

Functionals for Smooth Infinitesimals

Duncan Wilkie

25 January 2023

1 Introduction

The smooth infinitesimal analysis introduced by Lawvere and Kock, among others, through intense search for the proper topos for differential geometry, has promise as a better axiomatic implementation of the continuum. Chief among the reasons are the theory's computational simplicity, the indivisibility of the continuum, and constructive (and therefore computable) foundations—these are all exhibited in Bell's excellent monograph on the subject¹. One aspect that seems, at first glance, to impede its universal physical application is that *one cannot define discontinuous functions*. This paper argues that discontinuous functions that seem to arise in physics are actually never best regarded as functions at all. Using this insight, applicable to the ordinary reals, we tease out an analog in smooth worlds, and apply it to solve several *prima facie* discontinuous physical situations.

2 Physics Needs to Differentiate at Discontinuities

Consider the following example from Griffiths' *Introduction to Electromagnetism*.

Problem 1. *Imagine a very long solenoid with radius R , n turns per unit length, and current I . Coaxial with the solenoid are two long cylindrical (non-conducting) shells of length l —one, inside the solenoid at radius a , carries a charge $+Q$, uniformly distributed over its surface; the other, outside the solenoid, at radius b , carries charge $-Q$ (with $l \gg b$). When the current in the solenoid is gradually reduced, the cylinders begin to rotate. Where does the angular momentum come from?*

Solution 1. Initially, the fields are stationary: the electric field from the charges is (by Gauss's law, in the natural cylindrical coordinates)

$$\vec{E} = \frac{Q}{2\epsilon_0 l} \frac{1}{r} \hat{r} \quad (a < s < b)$$

and the magnetic is (by Ampere's law and some literal hand-waving)

$$\vec{B} = \mu_0 n I \hat{z} \quad (r < R)$$

When the current decreases, Faraday's law tells us that the changing magnetic field induces an electric field:

$$\oint_{\partial S} \vec{E} \cdot d\vec{\ell} = - \iint_S \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S}$$

¹Which, incidentally, is of such unparalleled quality to make me wish analytic philosophers wrote more pedagogical mathematics.

Taking S to be a disk slicing the whole cylindrical setup perpendicular to the common axis, one can chug out

$$E = \begin{cases} -\frac{1}{2}\mu_0 n \frac{dI}{dt} \frac{R^2}{s} \hat{\phi}, & (s > R) \\ -\frac{1}{2}\mu_0 n \frac{dI}{dt} s \hat{\phi}, & (s < R) \end{cases}$$

It's this electric field, tangent to the surface of the charges, that torques the cylinders in opposite directions; indeed this is what's physically observed. \square

Solution 2. With the same setup as before, apply Faraday's law in its equivalent form:

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

Pointwise, there's never a magnetic field outside the solenoid, so *there is no curl in the electric field outside the solenoid!* Namely, this means that there's no net perpendicular component to the electric field at the outer charged cylinder—and therefore no rotation! \square

What gives? The only difference is that we've changed which form of Maxwell's equations we're working with; they're supposed to be equivalent, right? The physicist's resolution of the paradox is to call it a failure of the idealized approximation: certainly, in any “very long solenoid,” there's some part of the magnetic field (however weak) that circles back to the other end, and it's probably just some jiggery-pokery with this mismatch causing the issue. However, this dodges the real meat of the issue: the integral formalism *just works*. A careful inspection of why the two forms of Maxwell's equations are (usually) equivalent reveals the error.

The theorem we're implicitly using when sensing a contradiction in these arguments is Stokes'.

Theorem 1. *Let S be a smooth, oriented (2-dimensional) surface in \mathbb{R}^3 with boundary ∂S . If a vector field $\vec{F}(x, y, z)$ is defined and has continuous first-order partials in a region containing S , then*

$$\iint_S (\nabla \times \vec{F}) \cdot d\vec{A} = \oint_{\partial S} \vec{F} \cdot d\vec{\ell}.$$

Immediately, we see that the primary hypothesis fails— \vec{B} jumps to 0 at the boundary of the solenoid. The takeaway is that the “more physical” version of Maxwell's equations are the integral forms; it's a strictly more general formalism. Indeed, this justifies many other subtle arguments deep in the weeds of E&M, such as the interface conditions for the fields when exiting media.

Is this just a quirk of electromagnetism? No—in fact, it's a much more general pattern in physics that the key mathematically-tractable examples are *fundamentally* discontinuous. There's a canonical example from quantum mechanics where the potential isn't even a *function*! The Dirac potential is a fundamental model for particle-particle scattering and strong-force capture.

The unifying mathematical machinery can be obtained by noticing that all examples seem to come from objects whose primary purpose is to “act” on functions to produce directly observable phenomena (in the end, a numerical measurement). For quantum potentials, this is especially obvious; for both electromagnetic fields and source distributions, it requires noticing that *trajectories* of charged particles are fundamental, and that fields and even sources are deduced from how trajectories are modified by their presence².

²If the lack of a physically measurable distinction between sources and the field they introduce bothers you, note that it's already implicit in the terminology! A source is nothing more than “that inducing a field,” and a field is nothing more than “that producing motion at a distance.” Review the great experimentalists' work if doubts remain; physics is the study of motion, after all, and it's only through analysis of motion that we deduce the physical existence of *anything*.

Nestruev noticed that the restriction of physical observation in a lab naturally places a smooth-manifold structure on anything classical. His excellent text develops the modern, sheaf-theoretic notions of differential geometry using this motivation; we'll adopt his viewpoint that a physics lab is a commutative, unital algebra; the sensors in that lab are elements of the algebra; physical states are algebra homomorphisms onto the trivial algebra structure on the reals; and sensor outputs are outputs of these homomorphisms. All this induces a manifold structure on the space of all physical states—exactly the configuration space.

These discontinuous objects arise from a higher-order consideration. An electric field is deduced by taking a particle that behaves trivially, changing something, and then observing (quantifying) how it changes. That is, such an action maps any measurement of a physical situation (a function from configuration space to the reals) to a measurement (real number) quantifying the effect of the change on the input measurement. The domain of this kind of a thing is clearly a vector space, and, given that putting additions and scalar multiplication circuits before or after the read-out shouldn't change the output, this identifies things like the electric field as **linear functionals**³.

3 Formalism

Such things are easily definable in smooth worlds, in which all subsequent notation ought to be interpreted (in particular, \mathbb{R} is the continuum of smooth infinitesimal analysis).

Definition 1. *The space $F_c(U)$ of compactly-supported functions on U is, for any open subset $U \subseteq \mathbb{R}^n$, the vector space of all functions $f : U \rightarrow \mathbb{R}$ such that $\{x \in U \mid f(x) \neq 0\}$ is contained in some closed hypercube in \mathbb{R}^n , with addition and scalar multiplication defined pointwise.*

Definition 2. *A distribution on $U \subseteq \mathbb{R}^n$ is a function $T : F_c(U) \rightarrow \mathbb{R}$ that is linear ($\forall v, w \in V : T(v + w) = T(v) + T(w)$) and homogeneous ($\forall v \in V \forall \alpha \in \mathbb{R} : T(\alpha v) = \alpha T(v)$).*

An important feature of classical distributions

In classical analysis, many distributions are justifiably confused with functions $\mathbb{R} \rightarrow \mathbb{R}$ by means of the Riesz representation theorem, whose realization for function spaces of interest allows one to identify many functionals with a partially applied inner product. Such confusion is far less generally possible in smooth worlds—often, the function fixed in the inner product is discontinuous. Whether this is a bug or a feature depends on whether the computational expedient is valued more than mandatory type-transparency.

³These, when the domain is a space of compactly-supported functions, are traditionally called **distributions**.