PHYSICAL REVIEW D VOLUME 46, NUMBER 4 15 AUGUST 1992

Why opposites attract

Y. Aharonov*
Physics Department, Boston University, Boston, Massachusetts 02215

A. Casher
School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv, Israel

Sidney Coleman
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138

S. Nussinov
School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv, Israel
(Received 11 May 1992)

We give a very simple proof that 't Hooft-Polyakov monopoles and antimonopoles attract at all separations, not just at large distances.

PACS number(s): 11.15.Kc, 14.80.Hv

Oppositely electrically or magnetically charged point particles attract at all distances. However, it is not obvious that this is also true for extended objects, like the magnetic monopoles discovered in spontaneously broken gauge theories by 't Hooft [1] and Polyakov [2]. Certainly the long-range part of the force between these objects is attractive, but one might suspect a short-range repulsion when their cores overlap. In this paper we show that this suspicion is groundless: oppositely charged 't Hooft–Polyakov monopoles attract at all distances, just like point particles. The most surprising part of this result is how easy it is to prove.

We stress that our proof is for 't Hooft–Polyakov monopoles as extended objects in classical field theory; we have nothing to say about quantum effects. Also, our proof is for the original 't Hooft–Polyakov monopoles only, the objects that arise in the theory of a triplet of real scalar fields with spontaneously broken SO(3) gauge symmetry. We shall discuss possible generalizations to other cases after we give the proof.

The energy of a field configuration in this theory is given by

$$W=\int d^3\mathbf{x}\left[\tfrac{1}{4}(G_{ij}^a)^2+\tfrac{1}{2}(D_i\phi^a)^2+\tfrac{1}{4}\lambda(\phi^a\phi^a-v^2)^2+\cdots\right] . \tag{1}$$

Here, as usual,

$$G_{ij}^a=\partial_i A_j^a-\partial_j A_i^a+g\epsilon^{abc}A_i^b A_j^c ,$$
$$D_i\phi^a=\partial_i\phi^a+g\epsilon^{abc}A_i^b\phi^c , \tag{2}$$

g, λ, and v are positive real numbers, all indices run from 1 to 3, and the sum over repeated indices is implied. The ellipsis in Eq. (1) denotes terms involving canonical momenta, all of which trivially vanish for the configurations we shall consider.

For the energy to be finite, $|\phi|$ must go to v at spatial infinity. Thus field configurations of finite energy are characterized by a topological charge, the winding number, the number of times $\hat{\phi}^a=\phi^a/|\phi|$ at spatial infinity winds around the unit sphere. The monopole (antimonopole) is the minimum-energy field configuration with winding number one (minus one). The topological charge can also be constructed from a knowledge of the behavior of ϕ at its zeros (assumed to be a finite set). With every zero of ϕ we can associate a winding number, constructed from the behavior of $\hat{\phi}$ on a small sphere surrounding the zero. It is then elementary that the winding number at infinity is the sum of the winding numbers at the zeros.

This observation is the key to defining what we mean by a monopole-antimonopole field configuration. It is a configuration with two (and only two) zeros of ϕ, one with winding number one and with winding number minus one.[1] The monopole-antimonopole solution is the

*Permanent addresses: School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv, Israel, and Department of Physics and Astronomy, University of South Carolina, Columbia, SC 29208.

[1]If we allowed more than two zeros, the minimum energy would be zero. This is because we could put a new monopole and antimonopole very close to our original antimonopole and monopole, with the new objects arranged so as to almost annihilate the original objects.

minimum-energy configuration with fixed zeros. It is a solution of the equations of motion everywhere except at the zeros; that is to say, in general, external forces must be applied to the centers of the monopole and antimonopole (the zeros) to keep them at fixed locations. Let us denote by $W(\rho)$ the energy of this solution, with ρ the distance between the centers. We shall show that for any positive number δ less than ρ,

$$W(\rho) > W(\rho-\delta) \ . \tag{3}$$

With no loss of generality, we can work in the axial gauge, $A_z^a=0$, and take the zeros to be on the z axis, at $z=\pm\rho/2$. Let us write $W(\rho)$ as the sum of three terms,

$$W(\rho) = W_1 + W_2 + W_3 \ . \tag{4}$$

W_1 is the contribution to Eq. (1) from region 1: $\infty > z > \delta/2$. W_2 is the contribution from region 2: $\delta/2 > z > -\delta/2$. W_3 is the contribution from region 3: $-\delta/2 > z > -\infty$. All three W's are manifestly nonnegative.

In fact, we can make a stronger statement than this: W_2 is strictly positive. The argument is as follows: If we follow 't Hooft [1] and define the magnetic field by

$$F_{ij} = G_{ij}^a \hat{\phi}^a - (1/g)\epsilon^{abc}\phi^a D_i \hat{\phi}^b D_j \hat{\phi}^c \ , \tag{5}$$

then the magnetic flux through any closed surface is $4\pi/g$ times the sum of the topological charges of the zeros enclosed by the surface. Let us consider the surface consisting of the plane $z=0$ and the upper half of the sphere at infinity. For a dipole configuration, the magnetic field falls off at large r like r^{-3}, so the flux through the hemisphere vanishes. There is one zero enclosed by the surface, so the flux through the plane must be $\pm 4\pi/g$. But if $W_2=0$, the field in region 2 is gauge equivalent to the classical vacuum, and F_{ij} vanishes on the plane.

Because W_2 is positive, at least one of the other two W's must be less than $W(\rho)/2$; let us choose the positive z direction so it is W_1. Now let us consider the following trial configuration:

$$\begin{aligned} \phi^a_{\text{trial}}(x,y,z) &= \phi^a(x,y\,|z|+\delta/2) \ , \\ A^a_{i\,\text{trial}}(x,y,z) &= A_i^a(x,y,|z|+\delta/2) \ . \end{aligned} \tag{6}$$

This is simply the original configuration of region 1, patched on to its reflection in the plane $z=\delta/2$.[2] This configuration is continuous, and therefore it is a good trial configuration for the functional (1), which only contains first derivatives. Because the space reflection of a monopole is an antimonopole, the trial configuration consists of a monopole and an antimonopole separated by $\rho-\delta$. Thus, by the variational principle,

$$W(\rho-\delta) \le W_{\text{trial}} = 2W_1(\rho) < W(\rho) \ . \tag{7}$$

This completes the argument.

Our proof would apply to a general theory of scalar and gauge fields, with a general pattern of symmetry breakdown, if we could only define, in general, what we mean by a monopole and an antimonopole separated by a fixed distance.[3] In our proof, this definition was based on identifying the centers of the monopole and antimonopole with the zeros of ϕ. This method (or an equivalent one) works for many simple models, but we have no idea of how to extend it to the general case, or even if it is possible to do so.

Our argument can also be applied to certain theories in two spatial dimensions. For example, it can be used to show that a vortex and an antivortex attract in the Abelian Higgs model. This is an instructive example because it shows there is no hope of proving what might seem at first glance to be the natural partner of our result, the statement that like charges always repel. In the Abelian Higgs model, whether two vortices attract or repel depends on the vector-to-scalar mass ratio, even at large distances.

This work was supported in part by NSF Grants Nos. PHY-87-14654 and PHY-88-7812.

---

[2]Note that this, and the subsequent arguments, would not be true were we not working in the gauge $A_z^a=0$.

[3]Space reflection turns a monopole into an antimonopole no matter what the theory; this is nothing but the statement that parity turns every element of $\pi_2(S^2)$ into its inverse.

[1] G. 't Hooft, Nucl. Phys. **B79**, 276 (1974).
[2] A. M. Polyakov, Pis'ma Zh. Eksp. Teor. Fiz. **20**, 430 (1974), [JETP Lett. **20**, 194 (1974)].