International Journal of Theoretical Physics, Vol. 17, No. 4 (1978), pp. 301–308

Monopoles, Topology, Symmetry[1]

Peter G. O. Freund

The Enrico Fermi Institute and the Department of Physics, University of Chicago, Chicago, Illinois 60637

Maxwell's electromagnetic theory is symmetric under the interchange of electric and magnetic field strengths: $\mathbf{E} \to \mathbf{H}$, $\mathbf{H} \to -\mathbf{E}$. To maintain this symmetry in the presence of sources one has to also interchange the electric and magnetic charges and therefore one has to introduce magnetic charges in the first place. In a quantum-theoretical context, as shown by Dirac, this requires a relation between possible values e and g of the electric and magnetic charges, viz.,

$$\exp(i4\pi eg) = 1$$

Dirac's theory of magnetic monopoles introduced the strings which clearly pointed at the topological repercussions of magnetic charge. This is more apparent in the Wu–Yang method of sections. Electromagnetism is a gauge theory of the (Abelian) $U(1)$-group. In the presence of magnetic charges this becomes extended to $U(1)_e \times U(1)_m$, i.e., the direct product of an electric and a magnetic $U(1)$ group. This "doubled" symmetry is not manifest in the original Dirac formulation, but can be made manifest at the expense of not manifestly exhibiting Lorentz invariance and/or locality. This is the contents of work by Zwanziger and Schwinger.

Can one also introduce magnetic charges in the Yang–Mills gauge theory of a non-Abelian group G? More importantly, can one also thereby double the symmetry to $G_e \times G_m$? These are obvious questions. Yet the methods of

[1] The first half of this talk reviews the topology of magnetic monopoles in non-Abelian gauge theory, while the second half reviews some recent developments and contains some new results. For the "older" work I have omitted a detailed bibliography as it can be found in the preprint "Magnetic Monopole Bibliography 1973–1976" by R. A. Carrigan, Jr., Fermilab-77/42. For the more recent papers full references are provided in the text. Work supported in part by the National Science Foundation: Contract No. PHY74-08833.

0020-7748/78/0400-0301$05.00/0

Zwanziger and Schwinger are ill suited for direct non-Abelian generalization, and one has to develop the topological aspects in much more detail. To establish the requisite ideas, I will briefly review the work of 't Hooft and Polyakov—discussed by Dr. Marciano—with a more topological emphasis.

One deals with a gauge theory in which the $SO(3)$ gauge symmetry is broken spontaneously to $O(2)$. In terms of the scalar and vector $SO(3)$ triplets ϕ^a and $A_\mu{}^a$ ($\mu = 0, \ldots, 3$, $a = 1, 2, 3$) the Lagrangian is

$$\mathscr{L}(e, a, \lambda, \phi^a, W_\mu{}^a) = -\tfrac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} + \tfrac{1}{2}D_\mu\phi^a D^\mu\phi^a - \tfrac{1}{4}\lambda(\phi^a\phi^a - a^2)^2$$

with

$$D_\mu\phi^a = (\partial_\mu\delta^{ac} + e\epsilon^{abc}W_\mu{}^b)\phi^c \qquad G^a_{\mu\nu} = \partial_\mu W_\nu{}^a - \partial_\nu W_\mu{}^a + e\epsilon^{abc}W_\mu{}^b W_\nu{}^c$$

Because of the potential (last) term, the ground state corresponds to $\phi^a\phi^a = a^2$. Selecting a direction $\phi_0{}^a$ for ϕ^a causes the spontaneous breakdown of $SO(3)$ to $O(2)$. The two scalar massless Nambu–Goldstone bosons combine with the two massless vector bosons W^+ and W^- into two massive vector bosons, as pointed out by Brout and Englert, by Higgs, and also by Guralnik, Hagen, and Kibble. Thus the particle spectrum of the theory contains two massive vector bosons W^+ and W^- both of mass $m_W = |e|a$, a massless "photon" $\gamma \equiv W^0$ and a "leftover" Higgs boson σ of mass $m_\sigma = (2\lambda)^{1/2}a$. Other possible vacua can be obtained from ϕ_0 by acting on it by any element g of $SO(3)$. Of course rotations around the direction of the (isovector) ϕ_0 leave it invariant. These rotations form an $O(2)$ group so that the manifold of all vacua is the coset space $SO(3)/O(2)$.

The field equations corresponding to the Lagrangian above have spherically symmetric (up to a gauge transformation) solutions

$$W_m{}^a = \epsilon^{amn}\frac{x^n}{er^2}[1 - k(r)] \qquad W_0{}^a = 0$$

$$\phi^a = -\frac{x^a}{r}f(r) \qquad r = |\mathbf{x}|$$

The condition that the energy be finite is $f(r) \xrightarrow[r\to\infty]{} a$. Particularly transparent is the Prasad–Sommerfield limit defined by letting λ go to zero while retaining $\boldsymbol{\phi}^2 = a^2$ at $r \to \infty$. In this limit the differential equations for $f(r)$ and $k(r)$ can be solved analytically with the result

$$k(r) = Cr/\sinh(Cr) \qquad f(r) = [Cr\coth(Cr) - 1]/er$$

where

$$C = m_W = |e|a$$

These classical solutions correspond in the quantum theory to particles M of mass

$$m_M = |g|a \qquad g \equiv \frac{4\pi}{e}$$

These particles carry a new type of conserved charge *topological* in nature as was noted by Arafune, Freund and Goebel, Monastyrskiĭ and Perelomov, Fateev, Tyupkin, and Schwarz, and by Coleman. We can see this as follows. At any point of configuration space at which $\phi^b\phi^b \neq 0$ the normalized scalar fields $\hat{\phi}^a = \phi^a/(\phi^b\phi^b)^{1/2}$ define a point on a two-sphere of unit radius S^2. In particular, for $r \to \infty \phi^b\phi^b \to a^2$ so that $\hat{\phi}$ exists and defines a map from a sphere $S_R{}^2$ of large radius R in configuration space to the unit sphere S^2. Topologically, such a map can be characterized up to a homotopy (i.e., a continuous deformation) by a degree or wrapping number d much like the map of a circle onto a circle by a winding number. For single-valued $\hat{\phi}$ this degree must be an integer. Remarkably this degree can be expressed as the space integral of a local density:

$$d = \frac{e}{4\pi}\int k^\circ d^3x \qquad k^\circ = \frac{1}{2e}\,\epsilon_{abc}\epsilon_{ijk}\partial_i(\hat{\phi}^a\partial_j\hat{\phi}^b\partial_k\hat{\phi}^c)$$

So far this is all topology and no dynamics has been used. But any dynamics will determine a continuous time evolution for the fields ϕ^a and hence for d. Yet at any time d must be integer. There is only one way a quantity can evolve continuously in time and equal an integer at all times, namely, by staying put. Thus the space integral of k° must be *time independent*. The corresponding differential conservation law has the form

$$\partial_\mu k^\mu = 0$$

with

$$k^\mu = \partial_\nu {}^*H^{\mu\nu} = \tfrac{1}{2}\epsilon^{\mu\nu\rho\sigma}\partial_\nu H_{\rho\sigma} \qquad H_{\mu\nu} = \frac{1}{e}\,\epsilon_{abc}\hat{\phi}^c\partial_\mu\hat{\phi}^b\partial_\nu\hat{\phi}^c = -H_{\nu\mu}$$

But the integer d is gauge invariant, while $H_{\mu\nu}$ is not. Hence there must exist a gauge-invariant antisymmetric tensor $F_{\mu\nu}$ such that

$$-\int \partial_\mu {}^*F^{\mu 0}d^3x = \int \partial_\mu {}^*H^{\mu 0}d^3x$$

Indeed the tensor

$$F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu - H_{\mu\nu}$$

with

$$B_\mu = \hat{\phi}^a W_\mu{}^a$$

has this property and is gauge invariant (it can be recast in manifestly gauge-invariant form). So we find that

$$M = \int d^3x \partial_\nu {}^*F^{\nu 0}$$

is a Lorentz and gauge-invariant conserved charge that can take values $4\pi/e \times$ integer. As explained its conservation is of purely topological origin, independent of any dynamics. As such it generates no symmetry and its conservation has nothing to do with Noether's theorem. In different contexts such "topological charges" were considered by Finkelstein and Misner, and by Skyrme.

In short, then, the scalar fields ϕ^a at large distances from the origin are vacuum fields $(\phi^a)^2 = a^2$ and they map a large two-sphere $S_R{}^2$ of configuration space onto the vacuum manifold G/H [in the $SO(3)/O(2)$ example considered above this was, of course, a two-sphere itself]. Magnetic charge is thus a topological characteristic of these maps from $S_R{}^2$ to the manifold G/H. It characterizes these maps up to a homotopy. Mathematically, these homotopy classes are described by the second (since the map originates on a two-sphere) homotopy group $\pi_2(G/H)$ of the vacuum manifold G/H. The conservation of magnetic charge expresses the compatibility between topology and dynamics. Yet this all puts magnetic charge conservation on a new topological foundation different from that of the ordinary Noetherian charges. This is quite unlike the Abelian case, where as we saw magnetic charge can be viewed as Noetherian. Could it be that there is an alternative description of the theory in which magnetic charge is Noetherian rather than topological? This question has been asked and answered at a conjectural level in a very interesting paper by Montonen and Olive (1977), which I shall now explain.

Earlier work of Englert and Windey (1976) and of Goddard, Nuyts, and Olive (1977) has shown that in a gauge theory of the non-Abelian group G spontaneously broken to the smaller group H, the magnetic monopoles can be, so to speak, classified according to a new "dual" group ${}^*H$. This dual group ${}^*H$ has the same dimension as but need not coincide with H. For example, ${}^*SU(3) = SU(3)/Z_3$ [Z_3 is the center of $SU(3)$], ${}^*\mathrm{Spin}\,(2r+1) = \mathrm{Sp}\,(2r)/Z_2$ [Spin (n) is the simply connected covering group of $O(n)$ and $\mathrm{Sp}\,(2r)$ is the symplectic group] and ${}^{**}H = H$. Unlike H, the dual group ${}^*H$ is not a manifest (Noetherian) symmetry of the theory. Rather, the quantum numbers associated with it are topological. Montonen and Olive (1977) then ask whether there could be an alternative picture of the theory in which ${}^*H$ or indeed a group ${}^*G$, dual in some sense to the "large" group G, becomes a manifest Noetherian symmetry and G and H become topological. The roles of Noetherian and of topological quantum numbers would be interchanged in the two pictures. In two dimensions an example along these lines is provided by the Sine–Gordon model and (the zero charge sector of) the massive Thirring model for which the work of Coleman, Luther, Emery, and others shows that

$$*(\text{Sine–Gordon theory}) = \text{massive Thirring model}$$

In four dimensions Montonen and Olive propose that the dual of the $SO(3)$ 't Hooft–Polyakov model reviewed above is a model of the same type with a different coupling parameter. Such a statement can be convincingly established only quantum theoretically, just as in the $2D$ example. But they note a number of features of the 't Hooft–Polyakov model already at the classical level that make the suggestion quite plausible. Consider, along with the Lagrangian $\mathscr{L}(e, a, \lambda, \phi^a, W_\mu{}^a)$ given above, also a dual Lagrangian $*\mathscr{L}$ defined as

$$*\mathscr{L} = \mathscr{L}(*e, *a, *\lambda, *\phi^a, *W_\mu{}^a)$$

Now take the Prasad–Sommerfield limits of both the $\mathscr{L}$ and $*\mathscr{L}$ theories (i.e., let λ and $*\lambda \to 0$ while keeping a and $*a$ fixed). In this limit the gauge boson, Higgs boson, and monopole spectra of the two theories are presented in Table 1.

TABLE 1. Particle spectra of the $\mathscr{L}$ and $*\mathscr{L}$ theories in the Prasad–Sommerfield limit.

Lagrangian	Particle: Role	Spin/$\hbar$	Name	Mass	Electric charge	Magnetic charge
$\mathscr{L}$	Gauge boson	1	W^+	$\|e\|a$	$+e$	0
			W^-	$\|e\|a$	$-e$	0
			γ	0	0	0
	Higgs boson	0	σ	0	0	0
	Monopole	?	M^+	$\frac{4\pi}{\|e\|}a$	0	$+4\pi/e$
			M^-	$\frac{4\pi}{\|e\|}a$	0	$-4\pi/e$
$*\mathscr{L}$	Gauge boson	1	$*W^+$	$\|*e\|*a$	$+*e$	0
			$*W^-$	$\|*e\|*a$	$-*e$	0
			$*\gamma$	0	0	0
	Higgs boson	0	$*\sigma$	0	0	0
	Monopole	?	$*M^+$	$\frac{4\pi}{\|*e\|}*a$	0	$+4\pi/*e$
			$*M^-$	$\frac{4\pi}{\|*e\|}*a$	0	$-4\pi/*e$

The spin of the monopoles is not determined in classical theory. Spherical symmetry of the monopole solutions just guarantees that the corresponding quantum states will have definite spin rather than being accompanied by higher spin orbital excitations. Assume now that the spin of the monopoles in units of $\hbar$ equals 1. Then for the choice (notice e and $^*e$ are both in rationalized units)

$$^*e = \frac{4\pi}{e} \qquad ^*a = a$$

the two particle spectra become identical provided one sets up the correspondence

$$^*M^{\pm} \equiv W^{\pm} \qquad ^*W^{\pm} = M^{\pm} \qquad ^*\gamma = \gamma \qquad ^*\sigma = \sigma$$

$$\begin{pmatrix}\text{electric}\\ \text{magnetic}\end{pmatrix} \text{charge for theory} \rightleftarrows \begin{pmatrix}\text{magnetic}\\ \text{electric}\end{pmatrix} \text{charge for theory.}$$

But $4\pi/e$ is precisely the rationalized value of the magnetic charge of the monopoles in the $\mathscr{L}$ theory. Thus we are presented with the possibility of two dual pictures of the 't Hooft model, the original ($\mathscr{L}$) picture and the dual ($^*\mathscr{L}$) picture, such that: (a) the monopoles of one are the massive gauge bosons of the other and vice-versa; (b) the photon and Higgs boson are the same for both pictures; and (c) electric and magnetic charge change roles between the two pictures. We thus see that the Noetherian and topological conservation laws are interchanged between the two pictures (just like in the Sine–Gordon–Thirring case). The real gauge group is then $O(3)_e \times O(3)_m$. This is not a manifest symmetry as in the Abelian case but rather in one picture $O(3)_e$ is manifest and $O(3)_m$ is topological, while it is the other way round in the dual picture. For this picture to make sense the monopoles must have spin 1, and this can be decided only at the quantum level. But even at the classical level there are further constraints. First of all, the number of spherically symmetric monopoles must equal that of massive gauge bosons of which for the $O(3) \rightarrow O(2)$ spontaneous breaking we know there are two. Now, according to a result of Guth and Weinberg (1976), there are precisely two spherically symmetric monopole solutions in this model and they have equal masses and equal and opposite magnetic charges. So this result agrees with the interpretation of monopoles as gauge bosons.

Finally, Manton (1977) has derived the long-range forces between monopoles in the Prasad–Sommerfield limit. He finds that opposite magnetic charges attract with double the normal Coulomb force, while like magnetic charges exert no long-range forces upon each other. But this is also the pattern of long-range forces between gauge bosons. Indeed, these forces are due to the exchange of the massless γ and σ quanta. The "photon" γ gives rise to a

Coulomb force, while the scalar σ exchange, like gravitation, is always attractive and minimal coupling determines its strength to be such that it doubles the strength of long gauge force between opposite charges, while precisely canceling the Coulomb repulsion between like charges.

There are further implications, e.g., on the dyon spectrum, which I will not discuss here. Rather, I would like to add some new results. An $SU(3)$-gauge field coupled to an octet of scalar fields with a renormalizable self-interaction potential will in general lead to a spontaneous breaking of $SU(3)$ down to $SU(2) \times U(1)$. In principle, breaking to $U(1) \times U(1)$ is possible due to quantum effects or to more than one octet of scalar fields. I assume that somehow $SU(3)$ has been spontaneously broken to $U(1) \times U(1)$ along the "λ_3 direction." Then six of the eight gauge bosons acquire mass in such a way that the heaviest two are degenerate as are the lightest four, and the heavier ones are precisely twice heavier than the lighter ones. There are also two Higgs bosons. In this case a Prasad–Sommerfield limit of the monopole spectrum is again exactly soluble. One finds six spherically symmetric monopole solutions, two of which have twice the mass of the remaining (degenerate) four, and as in the $SO(3)$ case the gauge-boson–monopole mass and quantum number spectrum is unchanged under an electric–magnetic interchange (higher special unitary gauge theories can be studied along the same lines). For the breakdown pattern considered here $SU(3) \to U(1) \times U(1)$ one finds the structure of two "photons" coupled to six electrically and six magnetically charged "gauge-bosons–monopoles." The symmetry is $SU(3)_e \times SU(3)_m$ and, although one has found the counterparts of the two Montonen–Olive pictures in which either of the two $SU(3)$'s is made manifest, one cannot but wonder whether a formulation *manifestly* exhibiting the *full* $SU(3)_e \times SU(3)_m$ symmetry exists, as in the Abelian case.

I would also like to observe that in general the two pictures will not be so similar. In both examples considered above the small group (H) is a product of $U(1)$ groups and as such identical to its dual $*H$. But for breaking of, say, $G = SU(3)$ to $H = SU(2) \times U(1)$, one finds $*H = SU(2)/Z_2 \times U(1) \neq H$, so that the two pictures should not be so similar. Moreover, in the presence of further scalar fields with appropriate quantum numbers one may encounter the "isospin becomes spin" phenomenon ('t Hooft, Hasenfratz, Jackiw, Rebbi, Goldhaber) and find fermionic dyons that have half-odd-integer spin. In such cases, say, an "electric" picture with a purely bosonic Lagrangian might lead to a "magnetic" picture in which some of the fields are fermionic, possibly quarks(?), and even supersymmetry may appear. But this is piling speculations upon speculations. Rather let us reemphasize the main impact of the Montonen–Olive conjecture. It restores the analogy to the Abelian case and essentially claims that spontaneously broken gauge theories have a second copy of the symmetry hidden in the monopole sector. The remarkable feature is that this doubling of the symmetry comes "for free," conditioned, as it were, by the topology of gauge theories.

REFERENCES

Englert, F., and Windley, P. (1976). *Physical Review D*, **14**, 2728.
Goddard, P., Nuyts, J., and Olive, D. (1977). *Nuclear Physics*, **B125**, 1.
Guth, A., and Weinberg, S. (1976). *Physical Review D*, **14**, 1660.
Manton, N. S. (1977). *Nuclear Physics*, **B126**, 525.
Montonen, C., and Olive, D. (1977). *Physics Letters*, **72B**, 117.