorial (instead of chiral), so the role of the anomaly is different. It is not excluded that the monopole catalysis will reappear in realistic Kaluza-Klein theories free of these undesirable features.

7. Are there monopoles around?

The theory of GUT monopoles and monopole catalysis of proton decay is reasonably well developed. On the experimental side, much effort is being expended in searching for monopoles and monopole catalysis. Unfortunately, no monopoles have been observed except two candidate events reported by Cabrera (1982) and Caplin *et al* (1986).

GUT monopoles are expected to be very heavy: predictions for their masses vary from 10^8 GeV (early unified theories) through 10^{16} GeV (minimal SU(5)) up to 10^{21} GeV (Kaluza-Klein theories). In any case, they can hardly be produced in the present universe, so the ultimate souce of monopoles is the big bang. Cosmological aspects of the monopole problem were reviewed by Langacker (1981) and Shafi (1986). Cosmic monopoles gain their kinetic energy by acceleration in galactic magnetic fields. The kinetic energy is estimated by (Goto 1963)

$$E_{\rm kin} \sim g_{\rm M} H_{\rm G} L_{\rm G} \sim 10^{11}\ {\rm GeV}$$

where $g_{\rm M}$ is the magnetic charge, and $H_{\rm G}$ and $L_{\rm G}$ are the magnitude of the galactic magnetic field and its spatial scale, respectively. So monopoles with $M_{\rm M} > 10^{11}$ GeV are slowly moving while monopoles with smaller masses are relativistic.

The first estimates by Domogatsky and Zheleznykh (1969), Zeldovich and Khlopov (1978) and Preskill (1979) of the present concentration of relic monopoles were in conflict with the observational and astrophysical bounds. However, several mechanisms were suggested for reducing the concentration; the estimates within various scenarios differed by many orders of magnitude (for a review see, e.g., Rubakov 1984). Thus, no decisive cosmological prediction for the present concentration is available at the moment.

Many astrophysical and cosmological bounds on the present monopole concentration and flux near the Earth have been reported (for a review see, e.g., Craigie 1986a). A reasonably safe yet restrictive bound was obtained by Parker (1970) and Chudakov (see Domogatsky and Zheleznykh 1969) who pointed out that too large a number of monopoles would exhaust galactic magnetic fields, in contradiction to observations. Roughly speaking, this bound is

$$F \leqslant F_{\mathrm{PC}} = 10^{-15}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}. \tag{7.1}$$

In fact, the actual bound depends slightly on the monopole mass (Lazarides *et al* 1981, Turner *et al* 1982). More stringent bounds come from the exhaustion of intracluster magnetic fields (Rephaeli and Turner 1983) as well as stellar magnetic fields (Ritson 1982, Dimopoulos *et al* 1982a, Harvey *et al* 1985). However, these bounds are less secure.

Another source of the astrophysical bounds is the catalysis of nucleon decay by monopoles trapped inside astronomical objects. The catalysis would heat up these objects; observational data on their luminosities can be translated into limits on the monopole flux. This observation was first made by Kolb *et al* (1982) and Dimopoulos *et al* (1982b) in the context of neutron stars, while Freese *et al* (1983) considered known pulsars. The reported neutron star bounds on the monopole flux were extremely stringent, but further discussions by Bais *et al* (1983), Kuzmin and Rubakov (1983), Harvey (1984) and others made it clear that the bounds are sensitive to unknown details of the inner core of a neutron star. A safer bound was obtained by Freese (1984) who considered white dwarfs. This bound depends on the catalysis cross section; a typical value of the allowed flux is 10^{-2}–$10^{-3}\,F_{\mathrm{PC}}$. There are also bounds coming from the Jovian planets (Arafune *et al* 1985) but these were also obtained with additional assumptions.

The present concentration of superheavy monopoles is also bounded by observational limits on the mass density of the universe at present or at the nucleosynthesis epoch (Zeldovich and Khlopov 1978, Preskill 1979, Lazarides *et al* 1981). These bounds are roughly the same as (7.1). Monopoles catalysing the proton decay could wash out the baryon asymmetry of the universe; the corresponding bound was found by Ellis *et al* (1982) not to be restrictive.

Thus, the conservative point of view is that the cosmological and astrophysical considerations do not totally exclude the monopole flux near the Earth of the order of F_{PC}. The flux of monopoles catalysing nucleon decay must presumably be smaller than $10^{-2}F_{\mathrm{PC}}$. One should realise, however, that none of the astrophysical or cosmological bounds is absolutely safe, so that the decisive limits come only from direct experimental searches for monopoles.

The best existing experiments have sensitivities to the monopole flux close to the Parker–Chudakov bound (for a review of the experimental searches for monopoles see Giacomelli 1986). The present limit from the Baksan scintillation telescope is (Alexeev *et al* 1985)

$$F < 1.5 \times 10^{-15}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}$$

valid in the velocity range $2 \times 10^{-4} < \beta < 6 \times 10^{-1}$. Almost the same limit for monopoles catalysing the proton decay comes from the water Čerenkov detectors IMB (Errede *et al* 1983, Stone 1985) and Kamiokande (Koshiba 1984). A search for monopole tracks in ancient mica performed by Price *et al* (1983) has led to a more stringent limit

$$F < 10^{-16}\text{-}10^{-17}\ \mathrm{cm^{-2}\,s^{-1}\,sr^{-1}}$$

for $3\times10^{-4} < \beta < 2\times10^{-3}$, valid under the assumption that monopoles capture nuclei while passing through the Earth (notice that monopoles catalysing the nucleon decay escape the mica limit; there are some other possibilities to evade this limit, see Ahlen *et al* 1983).

The sensitivities of the monopole detectors will be substantially increased in the near future. Induction experiments sensitive to $F \sim F_{\mathrm{PC}}$ are not unrealistic. A combined scintillation-track etch detector, MACRO, to be installed at Gran Sasso (De Marzo *et al* 1984) will be able to reach the flux $F \sim 10^{-16}\ \mathrm{cm^{-2}\,s^{-1}\,sr^{-1}}$ and even smaller. About the same sensitivity will be achieved by the track etch detector at Kamioka (Kawagoe *et al* 1983). Further improvements seem to come from Čerenkov detectors in open water. These installations are able to detect slowly moving monopoles catalysing nucleon decay at reasonably high rates (Domogatsky 1984, Rubakov *et al* 1983) as well as relativistic monopoles emitting a large amount of the Čerenkov light due to their large magnetic charge (Bezrukov *et al* 1984). Two single-string detectors of this type were installed at the Baikal Lake; recent limits on the flux of monopoles catalysing the proton decay are (Bezrukov *et al* 1986)

$$F < 6.5\times10^{-17}\ \mathrm{cm^{-2}\,s^{-1}\,sr^{-1}} \qquad \text{for } \sigma_0 = 10^{-26}\ \mathrm{cm^2}$$

$$F < 2.0\times10^{-16}\ \mathrm{cm^{-2}\,s^{-1}\,sr^{-1}} \qquad \text{for } \sigma_0 = 10^{-28}\ \mathrm{cm^2}$$

where σ_0 parametrises the catalysis cross section for a proton, $\sigma_{\mathrm{p}} = \sigma_0\beta^{-2}$. These limits are valid for $10^{-5} < \beta < 10^{-3}$. The reported Baikal limit on the flux of almost relativistic monopoles is (Bezrukov *et al* 1984)

$$F < 6\times10^{-16}\ \mathrm{cm^{-2}\,s^{-1}\,sr^{-1}} \qquad \text{for } \beta > 0.75.$$

Among the indirect searches, we point out one based on the observation by Rubakov (1981a), Dimopoulos *et al* (1982a) and Arafune and Fukugita (1983b) that the monopole catalysis of proton decay inside the Sun would lead to the emission of μ^+ (either directly in the reactions (5.17) and (5.18) or from π^+ decays) which in turn would decay producing ν_{e} (and $\bar{\nu}_\mu$). These neutrinos can be detected in underground proton decay installations, as well as in solar neutrino detectors. A search for these neutrinos has been performed at Kamiokande (Kajita *et al* 1985, Totsuka 1985). Translated into the limit on the monopole flux (although this translation is not quite adequate), the Kamiokande result is

$$F \leqslant 10^{-18}(10^{-23})\ \mathrm{cm^{-2}\,s^{-1}\,sr^{-1}}$$

for $\beta = 10^{-2}(10^{-4})$.

Thus, the experimental possibilities for monopole searches are far from exhausted. An intriguing feature of the present situation is that the existing detectors have reached conservative astrophysical and cosmological bounds and future experiments will perform searches well below these bounds. In view of the great importance of positive and even negative results, it is clear that monopole searches will continue to be developed. The question posed in the title of this section is yet to be answered.

Acknowledgments

The author is deeply indebted to his colleagues at the Institute for Nuclear Research of the Academy of Sciences of the USSR for their encouragement and helpful discussions during all stages of this work.

References

Abouelsaood A 1983a *Nucl. Phys.* B **226** 309–38

—— 1983b *Phys. Lett.* **125B** 467–9

Adler S 1969 *Phys. Rev.* **177** 2426–32

Ahlen S P, Price P B, Guo S and Fleischer R L 1983 *Proc. Workshop 'Monopole '83'* ed J L Stone (New York: Plenum) pp 383–90

Alexeev E N *et al* 1985 *Proc. 19th Int. Conf. on Cosmic Rays, La Jolla* **5** 52–63

Ambjørn J, Greensite J and Peterson C 1983 *Nucl. Phys.* B **221** 381–408

Ambjørn J and Rubakov V A 1985 *Nucl. Phys.* B **256** 434–48

Arafune J and Fukugita M 1983a *Phys. Rev. Lett.* **50** 1901–2

—— 1983b *Phys. Lett.* **133B** 380–4

Arafune J, Fukugita M and Yanagita S 1985 *Phys. Rev.* D **32** 2586–90

Bais F A 1978 *Phys. Rev.* D **18** 1206–10

Bais F A and Batenburg P 1984 *Nucl. Phys.* B **245** 469–80

—— 1985 *Nucl. Phys.* B **253** 162–86

Bais F A, Ellis J, Nanopoulos D V and Olive K 1983 *Nucl. Phys.* B **219** 189–219

Balachandran A P, Marmo G, Mukunda N, Nilsson J S, Sudarshan E C G and Zaccaria F 1983 *Phys. Rev. Lett.* **50** 1553–5

Balachandran A P and Schechter J 1983 *Phys. Rev. Lett.* **51** 1418–20

—— 1984 *Phys. Rev.* D **29** 1184–90

Bardeen W A 1969 *Phys. Rev.* **184** 1841–57

Belavin A A, Polyakov A M, Schwarz A S and Tyupkin Yu S 1975 *Phys. Lett.* **58B** 85–7

Bell J S and Jackiw R 1969 *Nuovo Cimento* **51** 47–68

Bennett D P 1985 *Phys. Rev.* D **31** 2323–31

Bernreuther W and Craigie N S 1985 *Phys. Rev. Lett.* **55** 2555–8

Bezrukov L B *et al* 1984 *Proc. Int. Conf. 'Neutrino-84', Heidelberg*

—— 1986 *Proc. 12th Int. Conf. 'Neutrino-86', Sendai*

Biebl K J 1982 *Proc. 16th Int. Symp., Ahrenshoop* pp 106–14

Blaer A S, Christ N H and Tang J-F 1981 *Phys. Rev. Lett.* **47** 1364–7

—— 1982 *Phys. Rev.* D **25** 2128–40

Bogomolny E B 1976 *Yad. Fiz.* **24** 861–8

Bracci L and Fiorentini G 1983 *Phys. Lett.* **124B** 493–6

Cabrera B 1982 *Phys. Rev. Lett.* **48** 1378–81

Callan C G 1982a *Phys. Rev.* D **25** 2141–6

—— 1982b *Phys. Rev.* D **26** 2058–68

—— 1983 *Nucl. Phys.* B **212** 391–400

—— 1984 *Proc. Workshop on Nonlinear Dynamics, La Jolla*

Callan C G and Das S 1983 *Phys. Rev. Lett.* **51** 1155–8

Callan C G, Dashen R F and Gross D J 1976 *Phys. Lett.* **63B** 334–40

—— 1978 *Phys. Rev.* D **17** 2717–63

Callan C G and Witten E 1984 *Nucl. Phys.* B **239** 161–76

Callias C J 1977 *Phys. Rev.* D **16** 3068–77

Caplin A D, Hardiman M, Koratzinos M and Schouten J C 1986 *Nature* **321** 402–6

Chetyrkin K G, Kazarnovsky M V, Kuzmin V A and Shaposhnikov M E 1981 *Phys. Lett.* **99B** 358–60

Christ N H 1980 *Phys. Rev.* D **21** 1591–602

Christ N H and Jackiw R 1980 *Phys. Lett.* **91B** 228–31

Coleman S 1975 *Phys. Rev.* D **11** 2088–97

—— 1978 *Proc. Int. School 'Ettore Majorana', Erice* ed A Zichichi (New York: Plenum)

—— 1981 *Proc. Int. School 'Ettore Majorana', Erice* ed A Zichichi (New York: Plenum)

Craigie N S 1986a *Theory and Detection of Magnetic Monopoles in Gauge Theories* ed N S Craigie (Singapore: World Scientific) pp 85-326, 379-400
—— 1986b *Phys. Rev.* D **33** 2930-43
Craigie N S and Nahm W 1984a *Phys. Lett.* **142B** 64-8
—— 1984b *Phys. Lett.* **147B** 127-31
Craigie N S, Nahm W and Rubakov V A 1984 *Nucl. Phys.* B **241** 274-308
Daniel M, Lazarides G and Shafi Q 1980 *Nucl. Phys.* B **170** 156-64
Dawson S 1983 *Proc. Workshop 'Monopole '83'* ed J L Stone (New York: Plenum) pp 229-38
Dawson S and Schellekens A N 1983a *Phys. Rev.* D **27** 2119-28
—— 1983b *Phys. Rev.* D **28** 3125-37
Dell'Antonio G F, Frishman Y and Zwanziger D 1972 *Phys. Rev.* D **6** 988-98
De Marzo C *et al* 1984 *Proposal for large area detector MACRO*
Dereli T, Swank J H and Swank L J 1975 *Phys. Rev.* D **11** 3541-4
Dimopoulos S, Glashow S L, Purcell E and Wilczek F 1982a *Nature* **298** 824-5
Dimopoulos S, Preskill J and Wilczek F 1982b *Phys. Lett.* **119B** 320-2
Dirac P A M 1931 *Proc. R. Soc.* A **133** 60-78
Dokos C P and Tomaras T N 1980 *Phys. Rev.* D **21** 2940-52
Domogatsky G V 1984 *Proc. Baksan School 'Partices and Cosmology'* (Moscow: INR) pp 34-42
Domogatsky G V and Zheleznykh I M 1969 *Yad. Fiz.* **10** 1238-44
Ebert D and Pervushin V N 1978 *Teor. Mat. Fiz.* **36** 313-20
Ellis J, Nanopoulos D V and Olive K A 1982 *Phys. Lett.* **116B** 127-32
Errede S *et al* 1983 *Phys. Rev. Lett.* **51** 245-8
Ezawa Z F and Iwazaki A 1983 *Z. Phys.* C **20** 335-45
—— 1984 *Nucl. Phys.* B **238** 638-52
Faddeev L D and Korepin V E 1978 *Phys. Rep.* **42** 1-87
Freese K 1984 *Astrophys. J.* **286** 216-20
Freese K, Turner M S and Schramm D N 1983 *Phys. Rev. Lett.* **51** 1625-8
Fujikawa K 1980 *Phys. Rev.* D **21** 2848-58
Gell-Mann M, Oakes R and Renner B 1968 *Phys. Rev.* **175** 2195-208
Georgi H and Glashow S L 1972 *Phys. Rev. Lett.* **28** 1494-6
—— 1974 *Phys. Rev. Lett.* **32** 438-41
Giacomelli G 1986 *Theory and Detection of Magnetic Monopoles in Gauge Theories* ed N S Craigie (Singapore: World Scientific) pp 407-99
Glashow S L 1979 *Harvard preprint* HUTP-79/A059
Goddard P and Olive D I 1978 *Rep. Prog. Phys.* **41** 1357-437
Goebel C J 1983 *Proc. Workshop 'Monopole '83'* ed J L Stone (New York: Plenum) pp 333-8
Goldhaber A S 1977 *Phys. Rev.* D **16** 1815-27
—— 1983 *Proc. 4th Workshop on Grand Unification* ed H A Weldon, P Langacker and P J Steinhardt (Basle: Birkhäuser)
Goto E 1963 *Prog. Theor. Phys.* **30** 700-12
Gross D J and Perry M 1983 *Nucl. Phys.* B **226** 29-48
Grossman B 1983 *Phys. Rev. Lett.* **50** 464-7
Grossman B, Lazarides G and Sanda A I 1983 *Phys. Rev.* D **28** 2109-11
Halpern M B 1975 *Phys. Rev.* D **12** 1684-99
Harvey J A 1984 *Nucl. Phys.* B **236** 255-73
Harvey J A, Ruderman M A and Shaham J 1985 *Princeton preprint*
Hasenfratz P and 't Hooft G 1976 *Phys. Rev. Lett.* **36** 1119-22
Ignatiev A Yu, Krasnikov N V, Kuzmin V A and Tavkhelidze A N 1978 *Phys. Lett.* **76B** 436-8
Isler K, Schmid C and Trugenberger C A 1986 *ETH preprint, Zurich*
Itoh K and Kazama Y 1985 *Phys. Lett.* **155B** 394-400
—— 1986 *Nucl. Phys.* B **266** 559-78
Jackiw R 1977 *Rev. Mod. Phys.* **49** 681-706
Jackiw R and Rebbi C 1976a *Phys. Rev. Lett.* **37** 172-5
—— 1976b *Phys. Rev.* D **13** 3398-409
Julia B and Zee A 1975 *Phys. Rev.* D **11** 2227-32
Kajita T *et al* 1985 *J. Phys. Soc. Japan* **54** 4065-70
Kawagoe K *et al* 1983 *University of Tokyo report*
Kazama Y 1983 *Prog. Theor. Phys.* **70** 1166-8
Kazama Y and Sen A 1984 *Nucl. Phys.* B **247** 190-224
Kazama Y, Yang C-N and Goldhaber A S 1977 *Phys. Rev.* D **15** 2287-99

Kazarnovsky M V, Kuzmin V A, Chetyrkin K G and Shaposhnikov M E 1980 *Zh. Eksp. Teor. Fiz. Pis. Red.* **32** 88–90
Klaiber B 1968 *Lectures in Theoretical Physics, Boulder Lectures* vol 10A (New York: Gordon and Breach)
Kobayashi M and Sugamoto A 1984 *Prog. Theor. Phys.* **72** 122–33
Kolb E W, Colgate S A and Harvey J A 1982 *Phys. Rev. Lett.* **49** 1373–5
Koshiba M 1984 *Proc. Int. Conf. on High Energy Physics, Leipzig* vol 2, ed A Meyer and E Wieczorek, pp 67–84
Kozakai H and Minakata H 1985 *Nucl. Phys.* B **249** 200–18
Krasnikov N V, Matveev V A, Rubakov V A, Tavkhelidze A N and Tokarev V F 1980 *Teor. Mat. Fiz.* **45** 313–28
Krasnikov N V, Rubakov V A and Tokarev V F, 1979a *J. Phys. A: Math. Gen.* **12** L343–9
—— 1979b *Yad. Fiz.* **29** 1127–34
Kuzmin V A 1970 *Zh. Eksp. Teor. Fiz. Pis. Red.* **13** 335–8
Kuzmin V A and Rubakov V A 1983 *Phys. Lett.* **125B** 372–5
Kuzmin V A, Rubakov V A and Shaposhnikov M E 1985 *Phys. Lett.* **155B** 36–42
Kuzmin V A and Shaposhnikov M E 1983 *Phys. Lett.* **125B** 449–51
Landau L D and Lifshitz E M 1964 *Quantum Mechanics* (Moscow: Nauka)
Langacker P 1981 *Phys. Rep.* **72** 185–385
Lavrelashvili G V 1986 *INR preprint, Moscow*
Lazarides G, Shafi Q and Walsh T F 1981 *Phys. Lett.* **100B** 21–4
Lowenstein J H and Swieca J A 1971 *Ann. Phys., NY* **68** 172–94
Mandelstam S 1975 *Phys. Rev.* D **11** 3026–30
Marciano W J and Muzinich I J 1983 *Phys. Rev. Lett.* **50** 1035–7
Marciano W J and Pagels H 1976 *Phys. Rev.* D **14** 531–4
Matveev V A, Rubakov V A, Tavkhelidze A N and Tokarev V F 1986a *Teor. Mat. Fiz.* **68** 3–17
—— 1986b *Teor. Mat. Fiz.* **69** 3–24
Mohapatra R and Marshak R E 1980 *Phys. Lett.* **94B** 183–6
Monastyrski M I and Perelomov A M 1975 *Zh. Eksp. Teor. Fiz. Pis. Red.* **21** 94–6
Nahm W 1986 *Theory and Detection of Magnetic Monopoles in Gauge Theories* ed N S Craigie (Singapore: World Scientific) pp 9–84
Nair V P 1983 *Phys. Rev. Lett.* **51** 631–3
Nelson P 1984 *Nucl. Phys.* B **238** 638–52
Nelson P and Coleman S 1984 *Nucl. Phys.* B **237** 1–31
Nelson P and Manohar A 1983 *Phys. Rev. Lett.* **50** 943–5
Nielsen H B and Ninomiya M 1983 *Phys. Lett.* **130B** 389–96
Nielsen N K and Schroer B 1977a *Nucl. Phys.* B **120** 62–76
—— 1977b *Phys. Lett.* **66B** 373–4
Olaussen K, Olsen H A, Osland P and Øverbø I 1983 *Nucl. Phys.* B **228** 567–87
—— 1984 *Phys. Rev. Lett.* **52** 325–8
Pak N, Panagiotakopoulos C and Shafi Q 1982 *Trieste preprint* IC/82/174
Parker E N 1970 *Astrophys. J.* **160** 383–9
Pati J and Salam A 1973 *Phys. Rev.* D **8** 1240–52
Peccei R and Quinn H 1977 *Phys. Rev. Lett.* **38** 1440–3
Pogrebkov A K and Sushko V N 1975 *Teor. Mat. Fiz.* **24** 425–9
Polchinski J 1984 *Nucl. Phys.* B **242** 345–66
Polyakov A M 1974 *Zh. Eksp. Teor. Fiz. Pis. Red.* **20** 430–2
Prasad M K and Sommerfield C 1975 *Phys. Rev. Lett.* **35** 760–2
Preskill J 1979 *Phys. Rev. Lett.* **43** 1365–8
Price P B, Guo S, Ahlen S P and Fleisher R L 1983 *Phys. Rev. Lett.* **52** 1265–8
Rajaraman R 1982 *Solitons and Instantons* (Amsterdam: North-Holland)
Rephaeli Y and Turner M S 1983 *Phys. Lett.* **121B** 115–8
Ritson D M 1982 *Stanford University preprint*
Ross G G 1981 *Rep. Prog. Phys.* **44** 655–718
Rubakov V A 1981a *INR preprint, Moscow* P-0211
—— 1981b *Zh. Eksp. Teor. Fiz. Pis. Red.* **33** 658–60
—— 1982 *Nucl. Phys.* B **203** 311–48
—— 1984 *Proc. Baksan School 'Particles and Cosmology'* (Moscow: INR) pp 16–33
—— 1985a *Zh. Eksp. Teor. Fiz. Pis. Red.* **41** 218–20
—— 1985b *Nucl. Phys.* B **256** 509–24
Rubakov V A and Serebryakov M S 1983 *Nucl. Phys.* B **218** 240–68
—— 1984 *Nucl. Phys.* B **237** 329–41

Rubakov V A, Stern B E and Tinyakov P G 1985 *Phys. Lett.* **160B** 292-6
Rubakov V A, Stern B E and Zheleznykh I M 1983 *Proc. Int. Colloq. on Matter Non-Conservation, ICOMAN-83, Frascati* ed E Bellotti and S Stipcick (Frascati: Serv. Doc. LNF) pp 329-33
Rubakov V A and Tavkhelidze A N 1985 *Phys. Lett.* **165B** 109-12
Sakharov A D 1967 *Zh. Eksp. Teor. Fiz. Pis. Red.* **5** 32-4
Sathiapalan B and Tomaras T 1983 *Nucl. Phys.* B **224** 491-507
Scadron M D 1981 *Rep. Prog. Phys.* **44** 215-92
Schellekens A N 1984a *Phys. Rev.* D **29** 2378-86
—— 1984b *Nucl. Phys.* B **246** 494-544
Schmid C 1983 *Phys. Rev.* D **28** 1802-3
Schmid C and Trugenberger C A 1984 *ETH preprint, Zurich*
—— 1986 private communication
Schwinger J 1962 *Phys. Rev.* **128** 2425-9
Sen A 1983 *Phys. Rev.* D **28** 876-81
—— 1984 *Phys. Rev. Lett.* **52** 1755-8
—— 1985 *Nucl. Phys.* B **250** 1-19
Seo K 1983 *Phys. Lett.* **126B** 201-6
Serebryakov M S 1985 *INR preprint, Moscow* P-0433
—— 1986 *PhD Thesis* INR, Moscow
Shafi Q 1986 *Theory and Detection of Magnetic Monopoles in Gauge Theories* ed N S Craigie (Singapore: World Scientific) pp 327-78
Skyrme T H R 1961 *Proc. R. Soc.* A **260** 127-39
Sorkin R D 1983 *Phys. Rev. Lett.* **51** 87-90
Stone J L 1985 *University of Michigan report* UM HE 85-02
Tamm I E 1931 *Z. Phys.* **71** 141-53
't Hooft G 1974a *Nucl. Phys.* B **75** 461-78
—— 1974b *Nucl. Phys.* B **79** 276-92
—— 1976a *Phys. Rev. Lett.* **37** 8-11
—— 1976b *Phys. Rev.* D **14** 3432-54
Totsuka Y 1985 *Proc. Int. Symp. on Lepton-Photon Interactions, Kyoto*
Turner M S, Parker E N and Bogdan T J 1982 *Phys. Rev.* D **26** 1286-305
Tyupkin Yu S, Fateev V A and Schwarz A S 1975 *Zh. Eksp. Teor. Fiz. Pis. Red.* **21** 91-3
Virasoro M A 1983a *Phys. Lett.* **125B** 161-4
—— 1983b private communication
Weinberg S 1978 *Phys. Rev. Lett.* **40** 223-6
Wilczek F 1978 *Phys. Rev. Lett.* **40** 279-82
—— 1982 *Phys. Rev. Lett.* **48** 1146-8
Witten E 1979 *Phys. Lett.* **86B** 283-7
—— 1982 *Phys. Lett.* **117B** 324-8
Yamagishi H 1983 *Phys. Rev.* D **27** 2383-96
Yan T M 1983 *Phys. Rev.* D **28** 1496-508
Yoneya T 1984 *Nucl. Phys.* B **232** 356-64
Yoshimura M 1978 *Phys. Rev. Lett.* **41** 281-4
Zeldovich Ya B and Khlopov M Yu 1978 *Phys. Lett.* **79B** 239-41