BERRY PHASES, MAGNETIC MONOPOLES, AND WESS-ZUMINO TERMS OR HOW THE SKYRMION GOT ITS SPIN*

By I. J. R. Aitchison

CERN, Geneva, Switzerland

(Received July 31, 1986)

An elementary discussion is given of the mechanism whereby the Wess-Zumino term determines the quantization of the Skyrme soliton. The work of Balachandran et al. is drawn upon to make explicit the remark of Wu and Zee that the Wess-Zumino term acts like a monopole in the space of scalar fields of the non-linear σ-model. The origin of the monopole structure, and its influence on quantization, is discussed in terms of the Berry (adiabatic) phase.

PACS numbers: 11.17.+y

1. *Introduction and outline*

The thing about Skyrmions [1] that is surely hardest to understand is how a lump-like solution (soliton) of a classical scalar field theory can, and in some cases even *must*, be quantized as a fermion. How can you add integers together and get a half left over? I want to draw together here some recent work on this subject, which has certainly helped me to understand how this marvellous trick is pulled.

It has of course been known for quite some time that a classical *extended* object (for example, a top [2, 3]) may be quantized as a fermion. A system which provides an explicit model of how this can come about — and one which is directly relevant to Skyrmions — is that of a particle of charge e in motion about a fixed magnetic monopole of strength g. Almost immediately after Dirac's 'monopole' paper [4], Tamm [5] studied the Schrödinger equation for this system, and found that the solutions of the angular equation are the rotation functions $\mathscr{D}^{j}_{m'm}(\theta, \phi)$, with $m = eg$ (in units $\hbar = c = 1$); when the product eg has the minimum non-zero value

$$eg = \tfrac{1}{2} \tag{1.1}$$

allowed by the Dirac quantization argument [4], or more generally the value $(n+1/2)$, the system has half-odd angular momentum and is a fermion. (Sometimes this circumstance

* This is an expanded version of the second of two lectures given at the XXVI Cracow School of Theoretical Physics, Zakopane, Poland, 1–13 June, 1986.

is used to run the argument the other way — i.e. that quantization of angular momentum yields the Dirac condition — but, in the present context at least, the Dirac condition will be fundamental.) More recently, field-theory examples of the charge-monopole system have been studied, with analogous results [6, 7].

In the two papers which initiated the recent burst of activity on Skyrmions [1] (and much else besides), Witten [8, 9] showed that the Wess-Zumino (W – Z) term [8–10] in the action for the scalar fields ϕ_a (whose solitons are Skyrmions) actually *determines* how these solitons are to be quantized. He obtained the remarkable result that the Skyrmion is a fermion if N_c is odd, and a boson if N_c is even: furthermore, the W – Z term also determines the pattern of spin-SU(3) multiplets ([1/2$^+$, **8**], [3/2$^+$, **10**]...) in the baryon spectrum [9, 11, 12]. Though obviously correct mathematically, these results were nevertheless still hard to explain in physical terms, especially to anyone who did not know what a W – Z term was — and even to those who did[1].

A good deal of light has been shed on this by the work of Balachandran and collaborators [13–17], Berry [18], Stone [19], and Wu and Zee [20, 21]. I shall try to state the major ideas in single sentences, which we will then examine in greater detail in the following sections.

(*i*) The W – Z term is a generalization, to the configuration space of scalar fields ϕ_a, of the charge-monopole interaction term in ordinary configuration space for particles. It acts like a monopole in ϕ-space.
(*ii*) Because Skyrmion field configurations are maps between field space and real space, the monopole structure of the W – Z term in field space induces, for such configurations, monopole structure in real space.
(*iii*) Upon quantization, fermionic behaviour will emerge via the well-known monopole mechanism referred to above.

These sentences state where we are trying to go, but they do not explain (*a*) where the W – Z term itself comes from, or (*b*) why it is like a monopole in ϕ-space. The short answer to (*a*) is: from the very fermion determinant which we studied in the previous lecture, but generalized to $SU(3)_f$, i.e. it is a term in the effective action for the ϕ fields which arises after integrating over the fermions [22, 23]. This is all very well in its way, but it too is mysterious: why does such an exotic term get induced in the boson sector when we integrate out the fermions? The technical answer to *this* is that the underlying fermion theory has anomalies, which can be calculated from single fermion loop diagrams. These diagrams generate effective vertices in the external fields (ϕ_a, gauge fields, etc.) coupled to the fermions. Hence any bosonic action obtained by integrating out the fermions — which is equivalent to summing all single fermion loop diagrams—*must* faithfully represent these anomaly-induced vertices. The W – Z action precisely encodes these anomalous vertices: if we only consider the 'ungauged' W – Z action, which is a function of the $SU(3)_f$ chiral field ϕ alone, we are representing correctly just the $SU(3)_f$ flavour anomalies of the underlying Fermion theory.

---

[1] For those who know that there is no W-Z term if the flavour group is SU(2), and wonder what happens then, see Section 6.

But anomalies are pretty mysterious too—are we not getting into an infinite regress of 'explanations'? It would be nice to have some kind of quantum mechanical *analogue*, at least, for what is going on. We can get a clue what to look for when we remember that the characteristic thing about anomaly-induced vertices is that they are independent of the fermion mass M; it is precisely this circumstance that allows the anomaly-cancellation mechanism discussed in the previous lecture, to work. Thus these 'anomalous' vertices will still survive in the fermion determinant with the correct coefficients, even as M becomes very large. This means that these particular vertices—or, equivalently, these particular contributions to the induced bosonic action — can be reliably calculated by the derivative expansion technique: $\partial\phi/M$ can be made as small as we like. (Some explicit examples of this way of calculating anomalous vertices are given in Ref. [22].) Now, a very large fermion mass M implies a large *gap* between the negative energy (sea) levels and the positive energy levels. Small values of $\partial\phi/M$ mean that the momenta and energies associated with shese 'slowly' varying ϕ fields are much less than the mass gap, and will therefore not induce tignificant fermionic excitations across the gap — indeed, in the limit of M *very* large, there will be no excitations at all, and we need only deal with the fermion vacuum (ground state).

This state of affairs is something we can find a quantum mechanical (rather than quantum field theoretic) analogue for. It arises quite frequently in many-body physics. Suppose we have a system described in terms of two sets of degrees of freedom: one (which we call $\boldsymbol{r}$) is 'fast moving' with 'large' differences between excitation levels, and the other ($\boldsymbol{R}$) is 'slow moving' with 'small' associated energy differences. We may think of the electronic (fast) and nuclear (slow) degrees of freedom (d.f.s) in a molecule for instance. It should make sense, when considering the $\boldsymbol{r}$ coordinates, to regard the $\boldsymbol{R}$'s as approximately constant; indeed this is called the adiabatic, or Born-Oppenheimer approximation in quantum mechanics. More precisely, if the $\boldsymbol{R}$ were constant, we would simply solve the stationary state Schrödinger equation for the $\boldsymbol{r}$'s, with the $\boldsymbol{R}$'s appearing parametrically:

$$H_{\boldsymbol{r}}(\boldsymbol{R})\psi_n(\boldsymbol{r}, \boldsymbol{R}) = E_n(\boldsymbol{R})\psi_n(\boldsymbol{r}, \boldsymbol{R}). \tag{1.2}$$

In reality, the $\boldsymbol{R}$'s are varying slowly with time, but not quickly enough to induce transitions from one E_n level to another. Thus the system, if started in a particular E_n level, 'stays with it' as the $\boldsymbol{R}$'s change. This is essentially the content of the adiabatic theorem: the 'fast' coordinates stay in the original eigenstate, which however itself changes slowly in response to the slow changes in the $\boldsymbol{R}$ coordinates which appear parametrically. This sounds very much like the situation of our fermion vacuum evolving slowly in response to the slowly varying ϕ's. But where is the quantum-mechanical analogue of the W—Z term? It must correspond to some non-trivial structure left behind in the space of the 'slow' parameters when we adiabatically decouple the 'fast' ones.

Here is where the work of Berry [18], and Kuratsuji and Iida [24] comes in. The adiabatic assumption tells us that, at any time t, the state of the system $|\psi(t)\rangle$ (adopting now a slightly more abstract notation) will essentially be the 'instantaneous' eigenstate $|n(\boldsymbol{R}(t))\rangle$, where

$$H(\boldsymbol{R}(t))\,|n(\boldsymbol{R}(t))\rangle = E_n(\boldsymbol{R}(t))\,|n(\boldsymbol{R}(t))\rangle, \tag{1.3}$$

if it was prepared to be in one of these states $|n(\boldsymbol{R}_0)\rangle$ at $t = 0$, where $\boldsymbol{R}_0 = \boldsymbol{R}(t = 0)$. In fact, $|\psi(t)\rangle$ will be related to $|n(\boldsymbol{R}(t))\rangle$ by a phase factor. What phase factor? The naïve answer would surely be

$$|\psi(t)\rangle = [\exp -i\int_0^t E_n(\boldsymbol{R}(t'))dt'] \cdot |n(\boldsymbol{R}(t))\rangle, \tag{1.4}$$

the expected integrated 'quasi-stationary state' phase. But this in *not* the whole story. An *additional* phase is generated during such an adiabatic change. That this is so in principle has been known for a long time (see, for example, Ref. [25]), but it had tended to be dismissed as unimportant physically ('just a phase'). Berry [18] pointed out a number of cases where the phase could be of considerable physical interest (see also Ref. [26]). In particular, a non-trivial phase can result from a closed path in $\boldsymbol{R}$ space, as we move along $\boldsymbol{R}_0 \to \boldsymbol{R}(t) \to \boldsymbol{R}_0$. Such 'Berry phases' depend on the actual path followed in $\boldsymbol{R}$-space — which may remind us of something...

The Berry phase is easily calculated [18]. We are looking for a solution of

$$H(\boldsymbol{R}(t))\,|\psi(t)\rangle = i\frac{d}{dt}|\psi(t)\rangle, \tag{1.5}$$

and we try the adiabatically-inspired ansatz

$$|\psi(t)\rangle = [\exp -i\int_0^t E_n(\boldsymbol{R}(t'))dt'] \cdot \exp i\gamma_n(t) \cdot |n(\boldsymbol{R}(t))\rangle. \tag{1.6}$$

Inserting (1.6) directly into (1.5) and using (1.3) yields

$$\gamma_n(t) = i\int_0^t \langle n(\boldsymbol{R}(t'))|\frac{d}{dt'}|n(\boldsymbol{R}(t'))\rangle dt'$$

$$= i\int_0^t \langle n(\boldsymbol{R})|\nabla_{\boldsymbol{R}} n(\boldsymbol{R})\rangle \cdot d\boldsymbol{R}, \tag{1.7}$$

the fundamental formula [18] for the Berry phase $\gamma_n(t)$.

Now — having dealt adiabatically with the $\boldsymbol{r}$ d.f.'s this way — let us turn our attention to the slowly varying $\boldsymbol{R}$'s, and consider *them* as quantum d.f.'s, not just classical parameters. In the same adiabatic approximation, we sit in one 'electronic' state n and look for solutions in which the total state function has the product form $\phi_n(\boldsymbol{R})|n(\boldsymbol{R})\rangle$, and ask: what Schrödinger equation does $\phi_n(\boldsymbol{R})$ obey? The answer is very interesting [25]: if $V(\boldsymbol{R})$ is the potential energy relevant to the $\boldsymbol{R}$ d.f.'s alone, then $\phi_n(\boldsymbol{R})$ obeys

$$[\text{'covariant kinetic energy'} + E_n(\boldsymbol{R}) + V(\boldsymbol{R})]\phi_n(\boldsymbol{R}) = i\frac{d}{dt}\phi_n(\boldsymbol{R}), \tag{1.8}$$

where by 'covariant kinetic energy' is meant that the gradient operator ∇_R in the normal R-kinetic energy terms is replaced by

$$\nabla_R \to \nabla_R + \langle n(R)|\nabla_R n(R)\rangle \tag{1.9}$$

$$\equiv \nabla_R - iA_n(R), \tag{1.10}$$

where (1.10) follows since the matrix element in (1.8) is easily seen to be pure imaginary. Thus a sort of gauge potential has been induced in R-space!

It is clear that this gauge potential is intimately related to the Berry phase; they are two facets of the same subtlety in the adiabatic approximation. Indeed we can see from (1.9) and (1.10) exactly what the local phase invariance corresponding to this 'gauge' structure is: an R-dependent phase change on $|n(R)\rangle$ induces a change in A_n of (1.10), which in turn causes a precisely compensating phase change in $\phi_n(R)$, so that the total state function $\phi_n(R)|n(R)\rangle$ is locally phase invariant. Thus a distinctly non-trivial structure has appeared in the 'slow' d.f.'s after adiabatic decoupling of the 'fast' d.f.'s. Note, incidentally, that the Berry phase is nothing but

$$\exp i\gamma_n(t) = \exp\left[i\int_0^t A_n(R)\cdot dR\right], \tag{.11}$$

so that we were right to be reminded of the *path-dependence* of the wave function for a particle in an electromagnetic potential A.

Thus we are getting nearer to understanding how funny phase factors — which might influence apparent rotational properties [26] — can arise via adiabatic decoupling. We can make closer contact with the field theory if we reformulate the adiabatic approximation in the path integral formalism. This was done by Kuratsuji and Iida [24]. In view of (1.6) and (1.11) we can almost guess what the result must be. We want the dynamics in R-space to correspond to a 'particle' moving in an (additional) 'vector potential' $A_n(R)$. Thus we expect to find a piece in the effective action $S_{\text{eff},n}$ in R-space which corresponds to the effective Lagrangian

$$\mathscr{L}_{\text{eff},n} = A_n(R)\cdot\frac{dR}{dt}. \tag{1.12}$$

Indeed, in that case

$$S_{\text{eff},n}(T) = \int_0^T \mathscr{L}_{\text{eff},n}dt = \int_0^T A_n(R)\cdot\frac{dR}{dt}\,dt, \tag{1.13}$$

and

$$\exp iS_{\text{eff},n}(T) = \exp i\gamma_n(T). \tag{1.14}$$

This is just what Kuratsuji and Iida obtain. By considering the trace of the evolution operator $\text{tr}\exp(-iHT)$ in the adiabatic approximation, they show that it is given by

$$K_{\text{eff}}(T) = \sum_n \int \mathscr{D}R \exp\left\{iS_0 - i\int_0^T E_n(R)dt' + i\gamma_n(T)\right\}, \tag{1.15}$$

where $\boldsymbol{R}(T) = \boldsymbol{R}_0$ (since for the trace we want to return to the same state at $t = T$), and where γ_n is now evaluated over closed loops $\boldsymbol{R}_0 \to \boldsymbol{R}(t) \to \boldsymbol{R}(T) = \boldsymbol{R}_0$ in $\boldsymbol{R}$-space:

$$\gamma_n(T) = i \oint \langle n(\boldsymbol{R}) | \nabla_{\boldsymbol{R}} n(\boldsymbol{R}) \rangle \cdot d\boldsymbol{R} = \oint \boldsymbol{A}_n(\boldsymbol{R}) \cdot d\boldsymbol{R} = \oint \mathscr{L}_{\text{eff},n} dt. \tag{1.16}$$

S_0 is the ordinary action for the $\boldsymbol{R}$ coordinates.

Now, finally, how can we understand the *specific* 'monopole-like' structure which corresponds (we have asserted) to the W–Z term? The secret, as Stone [19] pointed out, lies in a beautiful discovery by Berry [18]. We have assumed throughout that the eigenvalues $E_n(\boldsymbol{R})$ were well separated, and certainly not degenerate. But what happens if, for some particular value of the $\boldsymbol{R}$ d.f.'s, say $\boldsymbol{R}^*$, two of the E_n's coalesce? We expect some sort of catastrophe to show up in our adiabatic result. In fact, this point in $\boldsymbol{R}$-space is very likely to be a point at which the vector potential $\boldsymbol{A}_n(\boldsymbol{R})$ is singular! Such a vector potential would imply sources (δ-function singularities in the associated field strengths) — for example, magnetic monopoles. This is exactly what Berry found, explicitly, for the case in which the degeneracy is a spin-type degeneracy, the 'fast' coordinates are spin d.f.'s, and the 'slow' ones are angles describing the orientation of the (real!) magnetic field $\boldsymbol{B}$. The equation corresponding to (1.5) is then

$$\mu B \boldsymbol{S} \cdot \hat{\boldsymbol{B}} |\psi(t)\rangle = i \frac{d}{dt} |\psi(t)\rangle \tag{1.17}$$

and

$$E_n(\boldsymbol{B}) = \tfrac{1}{2} \mu B n, \tag{1.18}$$

where $n/2$ is the spin eigenvalue, which takes $2s+1$ values. Clearly these $2s+1$ states are degenerate when $\boldsymbol{B} = \boldsymbol{0}$ (the point $\boldsymbol{R}^*$). Berry found that the associated $\boldsymbol{A}_n(\boldsymbol{B})$ was precisely that of a monopole in $\boldsymbol{B}$-space located at $\boldsymbol{B} = \boldsymbol{0}$, having strength $-n/2$ (i.e. $eg = -n/2$). Thus spin-type degeneracies cause monopoles to lurk in the 'slow' space.

We can now see why the integral in (1.16) along a closed loop need not vanish. If we convert the line integral in (1.16) by a (multi-dimensional) Stokes theorem to a surface integral over the 'magnetic field' $\boldsymbol{B}_n = \nabla \times \boldsymbol{A}_n$, and thence to a volume integral via Gauss, we would normally get zero since div $\boldsymbol{B} = 0$. However, for the singular potential corresponding to a monopole div $\boldsymbol{B}_n \neq 0$ and a closed loop contributes a non-zero result. Actually we can go even further than this. The line integral over a closed loop C becomes

$$\oint_C \boldsymbol{A}_n(\boldsymbol{R}) \cdot d\boldsymbol{R} = \iint_S \boldsymbol{B}_n \cdot d\boldsymbol{S}, \tag{1.19}$$

where $\boldsymbol{S}$ is a surface spanning C. But what surface? Should we take an $\boldsymbol{S}_1$ (see Fig. 1) which is 'above' C, or an $\boldsymbol{S}_2$ which is 'below'? For consistency we must have

$$\iint_{S_1} \boldsymbol{B}_n \cdot d\boldsymbol{S} = \iint_{S_2} \boldsymbol{B}_n \cdot d\boldsymbol{S} + 2N\pi \tag{1.20}$$

(remember that these quantities are all *phases*). Since the normals for $\boldsymbol{S}_2$ and for $\boldsymbol{S}_1$ are oppositely oriented, we see that (1.20) is equivalent to

$$\oiint \boldsymbol{B}_n \cdot d\boldsymbol{S} = 2N\pi, \tag{1.21}$$

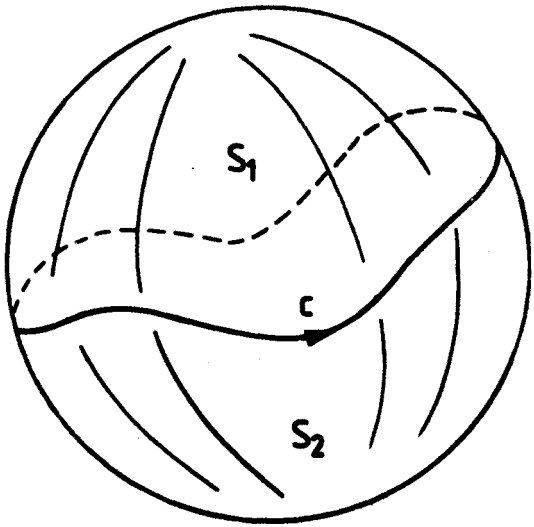

Fig. 1. Two surfaces spanning the curve C on S_2

where the integral is over a *closed* surface surrounding $\boldsymbol{R}$. Thus the total flux out of the 'monopole' is quantized — which is just the Dirac [3] condition ($eg = N \cdot 1/2$, in the real electromagnetic (e.m.) case). The above argument was a poor man's version of a deeper topological treatment, since we relied heavily on tacitly thinking of $\boldsymbol{R}$ as three-dimensional. Nevertheless, the result is correct.

But now notice a remarkable thing: the previous paragraph has shown quite generally that the monopole strength (total flux through a closed surface, divided by 4π) has the quantized value $N/2$, where N is an integer of topological significance. The paragraph before that stated the result that monopole-like structure arose from (1.17), in which the strength of the monopole is $-n/2$, where $n/2$ is the spin eigenvalue. The eigenvalue spectrum of (1.17) near $\boldsymbol{B} = \boldsymbol{0}$ (the point of degeneracy) seems to know something about topology!

We have learned that monopole-like structure can be generated in the 'slow' space $\boldsymbol{R}$ when we adiabatically decouple the 'fast' d.f.'s. Furthermore, the strength of the monopole interaction is an integer (divided by 2), from topological considerations, and this integer corresponds to some label of the energy spectrum of the 'fast' coordinates. This is as near as we are likely to get to a quantum mechanical analogue of the W−Z mystery. The W−Z term results [22, 23] from adiabatically decoupling the ψ's from the π's, starting from a Dirac equation

$$[-i\boldsymbol{\alpha} \cdot \nabla + \beta\mu \exp(i\boldsymbol{\lambda} \cdot \boldsymbol{\pi}\gamma_5 f^{-1})]\psi = i\frac{\partial \psi}{\partial t}, \tag{1.22}$$

which is the analogue of (1.17); μ is a mass parameter, $f \approx 93$ MeV, λ_a ($a = 1, 2, \ldots, 8$) are the Gell-Mann matrices, and the eight π fields are the analogues of the angle variables in $\hat{\boldsymbol{B}}$. The W−Z term (in the fields π) looks like a monopole in π space. Its coefficient is found to be an integer (which is, of course, N_c) by topological considerations [8] exactly

analogous to those given above for the $\boldsymbol{R}$-space monopole. There is one gap left to be closed: what is it in the spectrum of the Dirac equation (1.22) that 'knows' about topology (and hence about monopoles)? That is a *deep* question, the answer to which is provided by the mathematical subject called index theory. This way of looking at anomalies (remember?) is called the 'Hamiltonian approach' [27, 28], and is precisely the quantum field theoretical analogue of the quantum mechanical Berry-phase discussion outlined above.

Let us now see how all the foregoing works out in some simple cases.

2. *A simple example*

Consider, following Stone [19], a spin-1/2 particle in a magnetic field $\boldsymbol{B} = B\boldsymbol{n}$, where $\boldsymbol{n}^2 = 1$. The state function for the ('fast') spin d.f.'s satisfies

$$\mu\boldsymbol{\sigma} \cdot \boldsymbol{n}(t)|\psi(t)\rangle = i\frac{d}{dt}|\psi(t)\rangle, \tag{2.1}$$

where the magnitude B of the magnetic field has been absorbed into μ. The 'slow' d.f.'s are $\boldsymbol{n}$, since we shall only consider slow variations of $\boldsymbol{B}$ with fixed B. We consider the large μ limit (cf. large μ in (1.22)), so that the slow changes in $\boldsymbol{n}$ do not cause transitions between the two spin eigenstates $|\uparrow\boldsymbol{n}(t)\rangle$ and $|\downarrow\boldsymbol{n}(t)\rangle$, in the adiabatic approximation. Suppose at $t = 0$ we start in the state $|\uparrow\boldsymbol{n}(0)\rangle$, where $\boldsymbol{n}(0) = (\sin\theta_0\cos\phi_0, \sin\theta_0\sin\phi_0, \cos\theta_0)$. At a general time t, $\boldsymbol{n}(t) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, where the time-dependent d.f's θ and ϕ will vary over the surface of an S_2. The Berry phase $\gamma_\uparrow(t)$ is

$$\gamma_\uparrow(t) = i\int_0^t \langle\boldsymbol{n}(t')\uparrow\left|\frac{d}{dt'}\right|\uparrow\boldsymbol{n}(t')\rangle dt'. \tag{2.2}$$

We shall calculate this directly using an explicit wave function for $|\uparrow\boldsymbol{n}\rangle$; already here a crucial feature will emerge.

The wave function

$$\langle\theta\phi|\uparrow\boldsymbol{n}+\rangle = \begin{pmatrix}\cos\theta/2\\ \sin\theta/2e^{i\phi}\end{pmatrix} \tag{2.3}$$

is certainly an eigenfunction of

$$\mu\boldsymbol{\sigma}\cdot\boldsymbol{n} = \mu\begin{pmatrix}\cos\theta & \sin\theta e^{-i\phi}\\ \sin\theta e^{i\phi} & -\cos\theta\end{pmatrix} \tag{2.4}$$

with eigenvalue μ. Inserting (2.3) into (2.2) we find

$$\gamma_\uparrow^+(t) = -\int_0^t \tfrac{1}{2}(1-\cos\theta)\frac{d\phi}{dt'}dt' \tag{2.5}$$

for the Berry phase $\gamma_\uparrow^+(t)$, as θ and ϕ vary slowly over S_2. The reason for the + symbols will become clear in a moment.

According to what was advertised in Section 1, the integrand in (2.5) should be closely related to the vector potential of a magnetic monopole of strength $-1/2$, in θ–ϕ space, positioned at the origin. A standard expression for the vector potential of a Dirac monopole of this strength is

$$\boldsymbol{A}_+(\boldsymbol{r}) = \frac{-1}{2r}\,\frac{1}{z+r}\cdot(-y, x, 0), \tag{2.6}$$

where $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$. Thus with $\boldsymbol{r} = (x, y, z)$

$$\boldsymbol{A}_+(\boldsymbol{r})\cdot d\boldsymbol{r} = \frac{-1}{2r(z+r)}(x\,dy - y\,dx) = -\tfrac{1}{2}(1-\cos\theta)d\phi. \tag{2.7}$$

Hence indeed (cf. (1.11))

$$\gamma_\uparrow^+(t) = \int_0^t \boldsymbol{A}_+(\boldsymbol{n})\cdot\frac{d\boldsymbol{n}}{dt'}\,dt' \equiv \int_0^t \mathscr{L}_{\text{eff}}^+(\boldsymbol{n})dt', \tag{2.8}$$

and we see explicitly the 'monopole' character of the phase factor associated with adiabatic motion round the degeneracy point $\boldsymbol{n} = \boldsymbol{0}$. (We hope the reader will not be confused by the use of $\boldsymbol{n}$ for the slow d.f.'s in this Section, and of n as a label of a 'fast' eigenstate in the previous one.)

However, the wave function (2.3) is ill-defined at $\theta = \pi$ (what is the value of ϕ when $\theta = \pi$?). An alternative choice of $\uparrow$ wave function which is well defined at $\theta = \pi$ is

$$\langle\theta\phi|\uparrow\boldsymbol{n}-\rangle = \begin{pmatrix}\cos\theta/2e^{-i\phi}\\ \sin\theta/2\end{pmatrix}. \tag{2.9}$$

Repeating the above calculations we find that this leads to a Berry phase

$$\gamma_\uparrow^-(t) = -\int_0^t \tfrac{1}{2}(-1-\cos\theta)\,\frac{d\phi}{dt'}\,dt', \tag{2.10}$$

which is equivalent to a vector potential

$$\boldsymbol{A}_-(\boldsymbol{r}) = \frac{1}{2r}\cdot\frac{1}{z-r}(-y, x, 0). \tag{2.11}$$

Though good at $\theta = \pi$, (2.9) is ill-defined at $\theta = 0$ — and in fact we are hitting here the famous problem that, for a monopole field, no *single* vector potential exists which is singularity-free over the entire manifold S_2. The $\boldsymbol{A}_+$ which followed from the choice (2.3) has

a singularity along $z = -r$, i.e. the negative z-axis, or $\theta = \pi$. This line of singularities is called a 'Dirac string' [4]. Likewise, the $\boldsymbol{A}_-$ choice has a string along $\theta = 0$. But, comparing (2.5) and (2.10) we see that $\boldsymbol{A}_+$ and $\boldsymbol{A}_-$ differ by a gradient

$$\boldsymbol{A}_+ - \boldsymbol{A}_- = \nabla\phi, \tag{2.12}$$

that is, by a gauge transformation. Correspondingly,

$$\gamma_\uparrow^+ - \gamma_\uparrow^- = -[\phi(t) - \phi(0)], \tag{2.13}$$

so that for a closed path on S_2, $\gamma_\uparrow$ is unique.

What we see explicitly here for $\boldsymbol{A}_\pm$ is generally true. Any particular $\boldsymbol{A}$ will have a string singularity somewhere, and by doing a gauge transformation we merely shift the singularity somewhere else. The use of *two* $\boldsymbol{A}$'s (e.g. $\boldsymbol{A}_+$ and $\boldsymbol{A}_-$) was advocated by Wu and Yang [29, 30] as a way round the singularity problem, since we can use each in a region (or 'patch') where it is singularity-free, and then connect the two, in a convenient overlap region, by a gauge transformation.

The problem of singularities in the vector potential corresponding to a magnetic monopole would seem to be unavoidable since, if $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ and $\boldsymbol{A}$ is singularity-free, div $\boldsymbol{B} = 0$ and the magnetic charge must be zero. In our ultimate application of the Berry phase concept, the 'slow' d.f.'s will be the meson field variables, which we shall want to quantize. This is analogous to quantizing the $\boldsymbol{n}$ d.f.'s (i.e. θ, ϕ) in the present quantum-mechanical analogue. The presence of the (monopole) singularity at $\boldsymbol{n} = \boldsymbol{0}$ makes this quantization very awkward, and the Wu–Yang procedure is also not well-adapted to our later purpose.

Remarkably enough, however, it is possible to find a singularity-free Lagrangian for the monopole problem. Indeed, it is given to us automatically by the Berry phase formula, as we shall now describe. We then show how to obtain, from the Berry formula, the elegant Balachandran formalism [13–17], which is ideally suited to the Skyrmion application.

3. The Hopf fibration of S_2, and the Balachandran Lagrangian

Let us introduce the notation

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{3.1}$$

for the two-component spinor which is the eigenfunction of (2.1) (e.g. z could be (2.3) or (2.9)). The effective Lagrangian associated with the Berry phase is then

$$\mathscr{L}_{\text{eff}} = iz^\dagger \frac{dz}{dt} \tag{3.2}$$

(cf. (2.8) and (2.10), and check it by trying (2.3) or (2.9) for z). In the two z's considered explicitly so far ((2.3) and (2.9)) only two d.f.'s entered, namely θ and ϕ, the coordinates of a point on the surface of a two-dimensional sphere S_2. We set $\mathscr{L}_{\text{eff}} = -\boldsymbol{A} \cdot d\boldsymbol{n}/dt$ to

obtain the potentials A_+ and A_-, also on S_2. However, in principle the spinor z has *three* d.f.'s, since the normalization condition

$$z^\dagger z = 1 = |z_1|^2 + |z_2|^2 \tag{3.3}$$

is only one constraint on the two complex numbers z_1, z_2. Indeed, we may in general consider either (2.3) or (2.9) to be multiplied by an arbitrary *phase*, for example

$$z = \begin{pmatrix} \cos\theta/2 & e^{i\chi} \\ \sin\theta/2 & e^{i(\phi+\chi)} \end{pmatrix}. \tag{3.4}$$

The corresponding 'A' must now depend on three d.f.'s, and consequently is not restricted to the surface of an S_2: it turns out, as we shall now see, that it is actually defined on the surface of an S_3, and is non-singular!

Suppose we write

$$\left.\begin{aligned} z_1 &= x_1 + ix_2 \\ z_2 &= x_3 + ix_4 \end{aligned}\right\}. \tag{3.5}$$

Then $z^+z = 1$ becomes

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1, \tag{3.6}$$

and the x_i are the coordinates of a point on an S_3. Comparing (3.4) and (3.5) we find

$$\left.\begin{aligned} x_1 &= \cos\theta/2 \cos\chi, & x_2 &= \cos\theta/2 \sin\chi \\ x_3 &= \sin\theta/2 \cos(\phi+\chi), & x_4 &= \sin\theta/2 \sin(\phi+\chi) \end{aligned}\right\}. \tag{3.7}$$

The metric is

$$ds^2 = \tfrac{1}{4} d\theta^2 + d\chi^2 + \sin^2\theta/2 d\phi^2 + 2\sin^2\theta/2 d\phi d\chi. \tag{3.8}$$

It is more convenient to use orthogonal coordinates by introducing

$$\psi = \phi + \chi \tag{3.9}$$

in terms of which (3.8) becomes

$$ds^2 = \tfrac{1}{4} d\theta^2 + \cos^2\theta/2 d\chi^2 + \sin^2\theta/2 d\psi^2. \tag{3.10}$$

Thus '$A \cdot dn$' now has the form

$$A_\theta \tfrac{1}{2} d\theta + A_\chi \cos\theta/2 d\chi + A_\psi \sin\theta/2 d\psi. \tag{3.11}$$

Inserting (3.4) into (3.2) we find easily,

$$\text{'}A \cdot dn\text{'} = -iz^\dagger dz = d\chi + \tfrac{1}{2}(1-\cos\theta) d\phi \tag{3.12}$$

$$= \cos^2\theta/2 d\chi + \sin^2\theta/2 d\psi \tag{3.13}$$

whence, via (3.11),

$$A_\theta = 0, \quad A_\chi = \cos\theta/2, \quad A_\psi = \sin\theta/2. \tag{3.14}$$

These potentials are manifestly non-singular. By contrast, the 'S_2' forms (2.7), and the corresponding $\boldsymbol{A}_- \cdot d\boldsymbol{n}$ from (2.11), are singular. Consider, for example (2.7). On S_2 the metric is $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$ and so

$$A_{+,\theta} = 0,\ A_{+,\phi} = -\tfrac{1}{2}\frac{(1-\cos\theta)}{\sin\theta} = -\tfrac{1}{2}\tan\theta/2, \tag{3.15}$$

which is singular (as expected) at $\theta = \pi$. Likewise $A_{-,\phi}$ is singular at $\theta = 0$. In terms of the S_3 coordinates,

$$\boldsymbol{A}_+ \cdot d\boldsymbol{n} = -\tfrac{1}{2}(1-\cos\theta)d\phi = -\sin^2\theta/2 d\psi + \sin^2\theta/2 d\chi, \tag{3.16}$$

giving

$$A_{+,\theta} = 0, \quad A_{+,\chi} = \frac{\sin^2\theta/2}{\cos\theta/2}, \quad A_{+,\psi} = -\sin\theta/2 \tag{3.17}$$

and $A_{+,\chi}$ is singular at $\theta = \pi$. The S_3 components of $\boldsymbol{A}_-$ can be found similarly, and this time $A_{-,\psi}$ is singular at $\theta = 0$.

Thus a non-singular potential for the monopole can be found provided we enlarge the configuration space from S_2 to S_3, and use the full three d.f.'s available in z. Are we sure that the physics is really the same? The Lagrangian $\mathscr{L}_{\text{eff}}$ corresponding to (3.4) is, of course,

$$\mathscr{L}_{\text{eff}} = \mathscr{L}^+_{\text{eff}}(\boldsymbol{n}) - \dot{\chi}, \tag{3.18}$$

which differs from $\mathscr{L}^+_{\text{eff}}(\boldsymbol{n})$ by a total time derivative, and therefore leads to the same equations of motion. From (3.18) we learn that χ is acting like a U(1) gauge d.f. Thus the two d.f.'s of S_2 have been enlarged to three by the addition of a U(1) gauge d.f., χ. This is a well-known construction in mathematics, called the Hopf fibration of S_2. S_3 can be regarded as a principal fibre bundle with base space S_2 and a U(1) structure group. The (Hopf) projection map which takes us from S_3 to S_2 is given explicitly by

$$\boldsymbol{n} = z^\dagger \boldsymbol{\sigma} z, \tag{3.19}$$

as can easily be checked (Appendix B). Ryder [31] and Minami [32] were the first to introduce the Hopf map into monopole theory.

From (3.12) it is clear that the $\boldsymbol{A}_+$ potential is obtained (cf. (2.7)) by setting $\chi = 0$, and the $\boldsymbol{A}_-$ one by setting $\chi = -\phi$. Restricting χ in this way is called taking a 'section' of the fibre bundle. These two choices are each called 'local' sections, because they (and the potentials) are not smoothly defined globally over the entire S_2 : $\boldsymbol{A}_+$ is smooth for an upper patch of S_2 excluding $\theta = \pi$, and $\boldsymbol{A}_-$ is smooth for a lower patch excluding $\theta = 0$. It is, in fact, not possible to find any such section which is smooth globally, in this case: a minimum of two is required, as in the explicit examples of $\boldsymbol{A}_\pm$. Mathematically this corresponds to the fact that our (monopole) bundle is non-trivial, or — equivalently — to the fact that S_3 is only locally, but not globally, equivalent to $S_2 \times S_1$. Thus the monopole Lagrangian can be described in a singularity-free way by using a non-trivial bundle over S_2.

The above formulation is not yet quite suitable for our later application to Skyrmion physics. In that case, the d.f.'s in which we shall be interested are actually entries in an SU(3) matrix, and it is hard to see how to generalize z to such a matrix. On the other hand, S_3 is the group manifold of SU(2), and it is quite simple to reformulate the above results in terms of a basic dynamical variable $s(\theta, \phi, \chi) \in \mathrm{SU}(2)$, rather than $z(\theta, \phi, \chi)$. This will lead to Balachandran's form for $\mathscr{L}_{\mathrm{eff}}$, which will be directly analogous to the SU(3) case.

We can associate a general SU(2) matrix s with the components z_1, z_2 of z via

$$s = \begin{pmatrix} z_1 & -z_2^* \\ z_2 & z_1^* \end{pmatrix} \tag{3.20}$$

since the condition $|z_1|^2 + |z_2|^2 = 1$ guarantees $s^\dagger s = s s^\dagger = 1$. In terms of s, the $\mathscr{L}_{\mathrm{eff}}$ of (3.2) becomes

$$\mathscr{L}_{\mathrm{eff}} = i z^\dagger \frac{dz}{dt} = \frac{i}{2} \operatorname{tr} (\sigma_3 s^{-1} \dot{s}) \tag{3.21}$$

as may be verified explicitly. Equation (3.21) provides our desired (Balachandran) monopole Lagrangian in terms of $s \in \mathrm{SU}(2)$. It is pleasing to see this direct link between the Berry phase and the Balachandran Lagrangian.

The Hopf map can equivalently be described in terms of s. The counterpart of (3.19) is

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = s \sigma_3 s^{-1} \tag{3.22}$$

(note that $\boldsymbol{n}^2 = 1$ follows automatically upon squaring both sides). Under right multiplication of s by an element of U(1)

$$s \to s \exp i\sigma_3 \alpha, \tag{3.23}$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} \to s(\exp i\sigma_3\alpha)\sigma_3(\exp -i\sigma_3\alpha)s^{-1} = s\sigma_3 s^{-1} = \boldsymbol{\sigma} \cdot \boldsymbol{n} \tag{3.24}$$

and $\boldsymbol{n}$ is unchanged. Thus the space SU(2)/U(1) of right cosets (3.23) gets mapped by (3.22) into S_2 (see Fig. 2).

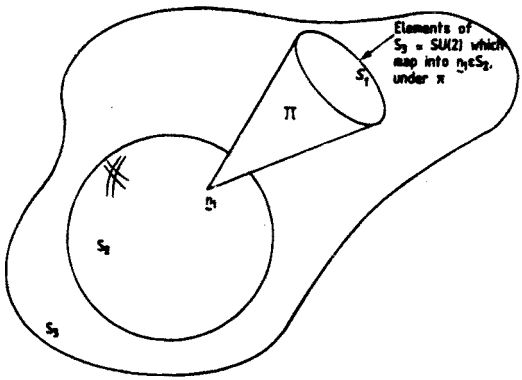

Fig. 2. The one-cycle S_1 of all points in S_3 related to a given point s_1 of S_3 by $s_1 \exp(i\sigma_3\alpha)$, as α varies, is mapped by the Hopf map Π into the single point $\boldsymbol{n}_1$ of S_2

To gain some confidence with this s-formalism, we can simply insert (2.3) into (3.20), obtaining

$$s_+(\boldsymbol{n}) = \begin{pmatrix} \cos\theta/2 & -\sin\theta/2e^{-i\phi} \\ \sin\theta/2e^{i\phi} & \cos\theta/2 \end{pmatrix} \tag{3.25}$$

whence

$$\mathscr{L}^+_{\text{eff}}(\boldsymbol{n}) = \frac{i}{2}\,\text{tr}\,(\sigma_3 s_+^{-1}\dot{s}_+) = \frac{-(1-\cos\theta)}{2}\,\frac{d\phi}{dt} \tag{3.26}$$

in agreement with (2.8). Alternatively, (2.9) gives

$$s_-(\boldsymbol{n}) = \begin{pmatrix} \cos\theta/2e^{-i\phi} & -\sin\theta/2 \\ \sin\theta/2 & \cos\theta/2e^{i\phi} \end{pmatrix} \tag{3.27}$$

so that

$$s_+(\boldsymbol{n}) = s_-(\boldsymbol{n})e^{i\sigma_3\phi} \tag{3.28a}$$

$$s_+\sigma_3 s_+^{-1} = s_-\sigma_3 s_-^{-1} = \boldsymbol{\sigma}\cdot\boldsymbol{n} \tag{3.28b}$$

$$\mathscr{L}^-_{\text{eff}}(\boldsymbol{n}) = \frac{i}{2}\,\text{tr}\,(\sigma_3 s_-^{-1}\dot{s}_-) = \frac{-(-1-\cos\theta)}{2}\,\frac{d\phi}{dt} \tag{3.28c}$$

$$\mathscr{L}^+_{\text{eff}}(\boldsymbol{n}) - \mathscr{L}^-_{\text{eff}}(\boldsymbol{n}) = -\dot{\phi}. \tag{3.28d}$$

Equation (3.28a) shows that, for given $\boldsymbol{n}$, s_+ and s_- are in the same coset, and get mapped into the same $\boldsymbol{n}$ (3.28b). Equation (3.28d) shows that the corresponding effective Lagrangians differ by a total derivative, and hence lead to the same equations of motion for the $\boldsymbol{n}$ d.f.'s. Indeed, the difference between the choice s_+ and s_- corresponds exactly to the gauge transformation on the associated vector potentials $\boldsymbol{A}_+$ and $\boldsymbol{A}_-$ considered earlier in (2.12).

In general, we may consider now the SU(2) matrix

$$\begin{aligned} s(\theta,\phi,\chi) &= s_+(\theta,\phi)e^{i\sigma_3\chi} \\ &= \begin{pmatrix} \cos\theta/2e^{i\chi} & -\sin\theta/2e^{-i(\phi+\chi)} \\ \sin\theta/2e^{i(\phi+\chi)} & \cos\theta/2e^{-i\chi} \end{pmatrix}, \end{aligned} \tag{3.29}$$

corresponding to (3.4). Then s_+ is the $\chi = 0$ section of this, while s_- is the $\chi = -\phi$ section. The Lagrangian following from (3.29) is, of course,

$$\frac{i}{2}\,\text{tr}\,(\sigma_3 s_+^{-1}\dot{s}_+) - \dot{\chi} \tag{3.30}$$

as in (3.18).

In concluding this Section we note (see [31] and [32]) that the foregoing can all be rephrased using the compact formalism of differential forms, and some elementary ideas

of homology and cohomology. The potential 1-form is

$$A = iz^\dagger dz = -x_1 dx_2 + x_2 dx_1 - x_3 dx_4 + x_4 dx_3, \tag{3.31}$$

and the field 2-form B is

$$B = dA = idz^\dagger \wedge dz = -2(dx_1 \wedge dx_2 + dx_3 \wedge dx_4) = -\tfrac{1}{2} \sin\theta d\theta \wedge d\phi, \tag{3.32}$$

where (3.7) has been used. B is just proportional to the area 2-form of S_2, and

$$\int_{S_2} B = -\tfrac{1}{2} 4\pi, \tag{3.33}$$

showing that these potentials and fields indeed correspond to a monopole of strength 1/2. B is certainly closed,

$$dB = 0$$

but it cannot be exact ($B = dA$) on S_2, since if it were we could use Stokes' theorem on S_2 to obtain

$$\int_{S_2} B = \int_{S_2} dA = \int_{\partial S_2} A = 0, \tag{3.34}$$

since S_2 has no boundary; (3.34) would then contradict (3.33). However, if B is regarded as a 2-form on S_3 it is exact, since $H^2(S_3) = 0$, and consequently an A such that $B = dA$ does exist on S_3.

4. Quantization of the n d.f.'s: the Dirac condition again

We now want to consider, following Balachandran et al. [13, 17], the problem of *quantizing* the d.f.'s $\boldsymbol{n}(\theta, \phi)$ — i.e. we want to promote the 'slow' d.f.'s, which hitherto in Sections 2 and 3 have been parameters, to dynamical variables. In path integral terms, this means — cf. (1.14) — that we want to consider

$$\int \mathscr{D}\boldsymbol{n} \exp\{i \int [\tfrac{1}{2} \dot{\boldsymbol{n}}^2 + \mathscr{L}_{\text{eff}}(\boldsymbol{n})] dt\}. \tag{4.1}$$

We know that the quantum theory of the charge-monopole system should only be consistent provided the Dirac condition holds. We are going to see where this arises in the s-formalism.

In the previous Section we have seen that the introduction of the new (SU(2)) d.f. χ allowed us to describe the monopole system by a non-singular Lagrangian — and so in quantizing this system we do not have the problem of singularities to contend with. On the other hand, we want the physics to be independent of χ. In the classical theory, as we have seen χ acts like a U(1) gauge d.f., and changing χ is like doing a gauge transformation, under which the equations of motion are invariant. In the quantum theory, we must ensure that a corresponding gauge invariance is correctly implemented. This requirement leads to the Dirac condition.

It is clear that the first term, $1/2\dot{\boldsymbol{n}}^2$, in the Lagrangian of (4.1) is *invariant* under a U(1) gauge transformation

$$s \to s e^{i\sigma_3 \alpha(t)} \tag{4.2}$$

since $\boldsymbol{n}$ remains invariant under (4.2) (see also Appendix C). Thus non-trivial constraints on the theory, associated with the implementation of gauge invariance under (4.2), must arise from the second ('monopole') term. Let us consider a general such term

$$\mathscr{L}_{\text{eff}}(\boldsymbol{n}) = -gi \operatorname{tr}(\sigma_3 s^{-1}\dot{s}), \tag{4.3}$$

where the monopole strength g is not yet determined. Then, under (4.2),

$$\mathscr{L}_{\text{eff}} \to \mathscr{L}_{\text{eff}} + 2g\dot{\alpha}. \tag{4.4}$$

In the quantum theory, s will be promoted to a quantum variable $\hat{s}$, and wave functions will be written as $\Psi(s)$. Consider the infinitesimal (quantum) version of (4.2):

$$\hat{s} \to \hat{s} + i\hat{s}\sigma_3 \delta\hat{\alpha}, \tag{4.5}$$

and let $\hat{G}$ be the generator of this transformation so that

$$[\hat{G}, \hat{s}] = \hat{s}\sigma_3. \tag{4.6}$$

Then, from Noether's theorem and (4.4), we deduce

$$\hat{G}\Psi = 2g\Psi \tag{4.7}$$

as a consistency condition on the state functions (it is a kind of 'Gauss Law' associated with gauge invariance under (4.2); see also Appendix C). For finite transformations we then have

$$\Psi'(s) \equiv (e^{i\hat{G}\hat{\alpha}}\Psi)(s) = \Psi(s e^{i\sigma_3\alpha}) = e^{2ig\alpha}\Psi. \tag{4.8}$$

The last two equalities of (4.8) give

$$\Psi(\theta, \phi, \chi+\alpha) = e^{2ig\alpha}\Psi(\theta, \phi, \chi), \tag{4.9}$$

which enforces a kind of 'Bloch' condition on the χ d.f. If we consider the particular case $\alpha = 2\pi$, then since

$$e^{2\pi i\sigma_3} = 1 \tag{4.10}$$

we deduce

$$e^{4\pi ig} = 1 \tag{4.11}$$

and hence

$$g = 0, \pm\tfrac{1}{2}, \pm 1, \ldots, \tag{4.12}$$

which is precisely the Dirac condition. Equation (4.9) is called an 'equivariance' condition on the wave function Ψ: in going from the S_2 of (θ, ϕ) to the S_3 of (θ, ϕ, χ) we have enlarged

the configuration space over which our wave functions are to be defined, but an *arbitrary* dependence on the additional variable χ is not consistent with the required gauge invariance (dynamical independence) with respect to χ. Only Ψ's satisfying (4.9) are allowed, with g satisfying (4.12). And, of course, our basic spinor

$$\begin{pmatrix} \cos\theta/2 & e^{i\chi} \\ \sin\theta/2 & e^{i(\phi+\chi)} \end{pmatrix} \tag{4.13}$$

does satisfy (4.12) with $g = -1/2$, the minimum non-trivial magnitude.

A general wave function $\Psi(s)$ can be expressed as a linear combination of the 'top' functions $\mathscr{D}^j_{m'm}(\theta, \phi, \chi)$, which carry irreducible representations of SU(2). It seems obvious from the fact that θ, ϕ, and χ are angles that j should indeed be the angular momentum quantum number; for those who doubt, some further discussion is given in Appendix D. Then

$$\Psi(s) = \sum_{j,m',m} c^j_{m'm} \mathscr{D}^j_{m'm}(\theta, \phi, \chi). \tag{4.14}$$

The constraint (4.9) must now be imposed. If we multiply s from the right by $\exp(i\sigma_3\alpha)$, the $\mathscr{D}$'s get changed by

$$\mathscr{D}^j_{m'm}(s \exp i\sigma_3\alpha) = e^{2\alpha m i} \mathscr{D}^j_{m'm}(s), \tag{4.15}$$

since m is the eigenvalue of $\sigma_3/2$. Thus from (4.15) and (4.9),

$$m = g = 0, \pm\tfrac{1}{2}, \pm 1, \ldots \tag{4.16}$$

and the possibility of 1/2-odd integral spin has emerged (since j is 1/2-odd integral if $2m$ is odd and integral if $2m$ is even). In fact, as stated in Section 1, the system has 1/2-odd angular momentum if the monopole strength g has the value $(n+1/2)$, for integer n.

5. The Skyrmion case

Our basic analogy is as follows:

$$\left.\begin{array}{rl} \text{fast d.f.'s : fermion Fock states} & \sim \text{spin states } |\uparrow\rangle, |\downarrow\rangle \\ \text{fermion vacuum } |0\rangle & \sim \text{spin state } |\uparrow\rangle \\ \text{slow d.f.'s: Goldstone boson fields } \phi_a & \sim \text{angular variables } \boldsymbol{n} \\ |0, \phi_a\rangle & \sim |\uparrow \boldsymbol{n}\rangle \end{array}\right\} \tag{5.1}$$

Just as a monopole structure appeared in the Berry phase associated with $|\uparrow, \boldsymbol{n}\rangle$ for slowly varying $\boldsymbol{n}$, so the W–Z term in the bosonic action is interpreted as a kind of Berry phase for $|0, \phi_a\rangle$.

We begin by introducing the commonly-used notation for the ϕ fields. In the case of $SU(2)_f$, we would have four ϕ's, written as $\phi = (\sigma, \boldsymbol{\pi})$, where $\sigma^2 + \boldsymbol{\pi}^2 = f^2$, quite analo-

gously to $\boldsymbol{n}^2 = 1$. However, this does not generalize to the required $SU(3)_f$ case. Instead, we first rewrite ϕ as

$$\phi = fU, \tag{5.2}$$

where

$$U = \exp(i\tau \cdot \pi/f) \tag{5.3}$$

is a unitary 2×2 matrix. This amounts to a reparametrization of the original σ, π in the expression $\phi = (\sigma,\pi)$. In $SU(3)_f$, (5.3) is generalized to (cf. (1.22))

$$U = \exp(i\lambda \cdot \pi/f), \tag{5.4}$$

where π is understood now to be an 8-component 'angle-type' field. The analogue of (4.1) is then

$$\int \mathscr{D}U \exp\{i \int \mathscr{L}_0(U)dt\} \exp iS_{W\text{-}Z}(U), \tag{5.5}$$

where $\mathscr{L}_0$ is all the rest of the Lagrangian for the U fields, apart from the W–Z term; for example,

$$\mathscr{L}_0 = \tfrac{1}{4} \operatorname{tr}(\partial_\mu U^\dagger \partial^\mu U) + \ldots, \tag{5.6}$$

where the dots represent other terms which are necessary to stabilize the soliton, for instance. Finally, the expression for the W–Z action is [8]

$$\exp iS_{W\text{-}Z} = \exp \frac{-iN_c}{240\pi^2} \int \varepsilon^{ijklm} \operatorname{tr}(U^\dagger \partial_i U\, U^\dagger \partial_j U\, U^\dagger \partial_k U\, U^\dagger \partial_l U\, U^\dagger \partial_m U) d^5x. \tag{5.7}$$

The integral in (5.7) is over a 5-dimensional 'disc' whose boundary is 4-dimensional Minkowskian space-time. This disc is the 5-dimensional analogue of the 2-dimensional surfaces considered in (1.19)–(1.21), and N_c is the analogue [8] of the monopole N in (1.21).

Now, we seem a long way from anything like the Balachandran monopole Lagrangian (3.21). However, we can actually make the connection quite explicit, as follows [16, 17]. Instead of treating the *full* quantum-mechanical problem (5.5), in which the whole of the U matrix is treated as a quantum field variable, we perform only a 'semi-classical' quantization. In such an approach, one starts from a solution $U_c(\boldsymbol{r})$ of the static classical field equations in the $SU(2)_f$ case, which is of standard Skyrmion type (cf. (5.3) with $\pi = f\hat{\boldsymbol{r}}\theta(r)$):

$$U_c(\boldsymbol{r}) = \cos\theta(r) + i\tau \cdot \hat{\boldsymbol{r}} \sin\theta(r). \tag{5.8}$$

It is clear that this solution is not rotationally invariant, nor is it invariant under isospin rotations. In fact, there are infinitely many such solutions, related to one another by spatial or isospin rotations, all of which are degenerate in energy since the original Lagrangian *is* invariant under space or isospin rotations. Actually these two kinds of rotation are effectively equivalent for (5.8), since

$$s\tau_i s^{-1} = \tau_j R_{ji}(s) \tag{5.9}$$

for $s \in \mathrm{SU}(2)$. The coordinates which distinguish these degenerate classical configurations are the parameters of the matrix s. The semi-classical quantization procedure consists in promoting these d.f.'s into quantum variables $s(t)$. Thus we write

$$U(\boldsymbol{r}, t) = s(t)U_c(\boldsymbol{r})s^{-1}(t). \tag{5.10}$$

Classical quantities will now have a subscript c, and quantum d.f.'s will be distinguished by having no subscript c, instead of by having a '$\hat{}$'. The $s(t)$ will behave just like the s of Sections 3 and 4.

We must now extend (5.8) and hence (5.10) to the $\mathrm{SU}(3)_f$ case, or else we get no W–Z term at all [8] (see further Section 6). This means that we have to 'embed' (5.8) inside an SU(3) matrix. The obvious way to do this would seem to be

$$U_c \to \begin{pmatrix} \cos\theta(r) + i\boldsymbol{\tau}\cdot\hat{\boldsymbol{r}}\sin\theta(r) & 0 \\ 0 & 1 \end{pmatrix} \equiv \tilde{U}_c(\boldsymbol{r}) \tag{5.11}$$

(alternative embeddings, which have different physical consequences, are discussed in Refs [12], [16] and [17]). So now,

$$U(\boldsymbol{r}, t) = s(t)\tilde{U}_c(\boldsymbol{r})s^{-1}(t),\ s \in \mathrm{SU}(3)_f. \tag{5.12}$$

We observe at once that $U(\boldsymbol{r}, t)$ is invariant under

$$s \to se^{iY\alpha(t)}, \tag{5.13}$$

where

$$Y = \tfrac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{5.14}$$

(the normalization is, of course, chosen for convenience). Thus, the configuration space for the s d.f.'s is not SU(3) but rather $\mathrm{SU}(3)/\mathrm{U}(1)_Y$; this is exactly analogous to our monopole example, where the required configuration space was SU(2)/U(1). In fact, the analogy is very close indeed, for when (5.12) is inserted into (5.7) one finds — after some calculation — that the term involving $\dot{s}$ (i.e. the piece involving the quantum d.f.'s, in this approximation) is just [16, 17]

$$\mathscr{L}_{\mathrm{W\text{-}Z}} = -\tfrac{1}{2}N_c B(U_c)\,\mathrm{tr}\,(Ys^{-1}\dot{s}), \tag{5.15}$$

where $B(U_c)$ is the winding number (= baryon number) of the classical configuration U_c. Equation (5.15) should be compared with (3.21).

We see, from this comparison, that indeed the W–Z term is acting so as to produce, in this semi-classical quantization, exactly a 'monopole in SU(3) space'. The procedure of Section 4 can be transcribed easily to SU(3). The gauge invariance analogous to (4.2) is the invariance of (5.13), under which, however, $\mathscr{L}_{\mathrm{W\text{-}Z}}$ changes according to

$$\mathscr{L}_{\mathrm{W\text{-}Z}} \to \mathscr{L}_{\mathrm{W\text{-}Z}} + \tfrac{1}{3}N_c B\dot{\alpha}. \tag{5.16}$$

In the quantized theory, s and Y are operators, and from Noether's theorem (corresponding to (4.7)) we have

$$\hat{Y}\Psi = \tfrac{1}{3} N_c B \Psi. \tag{5.17}$$

What is the analogue of the quantization constraint (4.12)? For this we note [17] that if we replace s by sh in (5.12), with $h \in \mathrm{SU}(2)$, this is equivalent to rotating U_c by some spatial rotation parametrized by h (cf. 5.9)). In particular, consider a rotation by 2π about the 3rd axis. This corresponds to

$$h = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.18}$$

and thus to the replacement

$$s \to s \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = s e^{3\pi i Y}. \tag{5.19}$$

According to (5.17), the allowed Ψ's must then pick up a phase factor

$$e^{i\pi N_c B}, \tag{5.20}$$

and hence *the allowed states for $B = 1$ are fermions if N_c is odd, and bosons if N_c is even*! [9].

The wave functionals Ψ are the SU(3) generalization of the SU(2) rotation functions $\mathscr{D}^{J}_{m'm}$ [s ∈ SU(2)] — namely

$$\mathscr{D}^{p,q}_{I,I_3,Y;I',I_3'Y'}(s \in \mathrm{SU}(3)), \tag{5.21}$$

where p and q label the irreducible representation of SU(3). In (5.21) the left-hand group of 'magnetic quantum numbers' refers to transformation properties under *left* multiplication of s by a matrix in SU(3), and hence (cf. (5.12)) to a flavour rotation of U; the right-hand indices refer to right multiplication. But we have already seen that the SU(2) part — in the sense of (5.11) — of any 'right multiplication' matrix corresponds to a spatial rotation. Hence I' and I_3' are actually the real spin and its third component. Now for $B = 1$ and $N_c = 3$ we need the eigenvalue $Y' = 1$ from (5.17). The lowest dimensionality SU(3) representations with $Y' = 1$ are the **8** and **10** (Fig. 3). In the former, the states with $Y' = 1$ have $I' = 1/2$, and hence spin 1/2, while in the latter they have spin 3/2. The left-hand indices give just the flavour quantum numbers corresponding to these SU(3) representations: thus we have an **8** of spin 1/2 and a **10** of spin 3/2.

Further details of Skyrmion quantization are given in Guadagnini [11] and Rabinovici et al. [12]: our concern here has been to place the 'monopole' form (5.15) of $\mathscr{L}_{\text{W-Z}}$ in the context of an adiabatic decoupling problem. From this point of view, the peculiar phase behaviour leading to 'fermion-ness' in the ϕ sector has arisen as a result of non-trivial structure left behind when the fermion vacuum is decoupled adiabatically from

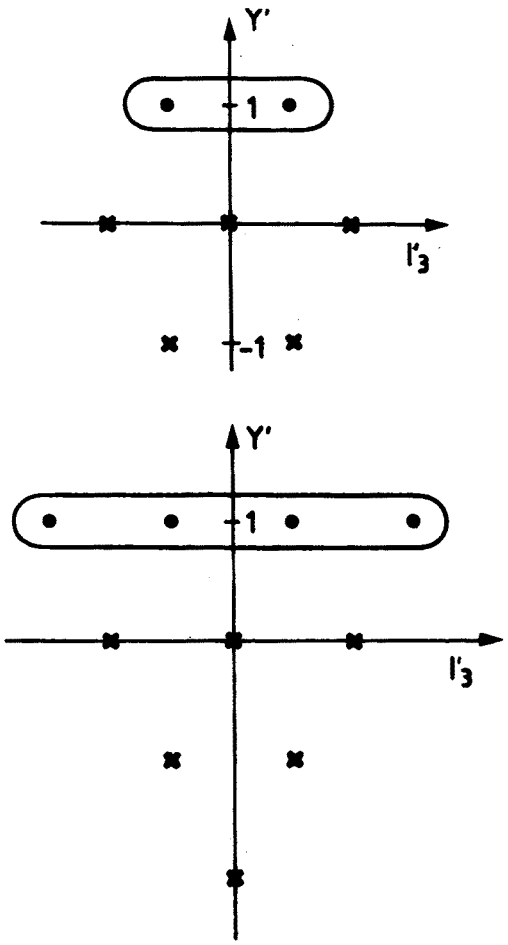

Fig. 3. The **8** and **10** representations of SU(3), showing the two allowed multiplets with $Y' = 1$

the ϕ's. If we use only the ϕ d.f.'s, and integrate the fermions away, we must include a W – Z term which embodies this structure. The ultimate reason that this structure has a 'monopole' form is to be found in the topological approach to anomalies [27, 28].

We may also remark that a similar mechanism holds for Skyrmions in 2+1 dimensions [33]. Here the Wess–Zumino term is replaced by the Hopf term [34–37], which can also be interpreted as arising from integrating out fermions [35]. When the Skyrmion is quantized semi-classically, the angular momentum has the value (*integer* + $\theta/2\pi$), where θ is the coefficient of the Hopf term [38]. The effective Lagrangian in this approximation is exactly analogous to that describing a charged particle moving in two dimensions in the field of a magnetic vortex lying perpendicular to the plane of motion. In that case, the angular momentum can have a value neither integral nor half-odd integral, with consequential 'fractional statistics' [39–41]. Thus just as, in 3+1, the W–Z term acts as a monopole in field space, so in 2+1 the Hopf term acts as a vortex in field space.

6. *Postscript: the case of only two flavours*

The above discussion has been predicated upon the existence of the W–Z term — whose presence determines the quantization of the Skyrmion (fermion if N_c odd, boson if N_c even). But if there are only two flavours, there is no W – Z term: when (5.3) is substituted into (5.7) the SU(2) trace vanishes[1]. Yet there are topological solitons since $\pi_3(\mathrm{SU}(2)) = \mathrm{Z}$. What determines their quantization?

The answer is that the $B = 1$ soliton can be quantized *either* as a boson *or* as a fermion: there is no restriction involving N_c, and one has to choose the fermionic option by hand [42]. The way in which fermionic quantization is *possible* (but not required) was discussed by Finkelstein [43], Finkelstein and Rubinstein [44], and Williams [45]. One way of putting it is as follows [46]. Since $\pi_4(\mathrm{SU}(2)) = \mathrm{Z}_2$, time-dependent soliton fields U fall into two distinct homotopy classes of maps from (compactified) space-time to SU(2). Functional integrals over the U's can therefore be separated into two topologically disjoint sectors (analogous to θ-vacua in QCD), corresponding to those U's which can be continuously deformed to the identity, and those which cannot. The contribution from these two sectors to the functional propagator can have a relative + sign or a relative − sign: in the former case the propagator contains all integral spins (bosonic), in the latter half-integral ones (fermionic).

This situation is mathematically the same as that of the spherical top [3], since $\pi_4(\mathrm{O}(3)) = \mathrm{Z}_2$ also, and the same boson/fermion option therefore exists. In this case one can say, alternatively, that since O(3) is doubly-connected, wave functions on O(3) need not be single-valued. One can define single-valued wave functions by passing to the universal covering space SU(2), but then one has to project back to O(3) via $\mathrm{SU}(2) \to \mathrm{O}(3) \simeq \mathrm{SU}(2)/\mathrm{Z}_2$, on which a double-valuedness can appear.

For $N_f > 2$, $\pi_4(\mathrm{SU}(N_f)) = 0$ and so this boson/fermion *option* is removed. But then, since $\pi_5(\mathrm{SU}(N_f)) = \mathrm{Z}$ we have the W – Z addition to the Lagrangian, and the N_c-related quantization is determined.

I am grateful to Jo Zuk for many very helpful discussions; and to Stephen Wilkinson for patient instruction in some of the relevant mathematics, and for carefully reading the manuscript. It is a pleasure to take this opportunity of thanking Drs M. Praszałowicz and W. Słomiński for organising such a stimulating and enjoyable School, and for their warm hospitality.

APPENDIX A

Monopole strength and (U)1 winding number

We have seen that the monopole strength g in

$$\mathscr{L}_{\mathrm{eff}}(\boldsymbol{n}) = gi\,\mathrm{tr}\,(\sigma_3 s^{-1}\dot{s}) \tag{A.1}$$

[1] Alternatively [8], in SU(2) G-parity invariance forbids amplitudes with an odd number of pions, while (5.7) would, if it were non-vanishing, allow them; in SU(3), (5.7) allows $\mathrm{K}\bar{\mathrm{K}} \to 3\pi$, which is not forbidden by G-parity.

is restricted to the values $g = p/2$ where $p = 0, \pm 1, \pm 2, \ldots$ In this Appendix we will show how p can be interpreted as a winding number associated with the U(1) gauge transformation (4.2).

Consider a *sequence* of gauge transformations

$$s \to s \exp i\sigma_3 \alpha(t) \tag{A.2}$$

parametrized by t, where

$$\alpha(t = 0) = 0, \quad \alpha(t = T) = 2\pi, \tag{A.3}$$

so that we have a closed loop in s-space,

$$s(t = 0) = s(t = T). \tag{A.4}$$

Then as we move through this sequence of t-values, the parameter α of the U(1) gauge group goes once round its circle (Fig. A1a):

Corresponding to the gauge transformation (A.2), we have the transformation

$$z = \begin{pmatrix} \cos \theta/2 e^{i\chi} \\ \sin \theta/2 e^{i(\phi+\chi)} \end{pmatrix} \to e^{i\alpha(t)} z \tag{A.5}$$

of the basic $\uparrow$ spinor (cf. (3.6) and (3.12)). Thus $\alpha(t)$ is just a variable phase for the associated spinor, and as we go round the sequence of gauge transformations in Fig. A1a, this phase swings round precisely once (Fig. A1b). Meanwhile, what is happening to $\mathscr{L}_{\text{eff}}(n)$? This becomes

$$_{\text{eff}}(n) \to \mathscr{L}_{\text{eff}}(n) - p\dot{\alpha}, \tag{A.6}$$

where $p = 2g$. The associated effective *action* therefore changes by

$$\exp -i \int_0^T p\dot{\alpha} dt = \exp(-2\pi i p), \tag{A.7}$$

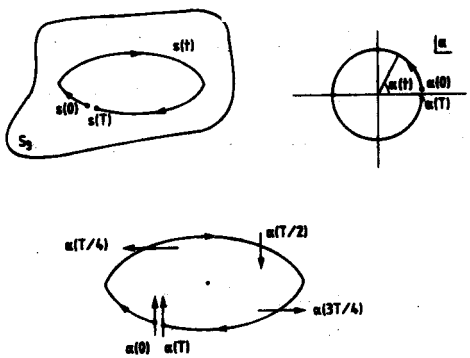

Fig. A1. Sequence of gauge transformations corresponding to a closed loop in s-space, and the associated variation of the spinor phase

i.e. its phase swings round p times as we follow the circuit of Fig. A1a. We can therefore interpret p as a winding number which counts the number of rotations of the action phase as we circulate once in α-space (i.e. one circuit in U(1) space).

APPENDIX B

More on the Hopf map

In Section 3 we gave two forms of the Hopf map, one in terms of z

$$\boldsymbol{n} = z^{\dagger}\boldsymbol{\sigma}z, \tag{B.1}$$

and the other in terms of s

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} = s\sigma_3 s^{-1}, \tag{B.2}$$

where

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \qquad |z_1|^2 + |z_2|^2 = 1 \tag{B.3}$$

and

$$s = \begin{pmatrix} z_1 & -z_2^* \\ z_2 & z_1^* \end{pmatrix}. \tag{B.4}$$

We make the connection between (B.1) and (B.2) as follows. Let us write $z_1 = x_1 + ix_2$, $z_2 = x_3 + ix_4$. Then from (B.4)

$$n_1 = (x_1 - ix_2 \quad x_3 - ix_4)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 + ix_2 \\ x_3 + ix_4 \end{pmatrix} = 2(x_1x_3 + x_2x_4) \tag{B.5}$$

and

$$n_2 = 2(x_1x_4 - x_2x_3) \tag{B.6}$$

$$n_3 = x_1^2 + x_2^2 - x_3^2 - x_4^2, \tag{B.7}$$

while

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1. \tag{B.8}$$

On the other hand, for our s of (3.12), connected to the parameters θ, ϕ of $\boldsymbol{n}$, we had

$$z_1 = \cos\theta/2 e^{i\chi}: \; x_1 = \cos\theta/2 \cos\chi, \qquad x_2 = \cos\theta/2 \sin\chi \tag{B.9}$$

$$z_2 = \sin\theta/2 e^{i(\phi+\chi)}: \; x_3 = \sin\theta/2 \cos(\phi+\chi), \qquad x_4 = \sin\theta/2 \sin(\phi+\chi), \tag{B.10}$$

whence from (B.5)–(B.7)

$$\left.\begin{aligned} n_1 &= \sin\theta\cos\phi \\ n_2 &= \sin\theta\sin\phi \\ n_3 &= \cos\theta \end{aligned}\right\} \tag{B.11}$$

as required.

The matrix s has a simple geometrical interpretation. Consider first the case of

$$s_+(\theta, \phi) = \begin{pmatrix} \cos \theta/2 & -\sin \theta/2 e^{-i\phi} \\ \sin \theta/2 e^{i\phi} & \cos \theta/2 \end{pmatrix}. \tag{B.12}$$

Let $\hat{\boldsymbol{u}}$ be the unit vector

$$\hat{\boldsymbol{u}} = (-\sin \phi, \cos \phi, 0) \tag{B.13}$$

and consider

$$\exp(-i\boldsymbol{\sigma} \cdot \hat{\boldsymbol{u}}\theta/2) = \cos \theta/2 - i \sin \theta/2 \boldsymbol{\sigma} \cdot \hat{\boldsymbol{u}} \tag{B.14}$$

$$= s_+(\theta, \phi). \tag{B.15}$$

This is a rotation of θ about $\hat{\boldsymbol{u}}$, which rotates $\hat{z}$ into $\boldsymbol{n}$ (Fig. B1). So Eq. (B.2), which is equivalent to

$$s_+^{-1} \boldsymbol{\sigma} \cdot \boldsymbol{n} s_+ = \sigma_3 \tag{B.16}$$

in this case, means simply that $\boldsymbol{n}$ has been rotated to be along the 3rd axis. The remaining factor $\exp(i\sigma_3 \chi)$ in (3.12) is then just a rotation about the 3 axis (the 'body' axis).

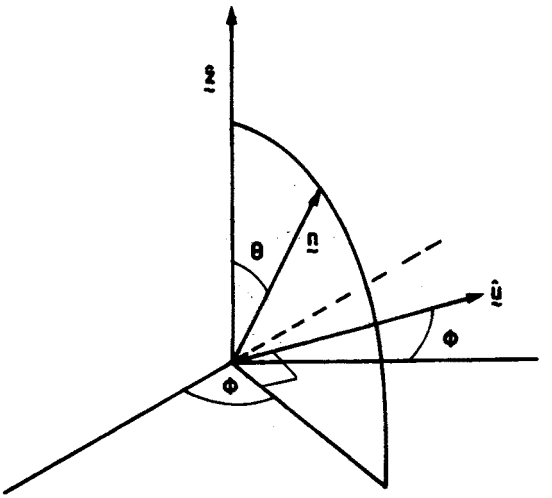

Fig. B1. A rotation of θ about $\hat{\boldsymbol{u}}$ rotates $\hat{z}$ into $\hat{n}$

There is yet one more way of writing the connection between z (or s) and $\boldsymbol{n}$ which the Hopf map enforces. It is

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = 2zz^\dagger - 1. \tag{B.17}$$

This is easy to verify: using (B.1) for the components of $\boldsymbol{n}$ in terms of z_1, z_2, we find

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = \begin{pmatrix} |z_1|^2 - |z_2|^2 & 2z_2^* z_1 \\ 2z_1^* z_2 & |z_2|^2 - |z_1|^2 \end{pmatrix} = 2zz^\dagger - 1, \tag{B.18}$$

with the help of $|z_1|^2 + |z_2|^2 = 1$.

APPENDIX C

An alternative quantization for monopoles [12]

Let us consider the monopole action in (4.1),

$$S = \int [\tfrac{1}{2}\dot{\boldsymbol{n}}^2 + \tfrac{1}{2} ip \operatorname{tr}(\sigma_3 s^{-1}\dot{s})]dt \tag{C.1}$$

$$= \int [\tfrac{1}{2}\dot{\boldsymbol{n}}^2 + ipz^\dagger \dot{z}]dt, \tag{C.2}$$

where the second step follows from (3.21), and g has been replaced by $p/2$ (Appendix A). We have not so far considered explicitly the *first* ('kinetic energy') term of (C.2) — let us attend to it now.

We have

$$\tfrac{1}{2}\dot{\boldsymbol{n}}^2 = \tfrac{1}{4}\operatorname{tr}(\boldsymbol{\sigma}\cdot\dot{\boldsymbol{n}})^2 \tag{C.3}$$

$$= 2[(\dot{z}^\dagger\dot{z}) + (z^\dagger\dot{z})^2], \tag{C.4}$$

using (B.17). Let us write

$$a = iz^\dagger\dot{z}; \tag{C.5}$$

then

$$\tfrac{1}{2}\dot{\boldsymbol{n}}^2 = 2\dot{z}^\dagger\dot{z} - 2a^2. \tag{C.6}$$

But also

$$\operatorname{tr}(\dot{s}^\dagger\dot{s}) = 2\dot{z}^\dagger\dot{z}, \tag{C.7}$$

by direct verification. Hence finally we can write this part of the action as

$$\int [\operatorname{tr}(\dot{s}^\dagger\dot{s}) - 2a^2]dt. \tag{C.8}$$

Consider now the behaviour of (C.8) under the U(1) gauge transformation:

$$z \to e^{i\alpha(t)}z, \tag{C.9}$$

which also corresponds to

$$s \to se^{i\alpha(t)\sigma_3}. \tag{C.10}$$

Under (C.9),

$$a \to a - \dot{\alpha}, \tag{C.11}$$

so that

$$-2a^2 \to -2a^2 + 4a\dot{\alpha} - 2\dot{\alpha}^2. \tag{C.12}$$

On the other hand, under (C.10) one finds easily

$$\operatorname{tr}(\dot{s}^\dagger\dot{s}) \to \operatorname{tr}(\dot{s}^\dagger\dot{s}) - 4a\dot{\alpha} + 2\dot{\alpha}^2. \tag{C.13}$$

Thus (C.9) and (C.10) are together an invariance of (C.8), as we stated in Section 4.

We can bring this invariance out by rewriting (C.8) as

$$\int \operatorname{tr}\{[(\vec{\partial}_t - ia\sigma_3)s^\dagger]\,[s(\overleftarrow{\partial}_t + ia\sigma_3)]\} \tag{C.14}$$

as can be simply checked, recalling that $a = \frac{1}{2}\,\mathrm{tr}\,(\sigma_3 s^{-1}\dot{s})$ also. Expression (C.14) is manifestly invariant under the combined transformations (C.9) and (C.10). Thus we are interested in the generating functional

$$Z = \int \mathscr{D}s \exp i[\oint (\mathrm{tr}\,\{[(\vec{\partial}_t - ia\sigma_3)s^\dagger]\,[s(\overleftarrow{\partial}_t + ia\sigma_3)]\} + pa)dt], \tag{C.15}$$

where the action is evaluated over closed loops in s-space. This can be rewritten with the aid of an auxiliary field $A(t)$ [12] as

$$Z \sim \int \mathscr{D}A\mathscr{D}s \exp i[\oint (\mathrm{tr}\,\{[(\vec{\partial}_t - iA\sigma_3)s^\dagger]\,[s(\overleftarrow{\partial}_t + iA\sigma_3)]\} + pA + \tfrac{1}{2}(p/2)^2)dt], \tag{C.16}$$

where a constant

$$\int \mathscr{D}A \exp\{2[A + (\tfrac{1}{4}p - a)]^2\}$$

has been ignored.

In (C.16) A acts as an independent gauge field, which changes by $A \to A - \dot{\alpha}$ under the gauge transformation (C.10), so that (C.16) is gauge invariant. We can work in the specific gauge $A = 0$, and require that the equation of motion obtained from the variation with respect to A (i.e. Gauss's law for this case) be realized as a constraint on the physical states. In this gauge the Lagrangian of (C.16) is just

$$\mathscr{L}(A = 0) = \mathrm{tr}\,(\dot{s}^\dagger \dot{s}) + \tfrac{1}{2}(p/2)^2 \tag{C.17}$$

and Gauss's law is

$$\mathrm{tr}\,(-i\sigma_3 s^\dagger \dot{s} + i\dot{s}^\dagger s\sigma_3) = -p. \tag{C.18}$$

Equation (C.18) is the equivalent of (4.7), since the l.n.s. can be identified with the generator of right transformations (C.10), as we discuss further in Appendix D. Indeed, as we also show there, the term $\mathrm{tr}\,(\dot{s}^\dagger \dot{s})$ is precisely $\frac{1}{2}\boldsymbol{J}^2$, the square of the angular momentum operator (the motion in $\boldsymbol{r}$ being ignored, only the angles varying). The Hamiltonian in this gauge is therefore

$$H(A = 0) = \tfrac{1}{2}\boldsymbol{J}^2 - \tfrac{1}{2}(p/2)^2 \tag{C.19}$$

and the eigenfunctions are again $\mathscr{D}^j_{m'm}(s)$ with m (which carries the right multiplications) restricted to the value $-p/2$, $p = 0, \pm 1, \pm 2, \ldots$. The eigenvalues are $\frac{1}{2}(j(j+1) - (p/2)^2)$, the allowed j being $j = |p/2|, |p/2|+1, \ldots$.

APPENDIX D

Angular momentum

The connection between the s-formalism and the conventional 'spherical top' formalism can be made explicit by parametrizing s by the *Euler angles* α, β, γ according to

$$s = \begin{pmatrix} e^{i(\alpha+\gamma)/2}\cos\beta/2 & e^{i(\gamma-\alpha)/2}\sin\beta/2 \\ -e^{-i(\gamma-\alpha)/2}\sin\beta/2 & e^{-i(\alpha+\gamma)/2}\cos\beta/2 \end{pmatrix}. \tag{D.1}$$

Straightforward calculation then yields

$$T \equiv \operatorname{tr}(\dot{s}^\dagger \dot{s}) = \tfrac{1}{2}\dot{\beta}^2 + \tfrac{1}{2}(\dot{\gamma} + \dot{\alpha}\cos\beta)^2 + \tfrac{1}{2}\dot{\alpha}^2 \sin^2\beta, \tag{D.2}$$

which may be compared with the expression for the spherical top kinetic energy given by Edmonds [47], p. 66. The momenta canonically conjugate to α, β, γ are then

$$p_\alpha = \frac{\partial T}{\partial \dot{\alpha}} = \dot{\alpha} + \dot{\gamma}\cos\beta, \quad \text{etc.}, \tag{D.3}$$

and the passage to quantum theory is made by

$$p_\alpha \to -i\frac{\partial}{\partial\alpha}, \quad \text{etc.} \tag{D.4}$$

We find

$$T = -\tfrac{1}{2}\left\{\frac{\partial^2}{\partial\beta^2} + \cot\beta\frac{\partial}{\partial\beta} + \operatorname{cosec}^2\beta\frac{\partial^2}{\partial\gamma^2} + \frac{1}{\sin^2\beta}\frac{\partial^2}{\partial\alpha^2} - \frac{2\cos\beta}{\sin^2\beta}\frac{\partial^2}{\partial\alpha\partial\gamma}\right\} \tag{D.5}$$

for the operator representing the (rotational) kinetic energy. This is, in fact, precisely the angular kinetic energy

$$T = \tfrac{1}{2}\boldsymbol{J}^2 \tag{D.6}$$

following Edmonds [47]. The eigenfunctions of T are then $\mathscr{D}^J_{m'm}(\alpha, \beta, \gamma)$. Alternative parametrizations of s, such as (3.29), are, of course, also possible.

Finally, we note that

$$i\operatorname{tr}(\dot{s}^\dagger s\sigma_3 - \sigma_3 s^\dagger \dot{s}) = 2(\dot{\alpha} + \dot{\gamma}\cos\beta) = 2p_\alpha. \tag{D.7}$$

In the quantum theory, p_α is the generator of rotations about the 3-axis, which are represented in terms of s by right transformations

$$s(\alpha, \beta, \gamma) = s(\alpha = 0, \beta, \gamma)e^{i\sigma_3\alpha/2}. \tag{D.8}$$

Thus $2p_\alpha$ is the generator associated with the transformation (C.10), as claimed in Appendix C, and its eigenvalues should be integral — as indeed is required by the constraint (C.18).

REFERENCES

[1] T. H. R. Skyrme, *Proc. R. Soc. London* **A260**, 127 (1961).
[2] F. Bopp, R. Haag, *Z. Naturforsch.* **5a**, 644 (1950).
[3] L. S. Schulman, *Phys. Rev.* **176**, 1558 (1968); see also L. S. Schulman, *J. Math. Phys.* **12**, 304 (1971).
[4] P. A. M. Dirac, *Proc. R. Soc. London* **A133**, 60 (1931).
[5] I. Tamm, *Z. Phys.* **71**, 141 (1931).

[6] R. Jackiw, C. Rebbi, *Phys. Rev. Lett.* **36**, 1116 (1976); P. Hasenfratz, G. 't Hooft, *Phys. Rev. Lett.* **36**, 1119 (1976).
[7] A. Goldhaber, *Phys. Rev. Lett.* **36**, 1122 (1976).
[8] E. Witten, *Nucl. Phys.* **B223**, 422 (1983).
[9] E. Witten, *Nucl. Phys.* **B223**, 433 (1983).
[10] J. Wess, B. Zumino, *Phys. Lett.* **37B**, 95 (1971).
[11] E. Guadagnini, *Nucl. Phys.* **B236**, 35 (1984).
[12] E. Rabinovici, A. Schwimmer, S. Yankelowicz, *Nucl. Phys.* **B248**, 523 (1984).
[13] A. P. Balachandran, G. Marmo, B.-S. Skagerstam, A. Stern, *Gauge symmetries and fibre bundles*, Springer-Verlag, Berlin etc. 1983.
[14] A. P. Balachandran, G. Marmo, B.-S. Skagerstam, A. Stern, *Nucl. Phys.* **B162**, 385 (1980).
[15] F. Zaccaria et al., *Phys. Rev.* **D27**, 2327 (1983).
[16] A. P. Balachandran, F. Lizzi, V. G. J. Rodgers, A. Stern, *Nucl. Phys.* **B256**, 525 (1985).
[17] A. P. Balachandran, TASI lectures at Yale University, 9 June-15 July 1985.
[18] M. V. Berry, *Proc. R. Soc. London* **A392**, 45 (1984).
[19] M. Stone, Illinois preprint ILL-(TH)-85-#55 (1985).
[20] Y. S. Wu, A. Zee, *Nucl. Phys.* **B258**, 157 (1985).
[21] Y. S. Wu, A. Zee, Santa Barbara preprint NSF-ITP-85-128 (1985).
[22] I. J. R. Aitchison, C. M. Fraser, *Phys. Rev.* **D31**, 2605 (1985).
[23] P. Simic, *Phys. Rev. Lett.* **55**, 40 (1985); and Rockefeller University preprint RU84/B/106 (1984).
[24] H. Kuratsuji, S. Iida, *Prog. Theor. Phys.* **74**, 439 (1985).
[25] A. Messiah, *Quantum mechanics*, North Holland, Amsterdam 1962, Vol. II, pp. 781–800.
[26] F. Wilczek, A. Zee, *Phys. Rev. Lett.* **52**, 2111 (1984).
[27] L. Alvarez-Gaumé, P. Ginsparg, *Nucl. Phys.* **B243**, 449 (1984).
[28] P. Nelson, L. Alvarez-Gaumé, *Commun. Math. Phys.* **99**, 103 (1985).
[29] T. T. Wu, C. N. Yang, *Phys. Rev.* **D12**, 3845 (1975).
[30] T. T. Wu, C. N. Yang, *Nucl. Phys.* **B107**, 365 (1976).
[31] L. H. Ryder, *J. Phys. A: Math. Gen.* **13**, 437 (1980).
[32] M. Minami, *Prog. Theor. Phys.* **62**, 1128 (1979).
[33] A. A. Belavin, A. M. Polyakov, *JETP Lett.* **22**, 245 (1975).
[34] F. Wilczek, A. Zee, *Phys. Rev. Lett.* **51**, 2250 (1983).
[35] A. M. Din, W. J. Zakrzewski, *Phys. Lett.* **146B**, 341 (1984).
[36] Y. S. Wu, A. Zee, *Phys. Lett.* **147B**, 325 (1984).
[37] T. Jaroszewicz, *Phys. Lett.* **146B**, 337 (1984); Harvard preprint HUTP/BO1 (1985).
[38] M. J. Bowick, D. Karabali, L. C. R. Wijewardhana, Yale University preprint YTP 85-20 (1985).
[39] F. Wilczek, *Phys. Rev. Lett.* **48**, 1144 (1982).
[40] F. Wilczek, *Phys. Rev. Lett.* **49**, 957 (1982).
[41] R. Jackiw, A. N. Redlich, *Phys. Rev. Lett.* **50**, 555 (1983).
[42] G. S. Adkins, C. R. Nappi, E. Witten, *Nucl. Phys.* **B228**, 552 (1983).
[43] D. Finkelstein, *J. Math. Phys.* **7**, 1218 (1966).
[44] D. Finkelstein, J. Rubinstein, *J. Math. Phys.* **9**, 1762 (1968).
[45] J. W. Williams, *J. Math. Phys.* **11**, 2611 (1970).
[46] J. S. Dowker, *J. Phys. A*, Ser. 2, No. **5**, 936 (1972); see also J. S. Dowker, *Lett. Nuovo Cimento* **4**, 301 (1972), and L. S. Schulman, *J. Math. Phys.* **12**, 304 (1971).
[47] A. R. Edmonds, *Angular momentum in quantum mechanics*, Princeton University Press, Princeton N. J. 1957.