



**HAL**  
open science

## Topology, holes and sources

Alexander Afriat

► **To cite this version:**

| Alexander Afriat. Topology, holes and sources. 2011. hal-00611265v2

**HAL Id: hal-00611265**

**<https://hal.science/hal-00611265v2>**

Preprint submitted on 13 May 2012 (v2), last revised 3 Jan 2013 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Topology, holes and sources

Alexander Afriat

May 13, 2012

## Abstract

The Aharonov-Bohm effect<sup>1</sup> is often called “topological.” But it seems no more topological than electrostatics or Newtonian gravity (or just about any radiation, propagation from a source).

## 1 The Aharonov-Bohm effect

A wavefunction is split into two, and these, having enclosed a (simply-connected) region  $\omega$  containing a solenoid, are made to interfere on a screen. The enclosing wavefunction is sensitive to any enclosed electromagnetism inasmuch as the electromagnetic four-potential  $A$  (a one-form) contributes a phase

$$\exp i \oint_{\partial\omega} A$$

to (the wavefunction along) the boundary  $\partial\omega$  and hence to the interference pattern on the screen. The electromagnetism on  $\omega$  is related to the circulation around the boundary by Stokes’ theorem

$$(1) \quad \oint_{\partial\omega} A = \iint_{\omega} dA.$$

The electromagnetic field<sup>2</sup>  $F = dA$  produced by the solenoid is circumscribed to a middle region  $\lambda \subset \omega$  surrounded by an isolating region  $\lambda' = \omega - \lambda$  where  $F$  vanishes but not  $A$ . The full *Aharonov-Bohm effect* can be considered the ‘differential’ or ‘incremental’ sensitivity of the interference pattern to variations in the current through the solenoid.

---

<sup>1</sup>Ehrenberg & Siday (1949), Aharonov & Bohm (1959)

<sup>2</sup>It is perhaps easiest to think of  $F$  as a magnetic field  $\mathbf{B} = d\mathbf{A}$  produced by a three-current  $\mathbf{J} = d*\mathbf{B}$  in the solenoid, where  $\mathbf{B}$  is a two-form,  $*\mathbf{B}$  the one-form Hodge-dual to it (in three dimensions), and the three-potential  $\mathbf{A}$  is three-quarters of the four-potential  $A \leftrightarrow (\varphi, \mathbf{A})$ .

## 2 The topological interpretation

The *topological interpretation*<sup>3</sup> of the effect can be formulated as follows: If  $A$  were closed throughout a simply-connected region  $\omega$  it would also be exact, and hence expressible as the gradient  $A = d\mu$  of a zero-form  $\mu$  (a real-valued function); the flux

$$\oint_{\partial\omega} d\mu = \iint_{\omega} d^2\mu$$

through the boundary  $\partial\omega$  would then vanish, since  $d^2 = 0$ . But here  $A$  is closed on  $\lambda'$ ; from  $dA = 0|_{\lambda'}$  it does not follow that  $A$  is exact, nor that the flux through the enclosing loop vanishes: it may or may not.

The existence of the source responsible for the effect is therefore ruled out by one topology ( $A$  closed throughout a simply-connected region) but not another.

The same applies to a simply-connected three-dimensional region  $\Omega$  enclosed by a two-dimensional boundary  $\partial\Omega$ . If the two-form  $E$  were closed throughout  $\Omega$  it would also be exact, and hence expressible as the curl  $E = d\zeta$  of a one-form  $\zeta$ ; the flux

$$\iint_{\partial\Omega} d\zeta = \iiint_{\Omega} d^2\zeta$$

through the boundary would then vanish. But if the region on which  $E$  is closed has a hole in it, the flux through the enclosing surface may or may not vanish.

This is precisely what we have in electrostatics, where the electric field  $E = *d\varphi$  is (Hodge-dual to) the gradient  $d\varphi$  of the scalar potential  $\varphi$ . The vanishing divergence  $dE$  expresses the conservation of electricity where none is created, away from the charges

<sup>3</sup>Aharonov & Bohm (1959, p. 490): “in a field-free multiply-connected region of space, the physical properties of the system still depend on the potentials.” Wu & Yang (1975b, p. 3845): “The famous Bohm-Aharonov experiment [...] showed that in a multiply connected region where  $f_{\mu\nu} = 0$  everywhere there are physical experiments for which the outcome depends on the loop integral [...] around an unshrinkable loop.” And p. 3856: “ $f_{\mu\nu}$  underdescribes electromagnetism because of the Bohm-Aharonov experiment which involves a doubly connected space region.” Nash & Sen (1983, p. 301): “We [...] consider the consequence of assuming the field  $F$  to be identically zero in some region  $\Omega$ . At first one may think that there will be no physically measurable electromagnetic effects in such a region  $\Omega$ . This is not so, effects may arise if the topology of  $\Omega$  is non-trivial, e.g. if  $\Omega$  is not simply connected. [...] In terms of parallel transport one says that zero curvature does not imply trivial parallel transport if the region in which the curvature is zero is not simply connected. This underlies the fact that there is a sense in which the connection is a more fundamental object than the curvature, even though a connection is gauge dependent and not directly measurable.” Ryder (1996, p. 101-4): “the Bohm-Aharonov effect owes its existence to the non-trivial topology of the vacuum [...]. The Bohm-Aharonov effect is the simplest illustration of the importance of topology in this branch of physics. [...] The relevant space in this problem is the space of the vacuum, i.e. the space outside the solenoid, and that space is not simply connected. [...] It is thus an essential condition for the Bohm-Aharonov effect to occur that the configuration space of the vacuum is not simply connected. [...] in other words, it is because the gauge group of electromagnetism,  $U_1$ , is not simply connected that the Bohm-Aharonov effect is possible. [...] The configuration space of the Bohm-Aharonov experiment is the plane  $\mathbb{R}^2$  [...] with a hole in, and this is, topologically, the direct product of the line  $\mathbb{R}^1$  and the circle [...]. There is, nevertheless, a positive effect on the interference fringes. The mathematical reason for this is that the configuration space of the null field (vacuum) is the plane with a hole in [...].” Martin (2003, p. 48): “in the case of non-trivial spatial topologies, the gauge-invariant interpretation runs into potential complications. [...] So-called holonomies [...] encode physically significant information about the global features of the gauge field.” Agricola & Friedrich (2010, p. 275): “so ist das verbleibende Gebiet  $\Omega - S$  der Ebene aus Sicht des Elektrons nicht mehr einfach zusammenhangend.” See also Nounou (2003).

that produce  $E$  according to the Maxwell-Poisson equation  $dE = d*d\varphi = \rho$  (the three-form  $\rho$  being the charge density). If the divergence  $dE$  vanished throughout the volume  $\Omega$ , there would be no electricity produced and hence none radiated through the enclosing surface.<sup>4</sup> But a charge in  $\Omega$ —say in a region  $\Lambda \subset \Omega$  isolated by  $\Lambda' = \Omega - \Lambda$ —would prevent electricity from being conserved throughout  $\Omega$ .

We have the same formalism in Newton-Poisson gravity, where  $\varphi$  is the gravitational potential,  $d\varphi$  and  $E$  both represent gravitational force, and  $\rho$  is the mass density. Gravity would therefore be another topological effect.

Again, the topology of the region where the ‘potential’<sup>5</sup> ( $A$  or  $E$  or whatever) is closed tells us relatively little: if the region were simply-connected, conservation would be *general* within the enclosing surface since there could be no holes containing sources; and if nothing were created inside the enclosing surface, the total radiation through it would vanish. But if the topology does *not* allow the presence of holes to be ruled out, the presence of sources in them cannot either; and sources would produce a flux through the enclosing surface.

A non-trivial topology cannot, on its own at any rate, rule out the *absence* of a source either. Nor does it provide the ‘amount’ or ‘intensity’ of the possible source (which would tell us the intensity of the effect—the flux through the enclosing surface). So the full Aharonov-Bohm effect, which can be considered ‘differential,’ is hardly accounted for by topology.

### 3 Gauge freedom

Since the electromagnetic field  $F = dA = dA'$  is easier to measure than  $A$ , it is customary to say that the freedom expressed by the substitution

$$(2) \quad A \mapsto A' = A + d\xi$$

( $\xi$  being a zero-form) is unobservable—and that hence one doesn’t know quite what to make of  $A$ . But even if the freedom (2) to bend and twist the level surfaces of  $A$ ’s local potential<sup>6</sup>  $\gamma$  can make one wonder about the physical meaning of  $A$ , all loops going around the solenoid once puncture the same number (1) of level surfaces.

It is natural to be troubled by the ontological ambiguities of  $A$  and implications of (2). Encouraged by the more reassuring properties of loops, Belot (2003)<sup>7</sup> and Healey (2007) have given them<sup>8</sup> much ontological legitimacy, even centrality.

<sup>4</sup>Over and above any divergence-free electrical background that may or may not be present.

<sup>5</sup>I am taking ‘potential’ to mean little more than ‘primitive’; just as  $A$  is the potential of  $F = dA$ ,  $E$  can be viewed as the potential of  $\rho = dE$ .

<sup>6</sup>For wherever  $A$  is closed it can be written locally as the gradient  $A = d\gamma$  of a scalar potential  $\gamma$ —just as  $E$  can be written locally, wherever it is closed, as the curl  $E = d\zeta$  of a one-form  $\zeta$ .

<sup>7</sup>P. 216: ‘‘holonomies [...] are well-defined quantities on the spaces of states of the standard formulations of Yang-Mills theories. If it is accepted that these theories describe reality, does not it follow that the quantities in question are as real as any others?’’

<sup>8</sup>Or classes of (homotopically) equivalent loops, perhaps with appropriate fields on them, or maybe integrals.

In three dimensions (2) would correspond to a freedom  $E \mapsto E' = E + d\beta$  invisible to the divergence  $dE = dE'$ —but again, any membrane containing the source<sup>9</sup> is punctured by the same number of field-lines of  $E$ , radiated by the source. The trouble is that in the two theories mentioned in §2, electrostatics and Newton-Poisson gravity, the curl  $d\beta$  vanishes since  $E = *d\varphi$  is derived from a potential  $\varphi$ . And no theory with just a source  $\rho$  and its potential  $E$  (but without the scalar potential  $\varphi$ ) comes to mind. If such a theory existed, it would attribute reality to membranes in much the same way as the Aharonov-Bohm effect, according to Belot and Healey, attributes reality to loops.

It could be argued that the real role of topology and loops appears most clearly in the non-Abelian case,<sup>10</sup> with  $\mathbb{S}\mathbb{U}(N)$ ; that here it is subtly concealed by the triviality of the structure group,  $\mathbb{U}(1)$ . The possibility of such non-Abelian ‘emergence’ goes beyond the scope of this paper. Moreover one can—claiming an appropriate ‘Abelian autonomy’—wonder whether the indispensability or centrality of topology would necessarily be inherited by the Abelian case: topology and loops can have one role with non-Abelian structure groups, quite another here.

My thanks to Dennis Dieks, Éricourgoulhon, Marc Lachièze-Rey and Jean-Philippe Nicolas for valuable clarifications and corrections.

## References

- Agricola, I. and T. Friedrich (2010) *Vektoranalysis: Differentialformen in Analysis, Geometrie und Physik*, Vieweg+Teubner, Berlin
- Aharonov, Y. and D. Bohm (1959) “Significance of electromagnetic potentials in the quantum theory” *Physical Review* **115**, 485-91
- Belot, G. (1998) “Understanding electromagnetism” *British Journal for the Philosophy of Science* **49**, 531-55
- Belot, G. (2003) “Symmetry and gauge freedom” *Studies in History and Philosophy of Modern Physics* **34**, 189-225
- Brading, K. and E. Castellani, editors (2003) *Symmetries in physics*, Cambridge University Press
- Ehrenberg, W. and R. E. Siday (1949) “The refractive index in electron optics and the principles of dynamics” *Proceedings of the Physical Society B* **62**, 8-21
- Healey, R. (2007) *Gauging what’s real: the conceptual foundations of contemporary gauge theories*, Oxford University Press, New York
- Martin, C. (2003) “On continuous symmetries and the foundations of modern physics” pp. 29-60 in Brading & Castellani (2003)

<sup>9</sup>In a neutral, abstract framework (where the details of the particular physics are left unspecified) the source is simply the support of the density  $\rho = dE$ .

<sup>10</sup>See Wu & Yang (1975a, p. 3843), Wu & Yang (1975b, pp. 3850-5), Belot (2003, pp. 216-7), Healey (2007, pp. 70-7, 184-99).

- Nash, C. and S. Sen (1983) *Topology and geometry for physicists*, Academic Press, London
- Nounou, A. (2003) "A fourth way to the Aharonov-Bohm effect" pp. 174-99 in Brading & Castellani (2003)
- Ryder, L. (1996) *Quantum field theory*, Cambridge University Press
- Wu, T. T. and C. N. Yang (1975a) "Some remarks about unquantized non-Abelian gauge fields" *Physical Review D* **12**, 3843-4
- Wu, T. T. and C. N. Yang (1975b) "Concept of nonintegrable phase factors and global formulation of gauge fields" *Physical Review D* **12**, 3845-57