On the relation between charge and topology

1977 J. Phys. A: Math. Gen. 10 717

(http://iopscience.iop.org/0305-4470/10/5/006)

J. Phys. A: Math. Gen., Vol. 10, No. 5, 1977. Printed in Great Britain. © 1977

On the relation between charge and topology

Rafael Sorkin

Department of Applied Mathematics and Astronomy, University College, PO Box 78, Cardiff CF1 1XL, UK

Received 1 October 1976

Abstract. The generalization of Stokes' theorem to non-orientable manifolds shows that a suitable topology can appear to carry *net* electromagnetic charge. By treating this as the origin of electric charge in nature one explains the non-existence of magnetic monopoles.

1. Introduction

Despite its drawbacks, the image of electric charge as nothing more than lines of force trapped in a multiply connected topology retains a certain attraction. According to the version described in Misner and Wheeler (1957) there might be a 'handle'† joining two apparently distant regions of space. In the neighbourhood of one of the ends of such a handle one would detect a net outflux of E field, in other words (as long as one did not look too closely) a 'charged' particle. In effect, an electric monopole would be understood as half of an electric dipole, the other half of which is removed arbitrarily far from it in the apparent metric.

On the other hand, there is no reason within electromagnetic theory why such a handle could not carry magnetic lines of force instead of—or in addition to—electric ones, leading to the existence of particles with net magnetic charge. But such particles have not been seen. Sometimes the existence of the vector potential is invoked to explain this absence, but when A is understood as a gauge field it turns out (Lubkin 1963) that magnetic monopoles are not ruled out after all: they correspond to a certain topological property of the U(1) bundle for which A_μ is a connection.

A second objection to the handle description of electric charge is of less consequence experimentally but perhaps more serious in the long run (after all, magnetic monopoles might be found). Namely the handle, which is hidden to coarse observation, establishes a connection between the two 'charges' corresponding to its ends even though these charges superficially are quite unrelated. It is hard to see how this could be reconciled with the quantum mechanical indistinguishability of all like charged particles except by introducing seemingly unnatural assumptions‡.

This last objection, at any rate, might be overcome if it were possible for the handle (or other suitable topology) to be a monopole rather than a dipole. Then one could bring its ends together to furnish a more local model of charge. The impression that this will not work—that topology cannot display *net* (apparent) electromagnetic charge—

† Called a 'wormhole' by Misner and Wheeler (1957).

‡ For example one might specify a 'backstage' topology which treated all the particles on the same footing; or one might perform some sort of quantum mechanical symmetrization over all possible backstage geometries.

rests on Gauss's law: since the net flux emerging from a region equals the net charge within a region, and since the density of the latter is everywhere zero, the total apparent charge must vanish as well (as long as the region includes the whole handle so it has no hidden boundaries through which flux might escape). But Gauss's law is a version of Stokes' theorem which customarily is formulated only for oriented manifolds. Our first task therefore, is to formulate Stokes' theorem in the non-orientable case.

2. Disorienting Stokes' theorem

Before it fell into the hands of mathematicians Stokes' theorem looked something like

$$\int_D \frac{\partial \mathfrak{A}^\mu}{\partial x^\mu} \mathrm{d}^4 x = \oint_{\partial D} \mathfrak{A}^\mu \, \mathrm{d}\sigma_\mu \tag{1}$$

in which $\mathrm{d}\sigma_\mu$ represents the outwardly directed surface element. After reformulation in the language of exterior calculus it became

$$\int_D \mathrm{d}\omega = \int_{\partial D} \omega \tag{2}$$

where ω is a totally skew covariant tensor or 'form'. Despite the elegance of this version, however, it is version (1) which generalizes most naturally to the non-orientable case.

The reason appears most clearly if one visualizes the theorem in terms of stationary fluid flow, $\mathfrak{A}^\mu$ being the current vector or rather vector *density*, and $\partial_\mu \mathfrak{A}^\mu$ being the rate of fluid creation per unit volume. The interpretation that (1) is equating the net outflux through ∂D to the total creation within D then requires that $\mathrm{d}\sigma_\mu$ be the *outwardly directed* surface element. But a directed surface element corresponds to a covariant vector (figure 1), or more properly, since $\mathfrak{A}^\mu \, \mathrm{d}\sigma_\mu$ must be a scalar, to a covariant vector density of weight -1.

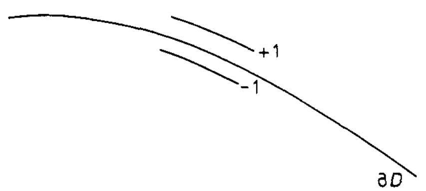

Figure 1. Picture of $\mathrm{d}\sigma_\mu$ showing why it is suited to describe a directed surface element.

The surface element (call it $\mathrm{d}\Sigma$) which is implicit in (2) is, in contrast, an oriented piece of ∂D. If D is itself oriented then a unique correspondence is set up between the internal orientation of $\mathrm{d}\Sigma$ and the inward or outward direction of $\mathrm{d}\sigma_\mu$. In general however there is no such correspondence and no version of Stokes' theorem in terms of forms.

On the basis of the fluid picture it is intuitively clear that (1) will obtain whether or not D is orientable. A bit more formally, imagine that D is decomposed into cubical cells, for each of which (1) is readily explicitly verified. If we sum over all the cells then, since we use the outwardly oriented surface element on each cell, all the boundary integrals cancel save on ∂D itself. The result is precisely (1).

Finally, we can even salvage (2) by a slight re-interpretation. First notice that we can still associate to $d\sigma_\mu$ the expression †

$$d\Sigma^{\alpha\beta\gamma} := d\sigma_\mu\, \epsilon^{\mu\alpha\beta\gamma} \tag{3}$$

where ϵ is the alternating symbol ($\epsilon^{0123} = +1$). This, however, is not a true ('polar') tensor but an *axial* one: it transforms with an extra minus sign under reflection. It is easy to check that, with this definition of $d\Sigma$, (2) will reduce to (1) if ω is the *axial* form

$$\omega_{\alpha\beta\gamma} = \mathfrak{A}^\mu \epsilon_{\mu\alpha\beta\gamma}. \tag{4}$$

We conclude that (2) will always be true as long as one uses *axial*, rather than *polar*, forms and surface tensors.

In fact it is not hard to conclude this directly granted that any non-orientable version of (2) which is *well defined* should be true. To show directly why our re-interpreted (2) is in fact well defined—and also as preparation for our final generalization of Stokes' theorem—it is convenient to present the notion of axial tensor in a more coordinate-free way.

An *orientation* for a vector space is a choice of one of the two disconnected classes of tetrads (or 'n-ads' if $n \neq 4$) in that space. An orientation at a point of the manifold M is a choice of orientation for the tangent space T_xM at x. An *axial* tensor may be thought of as a tensor which only assumes a value once an orientation has been specified, and changes sign when the orientation switches. (Similarly for a tensor density of weight w.) This is, in fact, exactly what one does when, by means of the 'right-hand rule', one treats the magnetic field as a vector.

Now, if we make explicit the volume element in (2), the volume integral appears as

$$\int d\omega \,.\, d\Sigma$$

in which $d\Sigma = d\Sigma^{\mu\nu\alpha\beta}$ should be *axial* since it can become a tensor only relative to an orientation for the volume element it represents. For the integrand to make sense independently of the orientation chosen for $d\Sigma$ it is thus necessary only that ω, and therefore $d\omega$, be axial as well. Similarly, the (outwardly directed) element $d^3\Sigma$ of ∂D becomes an oriented surface tensor once an orientation for D is specified locally, so that again the relative sign between it and the axial form ω is independent of the choice of overall orientation.

To summarize: by working locally and choosing a local orientation we turn (2) into the usual Stokes' theorem. But by proper definitions of ω, $d\Sigma$, we make both integrands independent of this choice, whence there is no need to be able to make it consistently throughout D.

Until now we have been treating D as the whole space under consideration. But when D is, e.g., part of a space-like hypersurface H, the physical fields will be tensors defined in the space–time, M, in which H is embedded, and one needs a Stokes' theorem expressed in terms of these fields.

Unfortunately, a result of the type we seek is not always available in M. To see why this is and what extra condition is needed, consider the Möbius strip, K, with its usual embedding in $\mathbb{R}^3$. The 'volume' integral over K would have to be of the form

$$\tfrac{1}{2}\int d\omega_{jk}\, d^2\Sigma^{jk}$$

† For definiteness we write all formulae for a four-dimensional D.

where ω and $d^2\Sigma$ are to be $\mathbb{R}^3$ tensors. But this is impossible since in circling the strip, $d^2\Sigma$ changes sign whereas $d\omega$ does not. Thus the integrand changes sign, which is absurd. It is impossible to avoid this problem because the sign change in $d^2\Sigma^{jk}$ comes from a *continuous* rotation in $\mathbb{R}^3$ and happens whether $d^2\Sigma$ is axial or polar. Therefore Stokes' theorem for an embedded manifold *cannot be formulated in the general case.*

Clearly the problem is that the Möbius strip is *one-sided.* To avoid this we require that our submanifold K be *externally orientable* in M—i.e. that there be a choice consistent for all x in K of orientation for $T_x(M)/T_x(K)$, the quotient of the tangent spaces.

Assuming this, we can reduce the 'embedded' Stokes' theorem to the case we have just discussed. Although it would be more natural to work with densities as in (1), it is quicker to base our extension on (2) (in its generalized form). In the following the notation assumes a three-dimensional K embedded in four-dimensional M.

If V is a (finite-dimensional) vector space and W a subspace then orientations for W and V/W, which we will call respectively internal and external orientations for W in V, imply one for V in an obvious way. Or if we are given an orientation for V/W then orienting V is equivalent to orienting W. Applying this to $W = T_xK$ and $V = T_xM$ shows that, assuming that $K \subset M$ is externally oriented, an orientation for T_xM is equivalent to one for T_xK. Thus the orientation which is needed in order that the axial form ω 'assume a value' can be taken as that of T_xK rather than that of T_xM, whence ω *restricted* to K becomes an axial form *on* K. Since, further, restriction to K commutes with exterior differentiation, both sides of

$$\tfrac{1}{2}\int_{\partial K} \omega_{\mu\nu}\, d\Sigma^{\mu\nu} = \int_K \tfrac{1}{6}\, d\omega_{\alpha\beta\gamma}\, d\Sigma^{\alpha\beta\gamma} \tag{5}$$

can be evaluated in terms of intrinsic quantities in K, so that (5) reduces to (2) (with 'D'→'K') and is thereby proved. On the other hand, by the same logic the elements $d\Sigma^{\mu\nu}$ and $d\Sigma^{\alpha\beta\gamma}$ can be construed as tensors in M so that (5) becomes the generalization we were seeking.

Theorem. Let K be a p-dimensional compact manifold with boundary which is externally oriented as a submanifold of M. Let $d^p\Sigma$ (or $d^{p-1}\Sigma$) be the volume element of K (or of ∂K outwardly directed) interpreted as an axial tensor in M by virtue of the orientation hypothesis. Then

$$\oint_{\partial K} \omega \,.\, d^{p-1}\Sigma = \int_K d\omega \,.\, d^p\Sigma \tag{6}$$

for an axial $(p-1)$-form ω defined on M.

If we re-interpret this in terms of (polar) densities it reads

$$\oint_{\partial K} \mathfrak{A} \,.\, d\sigma = \int_K \operatorname{div} \mathfrak{A} \,.\, d\sigma. \tag{7}$$

In particular, when $p = \dim M - 1$

$$\oint_{\partial K} \tfrac{1}{2}\mathfrak{A}^{\mu\nu}\, d\sigma_{\mu\nu} = \int_K \mathfrak{A}^{\mu\nu}{}_{,\nu}\, d\sigma_\mu. \tag{8}$$

3. Why are there no monopoles?

Let M be a time-oriented space–time and H a space-like hypersurface. H is thus externally oriented in M. To say that some topological feature T of H displays net charge is to say that net flux emerges through a sphere S 'enclosing' this feature. By discarding that part of H outside of S we get a compact three-manifold, K, whose boundary is precisely S (which is what it means to say S encloses T) and questions about a charge displayed by T reduce to ones about the total flux emerging through S (figure 2).

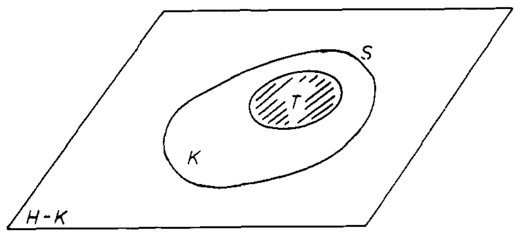

Figure 2.

Consider first the (apparent) electric charge

$$Q=\oint_S \tfrac{1}{2}\mathscr{F}^{\mu\nu}\, \mathrm{d}\sigma_{\mu\nu} \tag{9}$$

where $\mathrm{d}\sigma_{\mu\nu}$ has outward orientation in K and $F^{\mu\nu}=\mathscr{F}^{\mu\nu}/\sqrt{-g}$ is the Maxwell field, which we provisionally take, in accord with tradition, to be a polar ('true') tensor.

According to Stokes' theorem (8), Q is

$$\int_K \mathscr{F}^{\mu\nu}{}_{,\nu}\, \mathrm{d}\sigma_\mu \tag{10}$$

the vanishing of whose integrand is one of Maxwell's equations. It follows that T cannot display net electric charge†.

On the other hand the net (apparent) magnetic charge is given by

$$Q_{\mathrm{M}}=\oint_S \tfrac{1}{2}\,{}^*\mathscr{F}^{\mu\nu}\, \mathrm{d}\sigma_{\mu\nu} \tag{11}$$

where ${}^*\mathscr{F}^{\mu\nu}=\frac{1}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta}/\sqrt{-g}$ is the *dual* of $\mathscr{F}^{\mu\nu}$, and as such an *axial* rather than a *polar* tensor density. Now (11) still makes sense as a *definition* because, if we stay away from T, space appears orientable and we can adopt, say, the right-hand rule for defining B in $H-K$. However we can no longer prove $Q_{\mathrm{M}}=0$ in general because Stokes' theorem does not apply to the axial density ${}^*\mathscr{F}^{\mu\nu}$ when K is non-orientable.

In terms of the three-dimensional quantities $\mathbb{E}$, $\mathbb{B}$ we can express the situation as follows. For the polar vector $\mathbb{E}$ Gauss's law obtains with the consequence that the net electric charge displayed by T must vanish. For the axial vector $\mathbb{B}$ Gauss's law cannot be formulated properly when K is non-orientable, with the consequence, as we shall see,

† An analogous argument based on the Komar stress–energy tensor shows that in a stationary time-oriented vacuum the net mass associated with any space-like hypersurface vanishes.

that *T can* display *magnetic* charge. This is exactly wrong! Recall, however (see also § 5), that the traditional identification of $\mathbb{E}$ as the polar one of the pair $\mathbb{E}$, $\mathbb{B}$ is purely conventional. Henceforth we shall switch the roles of $\mathbb{E}$ and $\mathbb{B}$ by treating $F_{\mu\nu}$ as an axial tensor ($^*F_{\mu\nu}$ becoming the polar tensor) so that, in better accord with experience, the above situation is reversed, and it is *magnetic* monopoles which are excluded.

4. An electrically charged handle

So far we have shown that† no topology can display net magnetic charge. (Remember, $\mathbb{B}$ is now the polar vector.) To make the discussion more concrete we present a particular topology which *can* carry *electric* charge (figure 3). The space–time topology

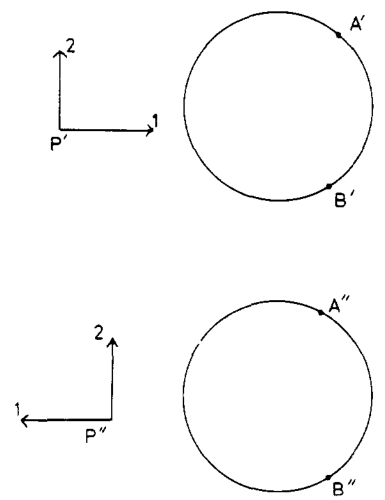

Figure 3. A topology which can display net electric charge.

in question is of the form $\mathbb{R}\times H$ and has for spatial cross section, H, the 'handle' produced by removing from $\mathbb{R}^3$ a pair of balls of equal radius and identifying those points on their surfaces which correspond under translation. To see that this topology is non-orientable consider carrying a triad e_1, e_2, e_3 (of which only two vectors are shown) through the handle from P′ to P″. The 'transverse' vectors e_2, e_3 will retain their direction but e_1, since its head must emerge from the handle before its tail (which entered later), will reverse itself, thereby reversing the orientation of the triad.

Now suppose (figure 4) that the S′ end of the handle carries an apparent electric charge of $\frac{1}{2}Q$—in other words that a flux of $+\frac{1}{2}Q$ (relative to the right-handed orientation at infinity) emerges through S′. By pulling the surface S′ through the handle to S″ we sweep out a three-volume V on which, since it is orientable, we can apply Stokes' theorem‡ to $\mathbb{E}^i \equiv \mathcal{F}^{0i}$. In terms of V therefore the (outward) fluxes of $\mathbb{E}$ through

† Assuming time-orientability.

‡ It is clearer to talk in three-dimensional terms. As shown above there is a perfectly equivalent formulation in terms of the four-tensor $\mathcal{F}^{\mu\nu}$. The sign of $\mathbb{E}$ is, of course, relative to the given time orientation. Notice by the way, that one can check directly that the value of $\mathcal{F}^{0i}$ is unchanged by any coordinate transformation which reduces to the identity on the hypersurface H.

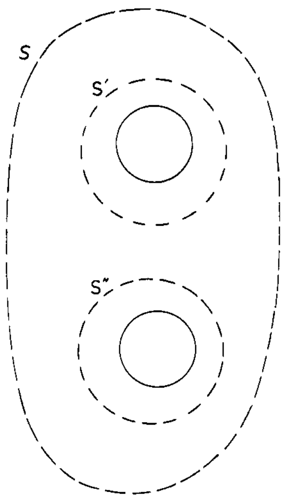

Figure 4. In charged particle models S would enclose a region of microscopic size.

S′ and S″ are equal and opposite. But if V is oriented to agree with the outside orientation at S′ it will disagree with the latter at S″. In terms of the outside orientation, then, both ends of the handle display the *same* charge. The handle as a whole displays charge $\frac{1}{2}Q+\frac{1}{2}Q=Q$.

Figure 5 is intended to show, using geometrical pictures of polar and axial vectors, why it is that the charges, which cancel in one case, reinforce each other in the other.

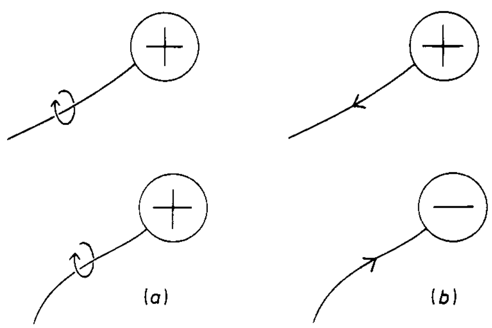

Figure 5. (a) Axial flux **E**. Apparent charge by right-hand rule. (b) Polar flux **B**; net apparent charge vanishes.

5. Speculation

The handle of § 4 is only one of an infinite number of 'quasi-localizable' topologies capable of displaying electric charge. If, in accord with the discussion of § 1, one treats such topologies as the origin of charged particles, one explains immediately the absence of magnetic monopoles (§ 3). One even gets a sort of model for a charge-exchange reaction by considering what happens if one end of a handle passes through a second handle (figure 6).

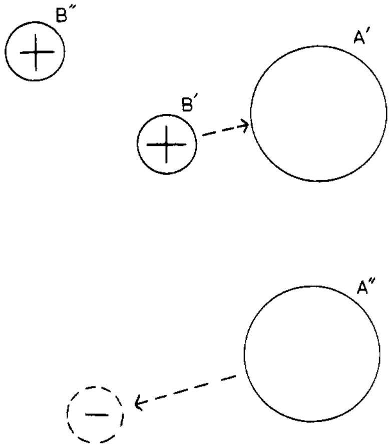

Figure 6. A charge exchange reaction between two handles like that of figure 3.

If the B′ end of handle B passes through the non-orientable handle A, as shown, then, as a bit of thought will verify, the orientation of B′ is reversed relative to B″. In other words B changes from a non-orientable to an orientable handle, or vice versa. But if B becomes orientable it no longer can display a net charge, which must mean an acquisition of B's charge by A. And this is exactly the conclusion one reaches by tracing through the fate of electric lines of force, as represented, e.g., in figure 5(a). In fact the result was inevitable once we treated $\mathbb{E}$ as an axial density and therefore Q_{E} as an *axial scalar.* When such a quantity undergoes spatial reflection (as does the charge of B′ by passing through A) its sign changes.

Still more wildly we could speculate that perhaps one could create magnetic monopoles after all by providing enough energy to pull apart the ends of a magnetic dipole handle (not that we know what holds the ends together in the first place!), all of which ignores possible counteracting quantum mechanical creation of dipole handles, etc. But it seems more prudent to underline some serious objections to this whole, so far essentially classical, picture of charge.

In the first place, it is not clear how half-integral spin (not to speak of quantum mechanics in general) will fit in. Furthermore, known classical solutions of the initial-value equations such as the Reissner–Nordstrøm metric which requires $Q<M$ are incapable of attaining anything like the e/m ratios occurring in real particles, which are of the order of 10^{18} in natural units. It may be that such a ratio is impossible in principle without singular metrics.

Finally there is one objection which *can* be answered concerning whether the assignment of axial character to $\mathbb{E}$ rather than $\mathbb{B}$ is contradicted by experiment. Of course, it may well be that CP non-invariance requires spatial orientability, which would dispose entirely of the charged handle hypothesis. But even if we neglect this problem (perhaps by recourse to the ideas of Lee 1974, on restoring CP symmetry) there is the question of whether, for example, the observed reflection invariance of quantum electrodynamics is compatible with an axial $\mathbb{E}$. Does not the $e\bar{\psi}\gamma^{\mu}\psi A_{\mu}$ coupling become parity violating if A_{μ} becomes axial?

The answer, of course, is that the current, $\bar{\psi}\gamma^{\mu}\psi$, also becomes an axial quantity: in effect the parity operation becomes what is usually called CP; there are still two invariances (neglecting T) but they correspond to CP and C, rather than P and C, of the usual scheme†.

References

Lee T D 1974 *Phys. Rep.* **9** 143–77
Lubkin E 1963 *Ann. Phys., NY* **23** 233
Misner C W and Wheeler J A 1957 *Ann. Phys., NY* **2** 525

† Thus parity in this sense was not overthrown at all—at least until much more recently!