Is the world made of loops?

Alexander Afriat

September 27, 2013

Abstract

In discussions of the Aharonov-Bohm effect, Healey and Lyre have attributed reality to loops σ_0 (or hoops $[\sigma_0]$), since the electromagnetic potential A is unmeasurable and can therefore be transformed. I argue that $[A] = [A + d\lambda]_\lambda$ and the hoop $[\sigma_0]$ are related by a meaningful duality, so that however one feels about $[A]$ (or any potential $A \in [A]$), it is no worse than $[\sigma_0]$ (or any loop $\sigma_0 \in [\sigma_0]$): no ontological firmness is gained by retreating to the loops, which are just as flimsy as the potentials. And one wonders how the unmeasurability of one entity can invest another with physical reality; would an eventual observation of A dissolve σ_0, consigning it to a realm of incorporeal mathematical abstractions?

1 Introduction

Thales, one gathers, had nothing but *water*; then came atoms, fire, air, earth, effluvia, fields, energy, waves and other complications. The history of ontological speculation (to say nothing of my garden, §5) has now been enriched by *loops*—and perhaps other boundaries too (§6).

The Aharonov-Bohm effect[1] (§2) involves a relationship between variations in the current through a solenoid and changes in the interference pattern on a screen. The relationship is puzzling, but one can try to make sense of it in various ways. Of the available elements (electromagnetic field F in the solenoid, loops enclosing the solenoid, wavefunction ψ, electromagnetic potential A, its circulation C around the loops, interference pattern P, topological features), some can be chosen, others left out—with inevitable philosophical tradeoffs. Accounts emphasizing the relationship between ψ or P and F (rather than A) have been found disturbingly nonlocal[2] (the tradeoff being *locality* for *invariance*); those preferring[3] A to F seem more local, but are vitiated by

[1]Ehrenberg & Siday (1949), Aharonov & Bohm (1959); see also Franz (1939, 1940, 1965), Olariu & Popescu (1985), Hiley (2013).

[2]Aharonov & Bohm (1959) p. 490: "we might try to formulate a nonlocal theory in which, for example, the electron could interact with a field that was a finite distance away. Then there would be no trouble in interpreting these results, but, as is well known, there are severe difficulties in the way of doing this." See also Healey (1997).

[3]Aharonov & Bohm (1959) pp. 490-1: "we may retain the present local theory and [...] try to give a further new interpretation to the potentials. In other words, we are led to regard $A_\mu(x)$ as a physical variable." See also Feynman *et al*. (1964) §15-5.

the kernel of $d : A \mapsto F = dA$, which produces the troublesome freedom

$$[A + d\lambda]_\lambda = d^{-1}F \tag{1}$$

eliminated by d.

A central notion here will be *measurability*: F and C can be measured, but not A—for the time being at any rate. Indeed measurability is a complicated matter: what is unmeasurable today may not be tomorrow (or vice versa); it depends on the current state of science, technology, ingenuity, economics and so forth; resources, instruments, capabilities and possibilities available in one space-time region may not be in another. I'll avoid the most absolute notion of measurability, as being too abstract to countenance, and will sometimes include a specification in square brackets: unmeasurable[today], measurability[Tuesday], unmeasurability[given the instruments available in space-time region R] or [with respect to instrument ι] and so forth. Different contexts require different notions of measurability; no notion will be given an absolute primacy, which transcends context. So I should really say something like "A is unmeasurable[now]" or "A is unmeasurable[in the current state of science and technology]." How can one know that "A will never be measured" or "A is unmeasurable *in principle*"? Maudlin (1998, p. 367) writes that "since potentials which differ by a gauge transformation generate identical effects, no amount amount of observation could reveal the ONE TRUE GAUGE," which I take to mean "since potentials which differ by a gauge transformation generate identical effects *for the time being*, no amount amount of observation[now] could reveal the ONE TRUE GAUGE."[4]

This is related to the matter of *invariance under certain transformations*. The point is not that A can be transformed whereas F and C cannot (for they can); but that A, which is unmeasurable[now], can be subjected to a transformation,

$$A \mapsto A' = A + d\lambda, \tag{2}$$

to which C and F are indifferent (despite being functions of A). Whereas it makes sense to say that "F (or C) is gauge invariant," the meaning of "A is not gauge invariant" or "A is gauge dependent" is less clear. Is it meant that A can be transformed? Of course it can—but so can C and F, and in many different ways: $\mathsf{C} \mapsto \mathsf{C} + 7$, $F \mapsto 2F$ *etc*. "A is gauge dependent" may mean something like "C and F, which are functions of A, are measurable, unlike A itself; and A can be subjected to transformations that leave C and F unchanged."[5] Shorthand I suppose, but not of the clearest sort.

[4] *Cf.* Healey (2007) pp. 113-4: "one cannot rule out a future extension of a Yang-Mills gauge theory that permits observations whose results depend on the existence of a privileged gauge [...]. If that were to happen, then his observations would discriminate in favor of an interpretation of the gauge theory that commits it to such a privileged gauge, and against a holonomy interpretation. This has not yet happened. But since we cannot be sure that it never will, it seems that we are in no position to answer the question as to whether a holonomy interpretation is correct." Belot (1998, footnote 17) seems to countenance the possibility of physically different but empirically indistinguishable potentials A, $A' \in [A]$—which makes the empirical indistinguishability appear particularly contingent, perhaps even temporary. See also Aharonov & Bohm (1959) p. 491: "we must be able to define the physical difference between two quantum states which differ only by a gauge transformation" and Healey (2009).

[5] Lyre has it the other way around—the symmetry *comes first*, A is unmeasurable as a result: (2001) p. S377 "**The Reality of Gauge Potentials.** "Only gauge-independent quantities are observable.""; p. S379

The *holonomy interpretation*[6] of the Aharonov-Bohm effect, which will concern us here (§3), can be summarized as follows: (1) rules out the reality of A; but the circulation C is indifferent to (1) (or rather (2)); since A isn't really there, and something ought to be, the obvious alternative as bearer of physical reality is the loop(s) σ_0 around which A is integrated to give C. Even more briefly: Unlike A, C is invariant, so σ_0 (or perhaps $[\sigma_0]$) is real but not A. Healey (2007) p. 51:

> If the value of the vector potential A_μ at each space-time point x in a region does not represent some qualitative intrinsic physical properties in the vicinity of x, it may be that some function of its integral around each closed curve C in that region does represent such properties of or at (the image of) C. [...] Since the gauge dependence of the vector potential made it hard to accept Feynman's view that it is a real field that acts locally in the Aharonov-Bohm effect, there is reason to hope that a gauge-invariant function of its line integral around closed curves might facilitate a local account of the action of electromagnetism on quantum particles in the Aharonov-Bohm effect and elsewhere.

And p. 105:

> The non-localized gauge potential properties view is motivated by the idea that the structure of gauge potential properties is given by the gauge-invariant content of a gauge theory. The most direct way to implement this idea would be to require that the gauge potential properties are just those that are represented by gauge-invariant magnitudes. [...] While the vector potential A_μ is gauge dependent, its line integral $S(C) = \oint_C A_\mu dx^\mu$ around a closed curve C is gauge invariant.

The point cannot just be that *invariance is ontologically relevant*, for that goes back[7] to Cassirer (1921) or perhaps Einstein (1916) or even Klein (1872); nor can it be anything like *the world is made of structures* or *structures should be taken seriously*,[8] which may be right but not new, either; nor can it be a mere extension of old ideas (about invariance or structures) to another theory. The claim, as I take it, is interesting, original, even spectacular: *the world is made of loops* (or perhaps *loops are ontologically preferable to potentials*). That's what I contest. Not that that transformation properties

"holonomies are gauge-independent quantities and therefore appropriate candidates of observable entities"; (2002) p. 82 "Der Eichsymmetrie zufolge lassen sich Eichpotentiale nicht direkt beobachten – nur eichinvariante Größen können observabel sein. [...] In den Eichtheorien sind diejenigen Entitäten, denen aufgrund observabler Konsequenzen Realstatus zugebilligt werden muß, einerseits klarerweise nur bis auf Eichtransformationen festgelegt [...]."

[6] Healey (1997, 2001, 2004, 2007), Belot (1998), Lyre (2001, 2002, 2004a,b); see also Wu & Yang (1975), Myrvold (2011). The first two letters of "anholonomy" don't look semantically irrelevant, quite on the contrary; the "a" looks very much like a transliterated alpha privative, which far from doing nothing at all would turn "holonomy" into its opposite, "$\neg$ holonomy." One might imagine that the removal of the initial "an" would restore the meaning of "holonomy." Not at all; by a prodigy of language and logic we have holonomy $\equiv \neg$ holonomy. "Holonomy" is often preferred to its anto-synonym (the French say *énantiosème*) "anholonomy." Whether "holonomy" means "anholonomy" or the opposite is settled by context.

[7] See Afriat & Caccese (2010) p. 18.

[8] On structural realism see for instance Poincaré (1902, 1905), Cassirer (1921), Russell (1927), Worrall (1989), French & Ladyman (2003), French (2006).

or structures can be ontologically relevant, nor that anholonomy is a useful mathematical resource.

I argue (§4) as follows: A and σ_0 are related by a significant duality, so that however one feels about the class $[A]$ (or its elements), it is no worse than the hoop[9] $[\sigma_0]$ (or its elements). I'll think of A as a set $\{\sigma_1, \ldots, \sigma_N\}$ of level curves, which can indeed be deformed (by (2))—but so can σ_0. The deformability of $\sigma_1, \ldots, \sigma_N$, or rather $\sigma_0, \ldots, \sigma_N$, is neither here nor there, and shouldn't be used to rule out the reality of A in particular.

I also explore (§§5,6) the relationship between the measurability of a quantity (say A) and the ontic status of a boundary; can the measurement of one entity undermine the physical reality of another, ferrying it off to a shadier realm of mathematical abstractions?

My arguments are no doubt best applied to the holonomy interpretation within an appropriate structural realism—which may or may not be the best description of Healey's position.

Healey (2007) devotes much attention to quantized non-Abelian Yang-Mills theory. I only consider electrodynamics—and hence $\mathbb{U}(1)$, rather than the non-Abelian structure groups of Yang-Mills theory—without quantization (beyond the use of a wavefunction).

2 The Aharonov-Bohm effect

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

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

to (the wavefunction along) the boundary $\partial\omega \equiv \sigma_0$ and hence to the interference pattern on the screen. The electromagnetism on ω is related to the circulation around the boundary by Stokes's theorem

$$\mathsf{C} = \oint_{\partial\omega} A = \iint_{\omega} dA.$$

The electromagnetic field[10] $F = dA$ produced by the solenoid is circumscribed to a middle region $\mu \subset \omega$ surrounded by an isolating region[11] $\mu' = \omega - \mu$ where F vanishes but not A.

Varying the current through the solenoid changes the arbitrarily distant interference pattern, which is surprising.

[9] This is the class of loops going around the solenoid once.

[10] It is perhaps easiest to think of F as a purely *magnetic* field $\mathbf{B}$ produced by the current density $\mathbf{J} = d{*}\mathbf{B}$ in the solenoid.

[11] It will be convenient to view μ and ω as concentric disks.

3 The reality of loops

The electromagnetic field F and inverse image $[A] = d^{-1}F$ are measurable, but not the individual potential A—which may therefore be physically meaningless, on its own at any rate. But surely the Aharonov-Bohm effect has to be conveyed by *something*. If the potential isn't really there, what's left? The solenoid, and the electromagnetic field it contains, are (arbitrarily) far from the wavefunction and the screen on which the effect is seen. The circulation C, which determines the interference pattern, has a promising indifference[12] to (2); but C is just a number, not enough on its own to convey or account for the effect—something more is presumably sought. The number is obtained by integrating any[13] $A \in [A]$ around any $\sigma_0 \in [\sigma_0]$; having ruled out A,[14] Healey (2001),[15] Lyre (2001),[16] Lyre (2004a),[17] Lyre (2004b)[18] and Healey (2007)[19]

[12]Healey (1997) p. 34: "since $S(C)$ is gauge-invariant, it may readily be considered a physically real quantity."

[13]To stress the ontic and conceptual autonomy of loops, Healey (2007, pp. 72-3—and perhaps Lyre too) would rather think of a holonomy as a number assigned to a loop than as the integral of a potential around a loop: *loops can do without potentials*. But can they? How does one know what number the loop gets? But suppose one works out the circulation by integration, then forgets about the original potential A—all that's remembered is the number C and its loop(s). Even if such associations underdetermine the individual potential A, they determine the equivalence class $[A]$ (see Anandan (1983), Barrett (1991), Healey (2004) pp. 626-7). So potentials are never too far away, and cannot be completely divorced from holonomies. See Healey (2001) p. 447.

[14]Lyre (2001) p. S377, Lyre (2004b) p. 665: "realists can hardly be satisfied by the gauge dependence of entities as imminent in the A-interpretation [. . .]." Healey (1997) p. 22: "there is reason to doubt that the magnetic vector potential is a physically real field, since $\mathbf{A}$ is not gauge-invariant, unlike the magnetic field $\mathbf{B}$ [. . .]." Healey (1999) p. 445, Healey (2001) pp. 435-6, 454, Healey (2007) pp. 25-6, pp. 55-6: "If there are new localized gauge properties, then neither theory nor experiment gives us a good grasp on them. Theoretically, the best we can do is to represent them either by a mathematical object chosen more or less arbitrarily from a diverse and infinite class of formally similar objects related to one another by gauge transformations, or else by this entire gauge-equivalence class" and p. 118. Healey's claim that *theory itself* (rather than experimental limitations) rules out a choice of gauge is answered (by himself) in footnote 4 above. *Cf.* Maudlin (1998) pp. 366-7, Leeds (1999) p. 610, Healey (2009) p. 707: "It is especially hard when it is the theory itself that provides our only initial access to those features of situations it represents by newly introduced structures—hard, but not impossible."

[15]P. 448: "it is the holonomies that represent the real physical structures in a gauge situation, rather than any particular bundle connection that is compatible with them." P. 449: "What is distinctive is not the properties represented by holonomies but the nature of the object whose properties they are. On the present view, holonomies represent global properties of a loop that are not determined by any intrinsic properties of the points on that loop."

[16]P. S379: "We [. . .] should consider holonomies as physically real." P. S380: "We may very well represent the physically significant structures in an ontological universe consisting of matter-fields, gauge field strengths, and holonomies."

[17]P. 116: "Die Entitäten der Eichtheorien sind Holonomien [. . .]."

[18]P. 665-6: "we may still consider holonomies as object-like entities, but to such an extent that our notion of an object becomes highly abstract." P. 667: "When we stick to more abstract and non-local entities we do not dismiss the notion of objects altogether. [. . .] A third "intermediate" possibility would be to stick with more abstract but still measurable or, as some authors like to phrase it, "structural" objects—and here non-separable holonomies turn ou a suitable case at hand. [. . .] a more sustained conclusion should perhaps rather be seen in the development of modern physics into more abstract—here spatiotemporally holistic—entities in accordance with an intermediate version of structural realism—less radical than the ontic version but more directly supported from science than the epistemic one." Lyre (p. 658) sees holonomies as "basic entities" and "genuine entities."

[19]P. xviii: "In the simplest case (classical electromagnetism interacting with quantum particles) such an account ascribes properties to (or on) a loop of empty space that are not fixed by properties of anything

attribute reality to[20] σ_0 (or $[\sigma_0]$) instead.[21]

α : *Since* C, *unlike* A, *is indifferent to (2),* A *is not real but* σ_0 *(or* $[\sigma_0]$*) is.*

4 Duality between loops and potentials

But there is a significant duality between loops and potentials: just as a vector $\dot{\sigma}_0'(x) \in T_x M$ and a covector $A(x)$ from the dual space $T_x^* M$ give a number $\langle A(x), \dot{\sigma}_0'(x)\rangle$, the loop[22] $\sigma_0 \equiv \partial\omega$ and potential give a number $(A, \sigma_0) = \mathsf{C}$. Both A and σ_0 can be deformed without affecting the circulation: the potential according to (2); a loop can be deformed into any other loop going around the solenoid once. Both could be replaced by their equivalence classes $[A]$ and $[\sigma_0]$, one could even write $([A], [\sigma_0]) = \mathsf{C}$.

It will be useful to understand the transformation (2) more geometrically, as a deformation of the level sets of A's local primitive[23] σ.[24] One can first imagine a purely 'angular' or 'radial' σ (with values running from zero to $2\pi k = \mathsf{C}$),[25] whose level lines are straight rays radiating through the annulus μ' from the inner disk μ to the edge $\partial\omega$.

located at points around the loop [. . .]." P. 30: "Suppose instead that one takes the holonomies themselves directly to represent electromagnetism and its effects on quantum particles." P. 31: "But if the holonomies directly represent electromagnetism and its effects, then there is still a sense in which the action of electromagnetism on the electrons is not completely local, since holonomies attach to extended curves rather than points." P. 56: "only gauge-invariant functions of these mathematically localized fields directly represent new electromagnetic properties; and these are predicated of, or at, arbitrarily small neighborhoods of *loops* in space-time—i.e. oriented images of closed curves on the space-time manifold." P. 74: "This makes it plausible to maintain that what an SU(2) Yang-Mills theory ultimately describes is not a localized field represented by a gauge potential, but a set of intrinsic properties of what I have simply called loops [. . .]." P. 106: "we arrive at the view that non-localized EM potential properties in a region are represented by the holonomies [. . .] of all closed curves in the region [. . .]. This is the interpretation of classical electromagnetism I shall defend." P. 118: "One can reformulate the theory as a theory of holonomy properties, so that it does not even appear to mention localized gauge potential properties." P. 185: "gauge potentials directly represent no localized gauge properties, but rather indirectly represent non-localized holonomy properties." P. 220: "[. . .] the Aharonov-Bohm effect and other related effects provide vivid examples of physical processes that seem best accounted for in terms of non-localized holonomy properties [. . .]." P. 221: "Should we believe that non-separable processes involving non-localized holonomy properties are responsible for phenomena like the Aharonov-Bohm effect? This belief may be encouraged by the predictive successes consequent upon introducing classical electromagnetism into the quantum mechanics of particles." P. 225: "This reinforces the conclusion that the evidence for contemporary gauge theories lends credence to the belief that these describe non-separable processes, while nothing in the world corresponds to or is represented by a locally defined gauge potential." And the last paragraph of the book pp. 227-8.

[20]A holonomy is a pair $\langle\sigma_0, \mathsf{C}\rangle$. It is hard to see how a loop and a number can be real without the loop itself being real.

[21]Belot (2003) p. 216 has a similar position—without, however, going so far as to claim that loops are *more* real: "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?" See also Belot (1998) p. 544: "we must also consider closed curves in space to be carriers of the electromagnetic predicates" and the final paragraph pp. 553-4.

[22]The one-dimensional manifold $\sigma_0 \subset M$ is the image of the mapping $\sigma_0' : I \to M\,;\, t \mapsto \sigma_0'(t)$, without its parameter t, which is not part of the boundary $\partial\omega$ (where $I \subset \mathbb{R}$ is an interval and the manifold M an appropriate base space).

[23]For wherever A is closed it can be written locally as the gradient $A = d\sigma$ of a zero-form σ.

[24]A similar construction is used in Afriat (2013).

[25]Such a σ cannot be continuous everywhere; we can imagine a single discontinuity, say on the ray with values $\sigma = 2\pi nk$, where the integer n is zero then one, $k = \mathsf{C}/2\pi$ being a constant.

A gauge transformation (2) would then deform the level rays, bending them without making them cross. It is easier to picture the denumerable set $\{\sigma_1, \ldots, \sigma_N\}$ of level curves at intervals of C/N than all of them; they will each be cut once[26] by any loop σ_0 going around the solenoid once.

In this construction we have $N+1$ deformable curves $\sigma_0, \ldots, \sigma_N$, which all seem pretty much on the same footing; α amounts to the surprising claim that

α' : *Only* σ_0 *is real because the other curves* $\sigma_1, \ldots, \sigma_N$ *can be deformed.*

Why should one curve σ_k be any better than the others? How about σ_7? It remains true that $\sigma_0, \ldots, \sigma_6, \sigma_8, \ldots, \sigma_N$ can be deformed.

To emphasise that loops are no better than A, we can even arrange for a gauge transformation to *induce* a loop deformation (thus strengthening the duality): Level rays of unit length determine a unit circle, which will then be 'deflated' into a smaller loop by a gauge transformation (2); to every such transformation there corresponds a different loop σ_0^λ. If a potential subject to (2) is too flimsy to exist, how can loops also subject to (2) be any better? Are vectors any more real than the covectors dual to them? Is a Lagrangian any less real than the Hamiltonian dual to it? Does momentum exist more than velocity?

Another approach is to represent the potential at a higher level of abstraction, at which the ontologically troublesome gauge freedom (2) disappears, leaving only a geometrical structure (horizontal distribution on the principal bundle[27]) that corresponds to $[A]$ but not to A, which it underdetermines.[28] If one is burdened with A and (2) and their ontological implications one has simply chosen the wrong way of thinking about potentials, the wrong level of abstraction, encumbered as it is by confusing and irrelevant clutter that just gets in the way.[29] The move has its costs (in mathematical intelligibility), but it does achieve a satisfactory ontological rehabilitation, which makes potentials sturdier than loops. The higher level of abstraction only makes sense, however, as long as the individual potential A remains unmeasurable. More on the surprising implications of measurability presently.

5 Vertical drop

Suppose I can only measure the curl $F = dA$ of my garden's gradient[30] (or infinitesimal vertical drop) A but not the gradient itself—some instruments and experimental

[26] One should really say *an odd number of times*, as Jean-Philippe Nicolas has pointed out to me. Crossings in opposite directions cancel, and add nothing to the integral.

[27] See Bleecker (1981), Baum (2009) pp. 73ff.

[28] *Cf.* Lyre (2002) p. 82: "Eine genauere Analyse zeigt jedoch, dass nicht den Eichpotentialen, sondern nur Äquivalenzklassen von Potentialen [. . .] bzw. den so genannten Holonomien ontologische Significanz zukommt [. . .]."

[29] Healey (1999 p. 446, 2001 p. 436ff., 2004 p. 628) argues that the gauge freedom, which remains in the form of vertical automorphisms, is in fact ineliminable. But even if such automorphisms, which do correspond to the gauge freedom, are indeed available, the horizontal distribution can be specified without any reference to them.

[30] Needless to say, most gardens have an exact gradient.

possibilities are available, others are not.[31] The indifference of $F = dA + d^2\lambda$ to the exact one-form $d\lambda$ can accordingly be called a *gauge freedom*. The vertical drop or rather circulation

$$\mathsf{C} = \oint_{\partial\omega} A = \iint_{\omega} dA$$

is also indifferent to (2), where $\partial\omega$ is the boundary of any region ω.

The unmeasurability[in an appropriate part of space-time] of A makes my garden an awkward tangle of real physical loops by producing the gauge freedom which in turn invests the boundary $\partial\omega$ (or $[\partial\omega]$) with physical reality. But to dispatch the loops to the shady regions (populated by innocuous mathematical ghosts) where they obstruct neither gardening nor evening strolls it is enough to *work out how to measure A*; for that will change the status of (2) and hence of the loops.[32]

We know that peeking can *kill a cat* (Schrödinger, 1935); but here, peeking at a gradient may well *carry off a loop*!

6 Electrostatics, gravity

The solenoid, or perhaps the field $F = dA$ it produces, is a source whose 'radiation' A is eventually caught by the boundary $\partial\omega$.

In electrostatics the source is a charge density three-form $\rho = d\mathbf{E}$, which radiates the electric two-form $\mathbf{E}$ eventually caught by the two-dimensional boundary $\partial\Omega$ of a region Ω containing ρ. Stokes's theorem again holds, and allows us to write

$$\mathsf{F} = \iint_{\partial\Omega} \mathbf{E} = \iiint_{\Omega} d\mathbf{E} = \iiint_{\Omega} \rho.$$

The difference here is that ρ's primitive $\mathbf{E}$ is measurable, and fixed by the condition $\mathbf{E} = *d\varphi$ (the electrostatic potential φ being a zero-form). But F and ρ can nonetheless be called 'gauge invariant,' in the sense that they are indifferent to the gauge transformation

$$\mathbf{E} \mapsto \mathbf{E}' = \mathbf{E} + d\beta, \tag{3}$$

where $d\beta$ is the curl of a one-form β.

But this gauge invariance may seem purely formal, vacuous, meaningless. To take it more seriously one would have to *forget how to measure* $\mathbf{E}$.[33] Once $\mathbf{E}$ is unmeasurable, (3) will acquire a different status, and so will the boundary $\partial\Omega$. If we then feel trapped inside an infinite gauge-invariant class $[\partial\Omega]$ of real physical membranes, all we have to do, to dissolve them all, is *remember how to measure* $\mathbf{E}$.

[31] Again, there can be other instruments and experimental possibilities in other space-time regions; a notion of experimental possibility may or may not propagate from one region to another.

[32] It may look as though I am *affirming the consequent*; [unmeasurable potential]⇒[real boundary] does not on its own imply [measurable potential]⇒[unreal boundary]. But it seems natural to assume that mathematical boundaries are merely mathematical, unless we have reason to believe they're not. Such a reason—"A is gauge dependent"—is exactly what Healey and Lyre provide. If that reason ceases to hold, however, the departure from the natural assumption is no longer justified.

[33] It is by no means impossible to make a quantity unmeasurable; one can destroy instruments, abolish know-how, banish specialists and so on. Measurability is as reversible as progress.

The same applies, *mutatis mutandis*, to Newton-Poisson gravity, where ρ is the mass density, φ the gravitational potential *etc*.[34]

The same also applies to the Aharonov-Bohm effect itself: Suppose an ingenious experimenter works out how to measure A. That would change the status of (2)—it would be taken less seriously—and hence of $\partial\omega$, which would undergo an ontic transition. We have something like a 'law of ontological conservation': if the reality is not *here* (A), and it has to be somewhere, it must be *there* ($\partial\omega$); but if it *is* here, it no longer has to be there ...

7 Final remarks

Picking a potential A out of $[A]$ is admittedly problematic[35] (as long as A remains unmeasurable). But picking a loop σ_0 out of the hoop $[\sigma_0]$ is just as bad. There are indeed reasons to prefer $[A]$ to any particular potential A, and to prefer $[\sigma_0]$ to any particular loop σ_0; but those reasons are not good reasons to prefer $[\sigma_0]$ to $[A]$, or any particular σ_0 to any particular A; nor are they good reasons to prefer $[\sigma_0]$ to any particular A (or $[A]$, for that matter, to any particular σ_0).

And it is remarkable that the unmeasurability of a differential form ζ defined on Ω should confer physical reality on the boundary $\partial\Omega$. If the 'potential' represented by ζ sooner or later becomes measurable, does the ontic status of the boundary $\partial\Omega$ change?

I thank Jean-Philippe Nicolas for valuable clarifications, Silvia De Bianchi for her suggestions, and Dimitri Kasprzyk for telling me about "énantiosèmes."

References

Afriat, A. (2013) "Topology, holes and sources" *International Journal of Theoretical Physics* **52**, 1007-12

Afriat, A. and E. Caccese (2010) "The relativity of inertia and reality of nothing" *Studies in History and Philosophy of Modern Physics* **41**, 9-26

Aharonov, Y. and D. Bohm (1959) "Significance of electromagnetic potentials in the quantum theory" *Physical Review* **115**, 485-91

Anandan, J. (1983) "Holonomy groups in gravity and gauge fields" in G. Denardo and H. D. Doebner (editors) *Proceedings of the Conference on Differential Geometric Methods in Physics, Trieste 1981*, World Scientific, Singapore

Barrett, J. W. (1991) "Holonomy and path structures in general relativity and Yang-Mills theory" *International Journal of Theoretical Physics* **30**, 1171-1215

Baum, H. (2009) *Eichfeldtheorie: eine Einführung in die Differentialgeometrie auf Faserbündeln*, Springer, Berlin

[34] See Afriat (2013).

[35] Healey (1999) p. 444: "The main problem with ONE TRUE GAUGE is epistemological: the theory itself entails that we could never have any evidence that the TRUE GAUGE was ONE rather than ANOTHER."

Belot, G. (1998) "Understanding electromagnetism" *The 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

Bleecker, D. (1981) *Gauge theory and variational principles*, Addison-Wesley, Reading

Cassirer, E. (1921) *Zur Einstein'schen Relativitätstheorie*, Bruno Cassirer, Berlin

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

Einstein, A. (1916) "Grundlage der allgemeinen Relativitätstheorie" *Annalen der Physik* **49**, 769-822

Feynman, R. P., R. B. Leighton and M. Sands (1964) *The Feynman lectures on physics, volume 2*, Addison-Wesley, Reading

Franz, W. (1939) "Elektroneninterferenzen im Magnetfeld" *Verhandlungen der Deutschen Physikalischen Gesellschaft* **20**, 65-6

Franz, W. (1940) "Elektroneninterferenzen im Magnetfeld" *Physikalische Berichte* **21**, 686

Franz, W. (1965) "Elektroneninterferenzen im Magnetfeld" *Zeitschrift für Physik* **184**, 85-91

French, S. (2006) "Structure as a weapon of the realist" *Proceedings of the Aristotelian society* **106**, 1-19

French, S. and J. Ladyman (2003) "Remodelling structural realism: quantum physics and the metaphysics of structure" *Synthese* **136**, 31-56

Healey, R. (1997) "Nonlocality and the Aharonov-Bohm effect" *Philosophy of Science* **64**, 18-41

Healey, R. (1999) "Quantum analogies: a reply to Maudlin" *Philosophy of Science* **66**, 440-7

Healey, R. (2001) "On the reality of gauge potentials" *Philosophy of Science* **68**, 432-55

Healey, R. (2004) "Gauge theories and holisms" *Studies in History and Philosophy of Modern Physics* **35**, 619-42

Healey, R. (2007) *Gauging what's real: the conceptual foundations of contemporary gauge theories*, Oxford University Press, New York

Healey, R. (2009) "Perfect symmetries" *The British Journal for the Philosophy of Science* **60**, 697-720

Hiley, B. (2013) “The early history of the Aharonov-Bohm effect” arXiv:1304.4736v1 [physics.hist-ph]

Klein, F. (1872) *Vergleichende Betrachtungen über neuere geometrische Forschungen*, Deichert, Erlangen

Leeds, S. (1999) “Gauges: Aharonov, Bohm, Yang, Healey” *Philosophy of Science* **66**, 606-27

Lyre, H. (2001) “The principles of gauging” *Philosophy of Science* **68**, S371-81

Lyre, H. (2002) “Zur Wissenschaftstheorie moderner Eichfeldtheorien” in A. Beckermann and C. Nimtz (editors) *Argument & Analyse – Sektionsvorträge*, Mentis, Paderborn

Lyre, H. (2004a) *Lokale Symmetrien und Wirklichkeit: eine Naturphilosophische Studie über Eichtheorien und Strukturenrealismus*, Mentis, Paderborn

Lyre, H. (2004b) “Holism and structuralism in $U(1)$ gauge theory” *Studies in History and Philosophy of Modern Physics* **35**, 643-70

Maudlin, T. (1998) “Healey on the Aharonov-Bohm effect” *Philosophy of Science* **65**, 361-8

Myrvold, W. (2011) “Nonseparability, classical and quantum” *The British Journal for the Philosophy of Science* **62**, 417-32

Olariu, S. and I. I. Popescu (1985) “The quantum effects of electromagnetic fluxes” *Reviews of Modern Physics* **57**, 339-436

Poincaré, H. (1902) *La science et l’hypothèse*, Flammarion, Paris

Poincaré, H. (1905) *La valeur de la science*, Flammarion, Paris

Russell, B. (1927) *The analysis of matter*, Kegan Paul, London

Schrödinger, E. (1935) “Die gegenwärtige Situation in der Quantenmechanik” *Naturwissenschaften* **23**, 807-12, 823-8, 844-9

Worrall, J. (1989) “Structural realism: the best of both worlds?” *Dialectica* **43**, 99-124

Wu, T. T. and C. N. Yang (1975) “Concept of nonintegrable phase factors and global formulation of gauge fields” *Physical Review D* **12**, 3845-57