Jay R. Yablon
March 7, 2015

Complete Mathematical Solution Set to the Wu-Yang Differential Equation for $U(1)_{em}$ Dirac Magnetic Monopoles

Jay R. Yablon
910 Northumberland Drive
Schenectady, New York 12309-2814
jyablon@nycap.rr.com

March 7, 2015

Abstract: Employing Dirac's suggested approach to "perfect and generalise the mathematical formalism that forms the existing basis of theoretical physics," and thereafter "try to interpret the new mathematical features in terms of physical entities," we obtain the complete set of mathematical solutions to the Wu-Yang differential equation for Dirac monopoles to find three general classes of solutions. The first set includes the Dirac Quantization Condition but generalizes to encompass fractional charges also. The second set includes charge quantization without any fractionalization, in the absence of magnetic monopoles. The third set includes electric and magnetic charges but exhibits a breaking of the electric-magnetic duality symmetry not only due to the low-energy experimental coupling ~1/137, but even at the theoretical level. We then offer some preliminary physical interpretations and suggest a path for experimental validation based on the fractionalized charge solutions which are part of the first solution set.

PACS: 11.15.-q; 14.80.Hv; 73.43.Cd; 65.

Contents

Jay R. Yablon
March 7, 2015

1. Introduction: Dirac Monopoles and their Advancement by 't Hoof-Polyakov and Wu-Yang; and Dirac's Philosophy on the "Steady Progress of Physics"

Ever since James Clerk Maxwell first unified the separate laws of Gauss, Ampere and Faraday for electricity and magnetism, the observation of electric charges taken together with the non-observation of magnetic charges has been a source of intensive theoretical and experimental pursuit. Making use of gauge potentials A^μ which subsequently became an indispensable fixture of quantum electrodynamics (QED), in the modern language of differential forms where the one form $A = A_\mu dx^\mu$ and the field strength two form $F = dA$, the differential equations governing electricity are $*J = d*F = d*dA$ where * represents the duality operation (see [1] section 3.5), while those governing magnetism are $0 = dF = ddA$. The zeroing out of what would otherwise be the magnetic charge density, *by identity*, is caused by the differential forms identify *dd*=0 which is a geometric statement that the "exterior derivative of an exterior derivative is zero." When we apply Gauss' / Stokes' theorem to obtain the integral formulation of Maxwell's equations, the electricity laws become $e = \iiint *J = \oiint *F$ where *e*, often referred to as the electric charge strength, represents the total electric charge contained within the three-dimensional volume, and by Gauss / Stokes, is equal to the net electric field flux through the two-dimensional surface enclosing that volume. In contrast, the magnetism laws simply become $\mu = \iiint 0 = \oiint F = 0$ which states that there is no magnetic charge contained within the volume and therefore no net magnetic field flux over the closed two-dimensional surface around that volume. These relationships taken together, and specifically the absence of $e \leftrightarrow \mu$ interchange symmetry because $e \neq 0$ while $\mu = 0$, are often understood to mean that while there are electric charges *e* in nature, there are no magnetic charges μ, a.k.a. magnetic monopoles.

In 1931 Dirac [2] discovered *if* magnetic charges μ were to hypothetically exist, *then* this would imply that the electric charge *e* must be quantized. While charge quantization had been known for several decades based on the experimental work of Thompson [3] and Millikan [4], Dirac was apparently the first to lay out a possible theoretical imperative for this quantization. Using a hypothesized solenoid of singularly-thin width known as the Dirac string to shunt magnetic field lines out to mathematical infinity, Dirac established that *if* there was a non-zero magnetic flux $\mu = \oiint F$ across a closed surface, *then* this would be related to the electric charge strength *e* according to $e\mu = 2\pi n\hbar c$, where *n* is an integer. This became known as the Dirac Quantization Condition (DQC), because it could easily be rearranged into $e = ne_u$ in which $e_u \equiv 2\pi\hbar c/\mu$ defines a unit of electric charge in terms of the reciprocal of the magnetic charge, and into $\mu = n\mu_u$ in which $\mu_u \equiv 2\pi\hbar c/e$ likewise defines a unit of quantized magnetic charge. Clearly, the DQC is invariant under $e \leftrightarrow \mu$ interchange, which Dirac described as "a symmetry between electricity and magnetism quite foreign to current views," namely, foreign to the classical Maxwellian view in which $e = \oiint *F \neq 0$ but $\mu = \oiint F = 0$ and there is no $e \leftrightarrow \mu$ symmetry.

The electric charge strength *e* in $e\mu = 2\pi n\hbar c$ of the DQC is the same one which is related to the "running" fine structure coupling via $\alpha = e^2 / 4\pi\hbar c$, which, at low probe energies / large impact distance, approaches the numeric value $\alpha = e^2 / 4\pi\hbar c \cong 1/137.036$ asymptotically, see, e.g., equation (1) in Dirac's [2] ($\hbar c / e^2 = 137$, which uses Gaussian units), and Witten's [5], pages 27 and 28. Indeed, *Dirac's original purpose* for the derivation in [2] was to "give a theoretical value for *e*" thus the number ~137. However, the DQC left this number "from the theoretical standpoint, completely undetermined," and to date, despite many efforts to explain this number, this still is an experimentally-derived number with no commonly-accepted theoretical explanation. Dirac perceived it to be "rather disappointing to find this reciprocity between electricity and magnetism, instead of a purely electronic quantum condition, such as [Dirac's equation number] (1)." In general, from here, we shall use the natural units of $\hbar = c = 1$, and only restore $\hbar$ and *c* when necessary to illustrate a point.

Of course, given that the magnetic pole μ in $e\mu = 2\pi n$ was found to be anything other than the zero of the classical $\mu = \oiint F = 0$, it became important to explain *how this magnetic pole might exist under some domain of physical conditions, even though its observation had never been made experimentally accessible.* Dirac recognized that the $e \leftrightarrow \mu$ symmetry of $e\mu = 2\pi n$ in combination with the theoretically-undermined experimental value $e^2 / 4\pi \cong 1/137$ meant that the DQC "does not, however, force a complete symmetry, analogous to the fact that the symmetry between electrons and protons is not forced when we adopt Oppenheimer's interpretation [of filled negative energy "hole" states]." Specifically, Dirac observed that "if we insert the experimental value 137 in our theory [namely into $e\mu = 2\pi n$], it introduces quantitative differences between electricity and magnetism so large that one can understand why their qualitative similarities have not been discovered experimentally up to the present."

Specifically, he said, for there to be a "complete symmetry" between electricity and magnetism beyond the formal $e \leftrightarrow \mu$ interchange invariance, for the "one-quantum poles" with $n = 1$, we would have to have $\mu = e$ *numerically*, whereby the unit DQC would then become $e^2 = \mu^2 = 2\pi = 4\pi\alpha$, which would mean that $\alpha = \frac{1}{2}$, rather than the $\alpha \cong 1/137.036$ actually observed with low-impact probes. By representing the unit magnetic charge of the DQC as $\mu_0 = \left(2\pi / e^2\right)e \cong \left(137/2\right)e$ using the empirical α as observed at low probe energy, Dirac at page 72 of [2] observed that the force between any two magnetic poles μ would be larger by an approximate factor of "$\left(137/2\right)^2 = 4692\frac{1}{4}$" than that between any two electric charges *e*. "This very large force," Dirac concludes, "may perhaps account for why poles of opposite sign have never yet been separated." This was the first explanation of *how* this DQC-predicted quantized magnetic surface flux in $\oiint F = \mu = 2\pi n / e$ might be able to exist, even though it was not empirically observed at attainable experimental energies.

Our understanding of these monopoles substantially advanced in 1974 when 't Hooft and Polyakov [6], [7] demonstrated how the spontaneous symmetry breaking of a simply-connected gauge group *G* of a grand unified theory (GUT) down to a smaller subgroup which includes a *U*(1) factor would yield solutions with topologically-stable magnetic monopoles. In the

asymptotic large-distance limit, these solutions obey the Schwinger condition $e\mu = 4\pi$ for "magnetic monopoles with twice the flux quantum" whereby "two of the original strings, oriented in the same direction, can now annihilate by the formation of a monopole pair," see [6] at 276-277 and equation [2.22] (which uses Gaussian units). In these theories, however, if one includes "isospin $\frac{1}{2}$ representations of the group SU(2) describing charges" $e_u = Q = eT_3 = \pm\frac{1}{2}e$ where T^3 is the third component of the weak isospin generators, then the "monopoles do not obey Schwinger's condition" $e\mu = 4\pi$ but do recover Dirac's quantization condition $e\mu = 2\pi n$. Specifically, with the generators establishing the unit charge $e_u = \pm\frac{1}{2}e$, the Schwinger condition may be rewritten as $e\mu = \pm 2e_u\mu = 4\pi$ i.e. $e_u\mu = \pm 2\pi$ i.e. $ne_u\mu = \pm 2\pi n$. Then with $e = ne_u$ describing multiple units of charge and with the $\pm$ sign absorbed into $n = 0, \pm 1, \pm 2, \pm 3...$, this becomes $e\mu = 2\pi n$ which recovers the DQC. In these models, the monopole quantum is dependent upon the particular GUT employed. So, for example, in the Georgi-Glashow $SO(3) \sim SU(2)$ model [8] used in [6], and "[i]n Weinberg's SU(3)×SU(3) the monopole quantum is the Dirac one and in models where the leptons form an SU(3)×SU(3) octet . . . the monopole quantum is three times the Dirac value. . . ."

In all of these GUT models, the possible existence of these 't Hooft-Polyakov (TP) monopoles is reconciled with their empirical non-observation by virtue of their large predicted masses, in contrast to Dirac's Coulomb force explanation using "$(137/2)^2 = 4692\frac{1}{4}$." For example, 't Hooft used "Georgi and Glashow [8], based on SO(3), [with which] we can construct monopoles with a mass of the order of M_W, where M_W is the mass of the familiar intermediate vector boson. In the Georgi-Glashow model, $M_W < 53\,\mathrm{GeV}/c^2$," see [6] at 278-279. 't Hooft points out that "[o]nly in the Georgi-Glashow model (for which we did this calculation) is the parameter M_W . . . really the mass of the conventional intermediate vector boson. In other models it will in general be the mass of that boson which corresponds to the gauge transformations of the compact covering group."

As to the non-observation of monopoles at experimentally-attainable energies, 't Hooft states that "even in [the Georgi-Glashow] model the mass is so high that that might explain the negative experimental evidence so far. If Weinberg's SU(2)×U(1) model wins the race for the presently observed weak interactions [as it now appears to have done from a 2015 vantage point], then we shall have to wait for its extension to a compact gauge model, and the predicted monopole mass will be again much higher."

Weinberg expounds upon this in his definitive treatise [9] at 442:

> "The Georgi-Glashow model was ruled out as a theory of weak and electromagnetic interactions by the discovery of neutral currents, but magnetic monopoles are expected to occur in other theories, where a simply connected gauge group *G* is spontaneously broken not to U(1), but to some subgroup *H'*×*U*(1), where *H'* is simply connected. . . . There are no monopoles produced in the spontaneous breaking of the gauge group *SU*(2)×*U*(1) of the standard electroweak theory, which is not simply connected. . . . But we do find

> monopoles when the simply connected gauge group *G* of theories of unified strong and electroweak interactions . . . is spontaneously broken to the gauge group $SU(3)\times SU(2)\times U(1)$ of the standard model. . . . The monopoles in this case are expected to have a mass larger by an inverse square gauge coupling constant than the vector boson masses $M \approx 10^{15} - 10^{16}$ GeV produced by this symmetry breaking. Such monopoles would have been produced when the universe underwent a phase transition in which *G* was spontaneously broken to $SU(3)\times SU(2)\times U(1)$, at a temperature *T* of order *M*."

Weinberg then points out at 443 that:

> "[t]his poses a problem for some cosmological models. [Although the monopole number density at $T \approx M \approx 10^{15}$ GeV would have been] smaller than the photon density . . . by a factor . . . of order 10^{-6}, [if] monopoles did not find each other to annihilate, then this ratio would remain roughly constant to the present, but with at least 10^9 microwave background photons per nucleon today, this would give at least 10^3 monopoles per nucleon, in gross disagreement with what is observed. This potential paradox was one of the factors leading to inflationary cosmological models, in which there was a period of exponential expansion, which if it occurred before the monopoles were produced would have greatly extended the horizon [in the early universe], and if it occurred after the production of monopoles (but before a period of reheating) would have greatly diluted the monopole density."

In sum, the Georgi-Glashow SU(2) model which could have yielded an M_W-based monopole mass, possibly accessible with modern accelerators, was ruled out by electroweak neutral currents. What is left is that all empirically-viable GUT groups *G* must break down to $SU(3)\times SU(2)\times U(1)$, and for all such groups the monopole masses are at least 10^{15} GeV/c^2 which is the vev *v* about which the Higgs scalar fields used for symmetry breaking are expanded via $\phi(x^\mu) = v + h(x^\mu)$. And indeed, these masses must be even "larger by an inverse square gauge coupling constant," and so would be not far from the Planck mass defined by $GM_P{}^2 \equiv \hbar c$ which is $M_P \approx 1.2209\times 10^{19}$ GeV / c^2. This, coupled with the view that inflationary cosmology diluted the abundance of these monopoles in the early universe to such an extent that the probability of one ever being observed is vanishingly small, is the prevailing modern view of why these monopoles a) probably do exist in nature but b) are not experimentally observed.

In many ways, especially with the 't Hooft-Polyakov (TP) monopole mass tied to the GUT mass ~10^{15} GeV/c^2, the problem of observing monopoles is rooted in the same soil as the problem of observing proton decay. Both magnetic monopoles and proton decay are believed to occur in nature. Both involve exceptionally-high energies rooted in the GUT energy ~10^{15} GeV/c^2. Given that our experiments cannot attain anywhere near these energies, detecting either would require sifting through trillions upon trillions of empirical collision events in the statistical hope of getting lucky with a small few of these events. And given all of this, it has to date

proven impossible to detect either monopoles or proton decay experimentally. Now let's return to Dirac, and his philosophy about what is needed for "the steady progress of physics."

In his monopole paper [2] at 60, Dirac began by reflecting on how "modern physical developments have required a mathematics that continually shifts its foundations and gets more abstract." He had in mind such things as "[n]on-Euclidean geometry and non-commutative algebra, which were at one time considered to be purely fictions of the mind and pastimes for logical thinkers, [but] have now been found to be very necessary for the description of general facts of the physical world." And he also had in mind Herman Weyl's then-new formulation of gauge theory [10], [11] which was a mathematical abstraction of the collective physical investigations of Gauss, Ampere, Faraday and Maxwell. Two decades later this would be further abstracted by Yang and Mills [12] into non-Abelian gauge theories for weak and strong interactions and GUT interactions using non-commutative algebras which had once been the quaternion playground of Hamilton and had already informed Heisenberg in $\left[x^i, p^j \right] = i\hbar\delta^{ij}$ [13] and Dirac himself via his gamma matrices which obeyed $\eta^{\mu\nu} = \frac{1}{2}\left\{ \gamma^\mu, \gamma^\nu \right\}$ to reproduce Minkowski space and commuted as $\sigma^{\mu\nu} = \frac{1}{2} i \left[\gamma^\mu, \gamma^\nu \right]$. Dirac also had in mind the non-Euclidean geometry of General Relativity [14] centered on a Reimann tensor defined via $R^\sigma{}_{\alpha\mu\nu} V_\sigma \equiv \left[\partial_\mu, \partial_\nu \right] V_\alpha$ as a measure of the degree to which spacetime derivatives were non-commuting. And Dirac would momentarily use this same mathematical abstraction of gauge theory derive his DQC $e\mu = 2\pi n$.

Reflecting upon all this (except Yang-Mills which was still in the future), Dirac felt it:

> "likely that this process of increasing abstraction will continue in the future and that advance in physics is to be associated with a continual modification and generalisation of the axioms at the base of the mathematics rather than with a logical development of any one mathematical scheme on a fixed foundation."

He then concluded that:

> ". . . fundamental problems in theoretical physics [will require solutions] beyond the power of human intelligence to get the necessary new ideas by direct attempts to formulate the experimental data in mathematical terms. . . . *The most powerful method of advance that can be suggested at present is to employ all the resources of pure mathematics in attempts to perfect and generalise the mathematical formalism that forms the existing basis of theoretical physics, and after each success in this direction, to try to interpret the new mathematical features in terms of physical entities . . .*"

And this finally brings us to Wu and Yang. Not long after 't Hooft and Polyakov demonstrated how topologically-stable magnetic monopoles having asymptotic large-distance behaviors matching the DQC $e\mu = 2\pi n$ could emerge from the spontaneous symmetry breaking of a simple compact GUT group *G*, Wu and Yang showed how to obtain "Dirac monopoles without strings." [15], [16] In fact the TP monopole did already replace the singular Dirac string

with a smooth Higgs scalar field $\Phi = T_i \phi_i$ used in the symmetry breaking where T_i commuting as $\left[T_i, T_j \right] = i f_{ijk} T_k$ with group structure constants f_{ijk} are the generators of *G*. But the Wu-Yang monopole development eschewed the type of physical arguments involving flux lines and contours and the physical monopole Lagrangian that had hitherto been used, in favor of a formal mathematical analysis based on $U(1)_{em}$ gauge theory alone. This opens up what Dirac would regard as a "most powerful method of advance," because what Wu and Yang did was to effectively lay bare the "mathematical formalism that forms the existing basis of theoretical physics" for Dirac monopoles.

Specifically, up until Wu and Yang, the DQC $e\mu = 2\pi n$ had been understood as an *algebraic relation* involving the electric and magnetic charges *e* and μ and a quantum number *n*. Now, starting with an electron wavefunction transforming as $\psi(x) \to \psi'(x) = e^{i\Lambda(x)} \psi(x)$ where $\Lambda(x)$ is a local gauge (really, phase) angle and where the field strength two form $F = (\mu / 4\pi) d\cos\theta d\varphi$, Wu and Yang essentially showed the DQC to be *an algebraic solution to the differential equation* $e^{-i\Lambda} de^{i\Lambda} / ie = (\mu / 2\pi) d\varphi$, when this equation is solved using the two azimuth angles $\varphi = 0$ and $\varphi = 2\pi$ which of course are the same as one another when viewed strictly in terms of their geometric orientations. By finding the differential equation $e^{-i\Lambda} de^{i\Lambda} / ie = (\mu / 2\pi) d\varphi$ "at the base of the mathematics" for which the DQC $e\mu = 2\pi n$ was simply <u>a</u> solution, Wu and Yang provided the means – which we shall exploit in this paper – to "perfect and generalise the mathematical formalism that forms the existing basis of [the] theoretical physics" underlying magnetic monopoles. And, once this is achieved, it then becomes possible "to try to interpret the new mathematical features in terms of physical entities."

In this paper, we shall complete the first step of this two-step prescription recommended by Dirac, as regards the mathematical formalism of Wu-Yang that is used to obtain Dirac's own DQC as one of its solutions. Specifically, approaching the problem mathematically, we shall simply study $e^{-i\Lambda} de^{i\Lambda} / ie = (\mu / 2\pi) d\varphi$, as a *mathematical* differential equation in which $de \neq 0$ and $d\mu \neq 0$, i.e., in which *e* and μ do not have constant values throughout spacetime as is well-evidenced by the renormalization-based "running" of the electric charge strength as a function of how spatially close one is able to collide a test charge with a charge being tested. As we shall see, if we allow these charges to run and do not impose any *a priori* constraint on their running, we are mathematically required to generalize the Wu-Yang differential equation to $e^{-i\Lambda} de^{i\Lambda} / ie = (\mu / 2\pi) d\varphi + \varepsilon$ where ε defines an observable distinctness between the north and south gauge field patches of the Wu-Yang analysis which arise precisely because of this permitted running.

We then develop *all* of the mathematically-predicted solutions for $\varepsilon = 0$ and $\varepsilon \neq 0$, and for all like-oriented azimuths $\varphi = 2\pi m$ where $m = 0, 1, 2, 3...$ is an integer. In doing so, we do not restrict ourselves to $\varphi = 0, 2\pi$ azimuths for which the DQC solves the $\varepsilon = 0$ differential equation, because *there is nothing in the mathematics which requires us to be so restrictive*. Indeed, after finding that the non-integrable derivatives of the wavefunction phase are synonymous with the potential of the electromagnetic field and so "gives us nothing new," Dirac

then highlighted "one further fact which must now be taken into account, namely, that a phase is always undetermined to the extent of an arbitrary integral multiple of 2π. This," he then said, "requires a reconsideration of the connection between the [non-integrable derivatives] and the potentials and leads to a new physical phenomenon" which turned out to be the DQC. In the Wu-Yang derivation, the azimuth also has a lead role in the differential equation $e^{-i\Lambda}de^{i\Lambda}/ie=(\mu/2\pi)d\varphi+\varepsilon$, and just like the phase, an azimuth is also "always undetermined to the extent of an arbitrary integral multiple of 2π." This too "leads to a new physical phenomenon," or – since we shall focus here on mathematics – leads to "new mathematical features" which include electric charge fractionalization.

While one might contend that such fractional charges violate the DQC and are therefore unacceptable, this is a misguided view for two reasons. First, the DQC is an algebraic equation which is *a mathematical solution* of the Wu-Yang differential equation, but not the only solution. What Dirac calls the "base of the mathematics" *is not the DQC, but is the Wu-Yang differential equation* $e^{-i\Lambda}de^{i\Lambda}/ie=(\mu/2\pi)d\varphi+\varepsilon$. The question is not whether something "violates" the DQC, but whether something violates Wu-Yang's $e^{-i\Lambda}de^{i\Lambda}/ie=(\mu/2\pi)d\varphi+\varepsilon$ as generalized to running charges with no *a priori* constraints. If the DQC is a solution to the differential equation, and if fractional charges are also a solution to the same differential equation, then fractional charges are consistent with the DQC because they both solve the same underlying differential equation.

To use an analogy from general relativity, the Schwarzschild solution to the Einstein equation was found in 1916. The Friedmann solution to the same differential equation was found in1924. To argue that fractional charges violate the DQC when they are both mathematical solutions to the same underlying differential equation is the same as arguing that the Friedmann solution violates the Schwarzschild solution. This is an apples-to-oranges comparison that does not stand up. And indeed, the Einstein equation is an excellent illustration of how a deep and general differential equation is much more powerful and explains many more phenomena that any individual solution to that differential equation. It is, in fact, highly unlikely that "the power of human intelligence" could have conceived of either cosmological or local gravitationally-collapsed conditions with mathematical specificity, absent having ascertained the correct differential equation from which these effects may be deduced as solutions. Similar things may be said of Maxwell's equations, and their extension into QED. That is why the differential equation $e^{-i\Lambda}de^{i\Lambda}/ie=(\mu/2\pi)d\varphi+\varepsilon$ of Wu and Yang, and not one of its algebraic solutions $e\mu=2\pi n$, must be the starting point for any generalized analysis of Dirac monopoles.

Second, such a contention blurs a clear line between mathematics and physics which Dirac wisely and accurately articulates. To be sure: after one is able to fully "generalise the mathematical formalism that forms the existing basis of theoretical physics" for Dirac monopoles, namely the Wu-Yang differential equation, there is much room to argue about how "to interpret the new mathematical features in terms of physical entities." One can and should engage in open, principled and transparent debate about how the full set of solutions to the Wu-Yang differential equation should be physically interpreted and whether these solutions do or do not apply to certain phenomena observed in the natural world. But the solutions to the mathematics are the solutions to the mathematics and cannot be bent by human will or

intelligence. How one thinks about the application of these solutions to physics is *independent* of what the solutions themselves actually are, and these must be separated in one's thinking. Otherwise one's preconceptions and wishes about the natural world cannot help but stand in the way of clear and receptive thinking about the natural world, because in modern physics it is often "beyond the power of human intelligence to get the necessary new ideas" *other than* through pure mathematical deduction from the "base of the mathematics . . . that forms the existing basis of theoretical physics." Dirac wisely and correctly observes that *deriving mathematical solutions*, and then *physically interpreting these solutions,* are two completely distinctive aspects of how to facilitate the "steady progress of physics." In so-doing he lays out a methodological discipline which is not nearly as heeded or practiced as it ought to be.

With these introductory observations, we shall now carefully enumerate *all* the available *mathematical* solutions of the Wu-Yang differential equation, recognizing that this may well be followed by lively debate about how "to interpret the new mathematical features in terms of physical entities." We open that debate by pointing out along the way, certain aspects of these mathematical solutions which are strikingly reminiscent of certain observed natural phenomena, especially charge fractionalization because there is a fractionalization observed in the Fractional Quantum Hall Effect (FQHE). But we do so without present claim about whether these reminiscences will end up being validated or contradicted following critical theoretical and experimental development and review. To the extent that we offer a few physical observations amidst the mathematical development, we do so simply to provide some initial thoughts about possible approaches to physically interpret the mathematical solutions that will now be laid out.

2. Wu and Yang and the Mathematization of Dirac Monopole Study – Constant Electric and Magnetic Charge Strengths

As just stated, we shall now develop mathematical solutions of Wu and Yang's differential equation $e^{-i\Lambda}de^{i\Lambda}/ie=(\mu/2\pi)d\varphi$ [15], [16] for electric and magnetic charge strengths for which $de\neq 0$ and $d\mu\neq 0$ and for which there is no *a priori* constraint imposed on how de is mathematically related to $d\mu$. This is distinct from Dirac's second step of trying to "interpret the new mathematical features in terms of physical entities," that is, of studying whether these solutions map sensibly to phenomena which are empirically observed in the natural world. To start, we study solutions for which the local charge strengths are taken to be constants such that $de=d\mu=0$. Then we study expanded solutions in which the local charge strengths are allowed to "run," i.e., for which $de\neq 0$ and $d\mu\neq 0$, with no *a priori* constraint between de and $d\mu$. As we shall see, the former correspond precisely to $\varepsilon=0$ with no observable difference between the gauge patches while the latter correspond with equal precision to $\varepsilon\neq 0$ in which there is an observable distinctness brought about by the running of the charge strengths. Of course, if we are going to develop and then study mathematical solutions to this Wu-Yang differential equation, it would be good to review why this equation is of physical interest to begin with. Zee at 220-221 of [17] provides a concise review of the Wu-Yang approach, which we briefly lay out here.

Jay R. Yablon
March 7, 2015

Wu and Yang begin with a local gauge (really, phase) transformation on a fermion wavefunction taken to be that of an electron, $\psi(x) \rightarrow \psi'(x) = e^{i\Lambda(x)}\psi(x)$. The one-form for the gauge field $A = A_\mu dx^\mu$ transforms as $A \rightarrow A' = A + e^{-i\Lambda}de^{i\Lambda}/ie$, where $e^{-i\Lambda}de^{i\Lambda}/ie = d\Lambda/e$ is a mathematically-convenient way to represent the exact closed one-form $d\Lambda = \partial_\mu \Lambda dx^\mu$ containing the phase gradient $\partial_\mu \Lambda$. If we represent the local field strength in spherical coordinates (t, r, φ, θ) as $F = (\mu/4\pi)d\cos\theta d\varphi$, then using *dd*=0, with μ presently assumed to be constant thus $d\mu = 0$, this will be reproduced via $F = dA$ by any $A = (\mu/4\pi)(\cos\theta - K)d\varphi$ with constant *K*.

Of course, the azimuth φ in $d\varphi$ is not defined everywhere, and in particular, it is undefined – which is to say that the longitude is indeterminate – at the north and south poles of the closed surface in $\mu = \oiint F$. This is an intrinsic feature of closed surfaces in three space dimensions. To remove this indefiniteness, with $K = \pm 1$ respectively, we define a north patch for the gauge field $A_N \equiv (\mu/4\pi)(\cos\theta - 1)d\varphi$ with the south pole undefined and a south patch $A_S \equiv (\mu/4\pi)(\cos\theta + 1)d\varphi$ with the north pole undefined. These patches are interrelated by $A_S = A_N + (\mu/2\pi)d\varphi$. So if we now "match" these north and south patches by regarding the south patch to differ from the north by no more than a gauge transformation, i.e., if we define $A'_N \equiv A_S$ and use this to write $A'_N \equiv A_S = A_N + (\mu/2\pi)d\varphi$, and if we then combine this with the gauge transformation $A_N \rightarrow A'_N = A_N + e^{-i\Lambda}de^{i\Lambda}/ie$ for the north patch, we obtain the above-referenced Wu-Yang differential equation:

$$\frac{1}{ie}e^{-i\Lambda}de^{i\Lambda} = \frac{\mu}{2\pi}d\varphi. \tag{2.1}$$

By making use of $A'_N \equiv A_S$, we are making the *assumption* that the north and south gauge field patches differ from one another by no more than a gauge transformation and so are not observably distinct. *Were there to hypothetically be* some observable distinctness between these patches, which distinctness we may define by ε in $A'_N \equiv A_S + \varepsilon$ where $\varepsilon \equiv \varepsilon_\mu dx^\mu$ is a differential one form and ε_μ so-defined must therefore be a four-vector field with dimensions of energy just like the gauge field in $A = A_\mu dx^\mu$, *then* (2.1) would become generalized to:

$$\frac{1}{ie}e^{-i\Lambda}de^{i\Lambda} = \frac{\mu}{2\pi}d\varphi + \varepsilon. \tag{2.2}$$

The above (2.2) serves in Dirac's words to "perfect and generalise the mathematical formalism" of (2.1) with an additional offset ε, whereby (2.1) is the $\varepsilon = 0$ specialization of (2.1). As we shall establish shortly, when μ and *e* viewed as mathematical entities are taken to be constant, i.e., with $d\mu = 0$ and $de = 0$, then $\varepsilon = 0$.

From here, the remainder of this paper will seek simply to obtain all mathematical solutions of (2.1) and (2.2), to lay the foundation, per Dirac, to "try to interpret the new mathematical features [of these solutions] in terms of physical entities." We start with (2.1), which is (2.2) with $\varepsilon = 0$, thus with observably-*indistinct* coordinate patches. This is the differential equation first developed by Wu and Yang.

For notational convenience, we define a "reduced azimuth" $\varphi \equiv \varphi / 2\pi$. The solution to the differential equation (2.1) in terms of φ is:

$$\exp\left(i\Lambda\right) = \exp\left(ie\mu\varphi\right), \tag{2.3}$$

as is easily seen by plugging $\exp\left(i\Lambda\right)$ from (2.3) into the left hand side of (2.2) and reducing. For $\varphi = 0$, (2.3) becomes $\exp\left(i\Lambda\right) = \exp\left(ie\mu \cdot 0\right) = 1$. Because $1 = \exp\left(i2\pi n\right)$ in general, this means that $\exp\left(i\Lambda\right) = 1 = \exp\left(i2\pi n\right)$, which has the general solution:

$$\Lambda = 2\pi n. \tag{2.4}$$

For similar notational convenience, we also define a "reduced phase" $\mathbb{\Lambda} \equiv \Lambda / 2\pi$, which enables us to write the solution (2.4) as $\mathbb{\Lambda} = n = 0, \pm 1, \pm 2, \pm 3...$. This solution represents the quantized number of windings of the phase Λ through the complex gauge space defined by $e^{i\Lambda} = \cos\Lambda + i\sin\Lambda = a + bi$.

From (2.3), if we require that $\exp\left(ie\mu\varphi\right) = \exp\left(ie\mu \cdot 0\right) = 1$ for all reduced azimuths in the quantized set $\varphi \equiv \varphi / 2\pi = 0, 1, 2, 3...$ due to the fact that each of these φ has *identical geometric orientation* (presently, we neglect so-called "entanglement" whereby like "versions" of a fermion have a 4π azimuth separation), then in view of $1 = \exp\left(i2\pi n\right)$, the mathematics enables us to write $\exp\left(i\Lambda\right) = \exp\left(ie\mu\varphi\right) = 1 = \exp\left(i2\pi n\right)$, which has the general solution

$$\Lambda = e\mu\varphi = 2\pi n. \tag{2.5}$$

The azimuth set $\varphi = m = 0, 1, 2, 3...$ used to obtain (2.5) represents the quantized number of (positive, right-handed) rotations or "windings" over a complete 2π circumference about the z axis in the physical space of spacetime. The integer *m* in $\varphi = m$ is a different integer from the integer *n* in $\mathbb{\Lambda} = n$ obtained from (2.4), the former representing physical space windings and the latter representing gauge space windings. As earlier noted, Dirac first derived his quantization condition upon pointing out "that a phase is always undetermined to the extent of an arbitrary integral multiple of 2π," which responsible for (2.4). The fact that an azimuth orientation is likewise undetermined up to an integer multiple of 2π is responsible for (2.5).

For the special case of an azimuth winding $\varphi = 1$, (2.5) becomes

$$\Lambda = e\mu = 2\pi n\,, \tag{2.6}$$

which is the Dirac Quantization condition. In other words, the DQC $e\mu = 2\pi n$ is the $\not{\varphi} = 1$ and $\varepsilon = 0$ solution of the Wu and Yang differential equation (2.2). Because $e = \sqrt{4\pi\alpha}$ where $\alpha \cong 1/137$ at low probe energies i.e., large distance from a bare charge, and because the charge quantum number $\not{A} = n$, we may regard a single $n = 1$ electron charge quantum as the $\not{\varphi} = 1$, $\not{A} = 1$, $\varepsilon = 0$ solution of the differential equation (2.2). This shows why (2.1) and its generalization (2.2) are of such interest. For, while (2.6) is merely an *algebraic equation*, (2.2) is the *differential equation* for which the DQC of (2.6) is <u>*a*</u> solution, but *not the only solution*. Thus, (2.2) provides the means to "generalise the mathematical formalism" associated with magnetic monopoles – or to be precise, associated with the asymptotic behavior of the TP monopoles which reproduce the DQC at large distance. Now it is time to examine some of the other *mathematical* solutions to the differential equation (2.2), beyond the specific solution $e\mu = 2\pi n$ which is the DQC and the specialization of that solution to $e\mu = 2\pi$ for the $\not{A} = n = 1$ unit electron charge.

Immediately, we restructure the generalized $\varepsilon = 0$ solution of (2.2), namely (2.5), into:

$$e = \frac{\not{A}}{\not{\varphi}}\frac{2\pi}{\mu} = \frac{\not{A}}{\not{\varphi}}e_{\mathrm{u}} = \nu e_{\mathrm{u}}\,, \tag{2.7}$$

where $e_{\mathrm{u}} \equiv 2\pi/\mu$ continues to define the unit electric charge as usual, and where:

$$\nu \equiv \frac{\not{A}}{\not{\varphi}};\quad \not{A} = 0, \pm 1, \pm 2, \pm 3...;\quad \not{\varphi} = 0, 1, 2, 3... \tag{2.8}$$

reveals that the solutions to (2.1) also admit *fractionalized* charges. Because the we are only seeking to obtain mathematical solutions to (2.1) and (2.2) separately from the question of physical interpretation, we do no more at the moment than make a mental note that the Fractional Quantum Hall Effect (FQHE) exhibits fractional charges albeit with denominators more restricted than those in (2.7), (2.8), and that the FQHE is an ultra-low temperature phenomenon observed only near 0K which therefore requires us to pay attention to the variable of temperature, and thus to thermodynamics.

Because $\not{\varphi} = 1$ describes quantized solutions of a whole electric charge, any version of this unit electron with a different azimuth but the same version (i.e., same orientation and entanglement) as the $\not{\varphi} = 1$ will differ from this by a 4π azimuth separation. Thus, the alike versions of the $\not{\varphi} = 1$ electron with a positive azimuth will have $\not{\varphi} = 1, 3, 5, 7... = 2l + 1$ with $l = 0, 1, 2, 3...$, which $\not{\varphi}$ is an odd integer, so that (2.8) for like-*versions* (orientation and entanglement) of the $\not{\varphi} = 1$ electron will become restricted to:

$$\nu \equiv \frac{\not{A}}{\not{\varphi}};\quad \not{A} = 0, \pm 1, \pm 2, \pm 3...;\quad \not{\varphi} = 2\left(l + \tfrac{1}{2}\right) = 2\left(l + s\right) = 2j = 1, 3, 5, 7...;\quad l = 0, 1, 2, 3...;\quad s = \tfrac{1}{2}. \tag{2.9}$$

Here, the charge fractionalization becomes odd-integer only, while the $\varphi = 0$ denominator which yields an infinite ν is naturally eliminated. With the exception of what would have to be a $\varphi = 2$, $l = \frac{1}{2}$, $s = \frac{1}{2}$ fraction also observed, this does describe precisely the FQHE fractionalization. These thoughts are intended merely to help us "to try to interpret the new mathematical features in terms of physical entities," as the second step of the Dirac approach, and are not intended to make any claim.

So for now, mathematically, we may make the following objective summary of the solutions we have found:

- The $\varepsilon = 0$, $\mu \neq 0$ solution of (2.2) includes fractionalized and quantized charges.
- The $\varepsilon = 0$, $\mu \neq 0$, $\varphi = 1$ solution of (2.2) is the DQC with quantized charges only.
- The $\varepsilon = 0$, $\mu \neq 0$, $\varphi = 1$, $\Lambda = 1$ solution of (2.2) describes a unit electron charge.
- The set of solutions which have the same orientation and entanglement as the $\varphi = 1$ unit electron charge, have a fractional charge denominator $\varphi = 2\left(l + \frac{1}{2}\right)$ which is an odd integer only.
- The quantum numbers *l* and *s* and *j* defined in (2.9) to characterize the same orientation and entanglement solutions happen to be identical to the Casimir numbers obtained in the operations $\mathbf{L}^2\left|\xi\right\rangle = l\left(l+1\right)\left|\xi\right\rangle$, $\mathbf{S}^2\left|\xi\right\rangle = s\left(s+1\right)\left|\xi\right\rangle$ and $\mathbf{J}^2\left|\xi\right\rangle = j\left(j+1\right)\left|\xi\right\rangle$ on a spinor $\left|\xi\right\rangle$ for orbital, spin and total angular momentum of an electron in an atomic shell.

Beyond these specific features, we may also observe, mathematically, that these solutions to (2.1) are *topologically quantized*, insofar as they *naturally introduce* the two quantum numbers Λ and φ which specify an integer number of windings, respectively, in the two-dimensional gauge space, and in the three-dimensional physical space about the z-axis.

Now, let's find the more general solutions to (2.2) for both $\mu \neq 0$ and $\varepsilon \neq 0$. Because the one-form ε must be a field $\varepsilon(x^\mu) = A'_N(x^\mu) - A_S(x^\mu)$ with structure and dimensionality similar to that of the gauge field, see the discussion prior to (2.2), let us posit a zero-form scalar field $\tau(x^\mu)$ related in some to-be-determined way to the one-form field $\varepsilon(x^\mu)$, as a mathematical means to solve (2.2). In a spherical coordinate system $x^\mu = \left(t, r, \varphi, \theta\right)$, the azimuth φ is one of the four spacetime coordinates of which these fields are a function. Using this $\tau(x^\mu)$ in a posited test expression we write:

$$\exp\left(i\Lambda\right) = \exp\left(ie\mu\varphi + ie\tau\left(t, r, \varphi, \theta\right)\right). \tag{2.10}$$

If we insert this in (2.2) and reduce, we find that this does indeed solve (2.2), *if and only if*

$$\varepsilon = d\tau = \varepsilon_\mu dx^\mu = \partial_\mu \tau dx^\mu. \tag{2.11}$$

Extracting the vectors, this means that $\varepsilon_\mu = \partial_\mu \tau$, that is, $\varepsilon_\mu(x^\mu)$ is the spacetime gradient of the posited scalar field $\tau(x^\mu)$. Via *dd*=0, this informs us that $d\varepsilon = dd\tau = 0$, and via Gauss / Stokes, that $\oint \varepsilon = 0$. This means that $\varepsilon = d\tau$ is a closed exact differential, and that τ may be thought of as a state variable.

At the moment, the particular coordinate in $\tau(t,r,\varphi,\theta)$ that piques our interest is the azimuth $\varphi \subset t,r,\varphi,\theta$ and its reduced form $\not{\varphi} = \varphi / 2\pi$ because this turns out to be the fractionalization denominator in (2.7) and (2.9). So to avoid visual clutter and focus in on $\tau(\not{\varphi})$ which is presently of greatest interest, let us define $\tau_{\not{\varphi}} \equiv \tau(\not{\varphi})$ as a shorthand notation to represent the azimuthal behavior of the scalar field τ. With this notation targeted to the spacetime coordinate presently of greatest interest, we rewrite (2.10) as:

$$\exp(i\Lambda) = \exp\left(ie\mu\not{\varphi} + ie\tau_{\not{\varphi}}\right). \tag{2.12}$$

Now, as we did at (2.4), let us first examine what happens in (2.10) when we set $\not{\varphi} = 0$ to obtain $\exp(i\Lambda) = \exp(ie\mu \cdot 0 + ie\tau_0) = \exp(ie\tau_0)$. If we multiply through by $\exp(-ie\tau_0)$ and then apply $1 = \exp(i2\pi n)$, we obtain $\exp(i\Lambda - ie\tau_0) = 1 = \exp(i2\pi n)$, which has the solutions:

$$\Lambda - e\tau_0 = 2\pi n\,. \tag{2.13}$$

Contrasting to (2.4) we see that the reduced gauge angles $\not{\Lambda} = n + e\tau_0 / 2\pi$ which solve (2.12) are still quantized, but they also have an offset $e\tau_0 / 2\pi$. Now, while discussing fermion wavefunctions ψ at page 63 of [2], and as is well-understood, Dirac points out that "[t]he indeterminacy in ψ then consists in the possible addition of an arbitrary constant to the phase γ [here, Λ]. Thus the value of γ at a particular point has no physical meaning and only the difference between the values of γ at two different points is of any importance." So while the quantum number *n* in (2.13) is an observable because it denotes a phase *difference*, the offset $e\tau_0 / 2\pi$ has no physical meaning and so can be set to zero to establish a "ground" state. Doing exactly that, we *define* a ground state phase:

$$\tau_0 = \tau(\varphi = 0) \equiv 0\,. \tag{2.14}$$

As a result, (2.13) becomes $\Lambda = 2\pi n$ which is the same as (2.4), and we continue to represent this via the reduced phase $\not{\Lambda} = n = 0, \pm 1, \pm 2, \pm 3...$ which is a topological winding number. With (2.14), this means that for $\not{\varphi} = 0$, (2.12) becomes $\exp\left(ie\mu\not{\varphi} + ie\tau_{\not{\varphi}}\right) = \exp(ie\mu \cdot 0 + ie\tau_0) = 1$.

Next, similarly to (2.5), we require that $\exp\left(ie\mu\not{\varphi} + ie\tau_{\not{\varphi}}\right) = 1$ in (2.12) for all reduced azimuths which have the same orientation as $\not{\varphi} = 0$, again, sans present consideration of

entanglement. These are all the azimuths in the quantized set $\not{\varphi} \equiv \varphi / 2\pi = 0,1,2,3...$. Again, just like the phase, the azimuth also is "undetermined to the extent of an arbitrary integral multiple of 2π." Then, (2.12) yields $\exp\left(ie\mu\not{\varphi} + ie\tau_{\not{\varphi}}\right) = 1 = \exp\left(i2\pi n\right)$ in view of $\exp\left(i2\pi n\right) = 1$. This has the solution:

$$\Lambda = e\mu\not{\varphi} + e\tau_{\not{\varphi}} = 2\pi\not{\text{A}}; \quad \not{\text{A}} = \Lambda / 2\pi = n = 0, \pm 1, \pm 2, \pm 3...; \quad \not{\varphi} = \varphi / 2\pi = m = 0,1,2,3.... \tag{2.15}$$

This is the full solution to the differential equation (2.2), for $\varepsilon = d\tau \neq 0$ *(see (2.11)) and* $\tau_{\not{\varphi}} \neq 0$ *generally*, with the unobservable ground state phase defined in (2.14) as $\tau_0 \equiv 0$. For the specialization $\tau_{\not{\varphi}} = 0$ generally, this reduces to $\Lambda = e\mu\not{\varphi} = 2\pi n$ which is (2.5). For $\tau_{\not{\varphi}} = 0$ and $\not{\varphi} = 1$ this further reduces to $\Lambda = e\mu = 2\pi n$ which is the DQC. And for $\tau_{\not{\varphi}} = 0$, $\not{\varphi} = 1$, $\not{\text{A}} = 1$ we obtain $\Lambda = e\mu = 2\pi$ for the unit electron charge. For states which have the same version as the $\not{\varphi} = 1$ electron, as in (2.9), the quantum number $\not{\varphi} = 2\left(l + \frac{1}{2}\right) = 2\left(l + s\right) = 2j = 1,3,5,7...$ will be restricted to odd integer values only. Per Dirac, (2.15) is how we "generalise the mathematical formalism that forms the existing basis of theoretical physics" for describing the asymptotic large-distance behavior of DQC magnetic monopoles.

The solution (2.15) now enables us to examine another specialization, namely that of $\tau_{\not{\varphi}} \neq 0$ but $\mu = 0$. This set of solutions cannot be seen at all just using the algebraic $\Lambda = e\mu = 2\pi n$ version of the DQC which for $\mu = 0$ becomes the trivial $\Lambda = 2\pi n = 0$. But from (2.15), when $\mu = 0$, we still have the non-trivial:

$$\Lambda = e\tau_{\not{\varphi}} = 2\pi n. \tag{2.16}$$

If we then isolate *e*, this becomes:

$$e = \not{\text{A}}\frac{2\pi}{\tau_{\not{\varphi}}} = n\frac{2\pi}{\tau_{\not{\varphi}}} = ne_{\text{u}}, \tag{2.17}$$

from which we may define a unit electric charge:

$$e_{\text{u}} \equiv \frac{2\pi}{\tau_{\not{\varphi}}}. \tag{2.18}$$

This now replaces the usual unit charge $e_{\text{u}} \equiv 2\pi / \mu$ of the DQC, and still gives us a quantized electric charge *even if we set* $\mu = 0$ *and so take away all magnetic monopoles*. In contrast to (2.7), there is no charge fractionalization in (2.17). There is only quantization. A simple variant of (2.16) is to isolate $\tau_{\not{\varphi}}$, and so write:

$$\tau_{\varphi} = \mathbb{A}\frac{2\pi}{e}. \tag{2.19}$$

Staying focused on mathematical solutions to (2.1) and (2.2), we simply regard the $\mu = 0$ solutions as ones in which a mathematical parameter μ has been set to zero. Of course in the natural world, this represents a magnetic charge strength observed to be equal to zero. So the objective mathematical statements we may make about this solution set are as follows:

- The $\varepsilon \neq 0$, $\mu = 0$ solution of (2.2) includes *only* quantized $\mathbb{A} = n$ electric charges with no fractionalization.
- For the $\varepsilon \neq 0$, $\mu = 0$ solution of (2.2), the $\mathbb{A} = 1$ electric charge quantum is $e_{\text{u}} = 2\pi / \tau_{\varphi}$.
- In contrast, for the $\varepsilon = 0$, $\mu \neq 0$ solution of (2.2), the electric charge quantum is the $e_{\text{u}} = 2\pi / \mu$ of the DQC.

Next, let us spend a few moments examining the complete solution (2.15) for the circumstance where both $\mu \neq 0$ and $\tau_{\varphi} \neq 0$. First it is helpful to isolate the new scalar field τ_{φ} in (2.15) by itself, thus:

$$\tau_{\varphi} = \mathbb{A}\frac{2\pi}{e} - \varphi\mu. \tag{2.20}$$

It is also helpful to isolate the electric charge strength as such:

$$e = \mathbb{A}\frac{2\pi}{\mu\varphi + \tau_{\varphi}}. \tag{2.21}$$

This displays vividly how when $\tau_{\varphi} = 0$ and $\mu \neq 0$ there is a fractional quantization $e = \left(\mathbb{A} / \varphi\right)\left(2\pi / \mu\right)$ with unit charge $e_{\text{u}} \equiv 2\pi / \mu$, see (2.7) and (2.8), while when $\mu = 0$ and $\tau_{\varphi} \neq 0$ there is a quantization-only with no fractionalization, in the form $e = \mathbb{A}\left(2\pi / \tau_{\varphi}\right)$ with unit charge $e_{\text{u}} \equiv 2\pi / \tau_{\varphi}$, see (2.17) and (2.18).

As to the physics which we defer until we "try to interpret the new mathematical features in terms of physical entities," we simply make a mental note that fractional electric charge quantization of the sort that solves (2.2) when $\mu \neq 0$ and $\tau_{\varphi} = 0$ has only been observed in nature near absolute zero T=0, and that electric charge quantization with no fractionalization of the sort that solves (2.2) when $\mu = 0$ and $\tau_{\varphi} \neq 0$ is observed throughout the natural world when the temperatures $T > 0$ are not right near absolute zero. So (2.21) where both $\tau_{\varphi} \neq 0$ and $\mu \neq 0$ reveals a sort of hybrid of these two results in which the *electric charge is always quantized*, but the denominator $\mu\varphi + \tau_{\varphi}$ contains a quantized azimuth winding φ offset with the field τ_{φ}

which can drive a modified form of fractionalization. Finally, we isolate the magnetic charge strength in (2.15), as such:

$$\mu = \frac{1}{\varphi}\left(\hbar \frac{2\pi}{e} - \tau_\varphi \right). \tag{2.22}$$

Here, *the magnetic charge is always fractionalized*, but this fraction contains a quantized numerator $\hbar\left(2\pi / e\right)$ offset by the scalar field τ_φ.

Next, as we examine the $\mu \neq 0$ and $\tau_\varphi \neq 0$ solution (2.15), we are mindful of Dirac's statement in [2] that his DQC "shows, in fact, a symmetry between electricity and magnetism quite foreign to current views [but that it does not] force a complete symmetry," and that "if we insert the experimental value 137 in our theory, it introduces quantitative differences between electricity and magnetism so large that one can understand why their qualitative similarities have not been discovered experimentally up to the present." That is, "[t]he experimental result . . . shows that there must be some cause of dissimilarity between electricity and magnetism . . . as the result of which we have, not $\mu_0 = e$, but $\mu_0 = \left(137/2\right)e$."

With this in mind, we see that the complete solution (2.15) of the differential equation (2.2) <u>does</u> *break this duality symmetry, even at the theoretical level*. Specifically, if we take (2.15) and interchange $e \leftrightarrow \mu$, then this will become $\Lambda = e\mu\varphi + \mu\tau_\varphi = 2\pi n$ which *is not the same* as the original $\Lambda = e\mu\varphi + e\tau_\varphi = 2\pi n$. In fact, let us numerically compare (2.15) with an $e \leftrightarrow \mu$ interchanged (2.15), using the approximate numeric value $\alpha = e^2/4\pi \cong 1/137$ for the running electric coupling "constant" at low probe energies, i.e., at large spatial separation of a test charge from the bare charge being tested. For (2.15) itself:

$$2\pi\hbar = e\mu\varphi + e\tau_\varphi = e\mu\varphi + 2\sqrt{\pi/137}\,\tau_\varphi = e\mu\varphi + .033\cdot\tau_\varphi. \tag{2.23}$$

The final term $.033\cdot\tau_\varphi$ is a measure of the degree to which the $e \leftrightarrow \mu$ symmetry is broken, because were this to be zero, this symmetry would be restored. On the other hand, if we explicitly use the low energy $e \cong 2\sqrt{\pi/137}$ in (2.22) for the magnetic charge and so write $\mu = \left(\hbar\sqrt{137\pi} - \tau_\varphi\right)/\varphi$, and if we then use this after we interchange $e \leftrightarrow \mu$ in (2.15), the duality-interchanged version of (2.15) becomes:

$$2\pi\hbar = e\mu\varphi + \mu\tau_\varphi = e\mu\varphi + \frac{\hbar}{\varphi}\sqrt{137\pi}\,\tau_\varphi - \frac{\tau_\varphi{}^2}{\varphi} = e\mu\varphi + \frac{\hbar}{\varphi}20.746\tau_\varphi - \frac{\tau_\varphi{}^2}{\varphi}. \tag{2.24}$$

So we see that for the empirical low-impact $\alpha = e^2/4\pi \cong 1/137$, the latter equation (2.24) will break the duality symmetry to a greater degree than the former (2.23) so long as $\tau_\varphi < 20.746\hbar - .033\varphi$. For $\hbar = \varphi = 1$ this condition becomes $\tau_1 < 20.746 - .033 = 20.713$. While (2.15) does not tell us why this 137 number has the value it does – which would be the

"purely electronic quantum condition" Dirac was originally hoping to find with his derivation in [2] – a comparison of (2.23) with (2.24) does shed some light on why this numeric value ~1/137 is smaller and not larger: the expression $\Lambda = e\mu\cancel{\varphi} + e\tau_{\cancel{\varphi}} = 2\pi n$ will generally yield a higher degree of $e \leftrightarrow \mu$ symmetry than $\Lambda = e\mu\cancel{\varphi} + \mu\tau_{\cancel{\varphi}} = 2\pi n$, so long as *e* – whatever its experimental value – is generally a smaller number than μ. That is, the smaller the value of *e*, the closer we get to duality symmetry. The fact that $\alpha = e^2 / 4\pi \cong 1/137$ is smalls tell us that nature tends toward a duality symmetry, but that this symmetry is broken to the degree that 1/137 is not zero.

This now begs the next question: when happens to this solution when $e = \mu$? Dirac was certainly thinking about this question when he pointed out that the "dissimilarity between electricity and magnetism [causes us to] have, not $\mu_0 = e$, but $\mu_0 = (137/2)e$." But in 1931 we were a long way from understanding that these electric and magnetic charge strengths were "running" charges that were really not "constant," but were rather a function of a renormalization scale *Q* which reflected what would happen if we could probe more deeply into a charge, that is, if we could penetrate the polarized charge screen and get a "test" charge closer to the "bare" charge being tested. Nor was it close to being understood in 1931 that there might be GUT theories using Yang-Mills gauge groups that would not be invented for another two decades, and that a key hallmark of these theories would be that at ultra-high energies, all of the running couplings α merge into the same numerical value, so that their associated charges would also all become equal at least up to a Clebsch-Gordon coefficient. So in 2015, when we ask what happens when $e = \mu$, we know that we are implicitly asking what happens at GUT energies, at which there are also extraordinarily high temperatures associated with what many regard to be conditions in the very early universe. So, let's ask: what happens for $e = \mu$?

Because (2.15) is the most general solution to (2.2), we simply insert $e = \mu$ to obtain:

$$e^2\cancel{\varphi} + e\tau_{\cancel{\varphi}} = \mu^2\cancel{\varphi} + \mu\tau_{\cancel{\varphi}} = 2\pi\cancel{\mathrm{A}} \,. \tag{2.25}$$

This is quadratic in $e = \mu$, and it can be solved to show that:

$$2e = 2\mu = -\frac{\tau_{\cancel{\varphi}}}{\cancel{\varphi}} \pm \sqrt{\frac{\tau_{\cancel{\varphi}}^{\,2}}{\cancel{\varphi}^2} + 8\pi\frac{\cancel{\mathrm{A}}}{\cancel{\varphi}}} \,. \tag{2.26}$$

Then using this to obtain the running coupling $\alpha = e^2 / 4\pi = \mu^2 / 4\pi$, we obtain:

$$\alpha = \frac{e^2}{4\pi} = \frac{\mu^2}{4\pi} = \frac{\tau_{\cancel{\varphi}}^{\,2}}{8\pi\cancel{\varphi}^2} + \frac{\cancel{\mathrm{A}}}{2\cancel{\varphi}} \mp \frac{\tau_{\cancel{\varphi}}}{8\pi\cancel{\varphi}}\sqrt{\frac{\tau_{\cancel{\varphi}}^{\,2}}{\cancel{\varphi}^2} + 8\pi\frac{\cancel{\mathrm{A}}}{\cancel{\varphi}}} \,. \tag{2.27}$$

Making a mental note once again for when we need to "try to interpret the new mathematical features in terms of physical entities," we observe that because (2.25) through (2.27) rely upon setting $e = \mu$, these would only be valid under the physical conditions of a GUT, which is the

energy and temperature domain in which the stable 't Hooft and Polyakov [6], [7] magnetic monopoles are expected to become observable at ultra-high energies $>10^{15}$ GeV. We should also keep in mind, because the DQC corresponds with (some) TP monopole solutions asymptotically at large distance, and because any GUT discussion necessarily entails ultra-short distance impacts with the "bare" monopole, that the question arises whether the extended relationships such as (2.26) which arise as solutions to the same differential equation (2.2) might be shown to correspond to (some) of the *non-asymptotic* TP monopole solutions at closer range.

Finally, much of the development thus far has been predicated on holding the charge strengths e and μ constant, $de = d\mu = 0$. Let us now explore what needs to change in our analysis in the event one or both of these charges strengths is allowed to vary, such that $de \neq 0$ and / or $d\mu \neq 0$. Physically, we note that a charge strength which locally varies such that $\partial_\mu e \neq 0$ and / or $\partial_\mu \neq 0$ is a *running* charge strength whereby as one is able to move a test charge spatially closer to the bare charge being tested past some of the polarization screen of that bare charge, such that the effective charge strength is observed to vary as a function of this collision penetration. For example, based on renormalization group theory, $\alpha = e^2/4\pi$ grows logarithmically as the energy scale is increased. A empirical sampling of this is the observation that $\alpha \cong 1/137.036$ when the test charge is far from the observed charge with no penetration at all, $Q^2 = 0$, but that this grows to $\alpha \cong 1/128$ for a closer penetration of $Q^2 = {M_W}^2 \cong (80.385 \text{ GeV})^2$, see PDG's [18]. Indeed, the essential purpose of a collider is to get two particles to collide as strongly as possible, i.e., to penetrate one another as deeply as possible, and then characterize what is observed. And an important part of what is observed is that the charge strength then varies as a function of the collision depth. For the moment, we simply represent this local running of the charge strength via the very elemental mathematical statement that $de \neq 0$ and / or $d\mu \neq 0$ in spacetime.

If the charge strength is allowed to run (as it does in the real world), it first becomes mathematically important to carefully distinguish a *global* equation such as $\mu = \iiint dF = \oiint F$ from a *local* equation such as $F = (\mu/4\pi)d\cos\theta d\varphi$. The former represents a *total magnetic charge* μ enclosed within a *finite* volume which via Gauss / Stokes is equal to the *total net magnetic field flux* $\Phi_\mu = \mu$ across the *finite* surface of that same volume, so that μ is being measured *globally*. In the latter, F is a *local* field measured *infinitesimally* at each event in spacetime, so that μ represents a *local* measure of the charge strength at that same event. *It is only when we assume that* $de = d\mu = 0$, *that the* μ *appearing in the global* $\mu = \iiint dF = \oiint F$ *is identical to the* μ *appearing in the local* $F = (\mu/4\pi)d\cos\theta d\varphi$, *i.e., that* $\Phi_\mu = \mu$. If we allow μ to run, $d\mu \neq 0$, then this is no longer the case. When $d\mu \neq 0$ over the surface, $\Phi_\mu \neq \mu$. A detailed calculation explicitly showing how this occurs and how this introduces Dirac strings in a possible novel fashion is included in Appendix A.

As to the particular development here, let us return to (2.11), where we found that (2.10) is the solution to the Wu-Yang differential equation (2.2) *if and only if* the fields $\varepsilon(x^\mu) = d\tau(x^\mu)$

. At (2.12) we began focusing on the azimuth coordinate $\varphi \subset t,r,\varphi,\theta$, so using the notation introduced at (2.12) we rewrite (2.11) as $\varepsilon_\varphi = d\tau_\varphi$. In (2.15) we found the full solution to (2.2), which upon isolating τ_φ in (2.20) became $\tau_\varphi = (2\pi/e)\mathcal{A} - \mu\varphi$. So if we substitute (2.20) into $\varepsilon_\varphi = d\tau_\varphi$ in order to obtain an explicit expression for ε_φ, then given that $\mathcal{A}$ and φ are integer numbers, and also employing $\alpha = e^2/4\pi$ and the original definition $\varepsilon_\varphi \equiv A'_{N\varphi} - A_{S\varphi}$ written to explicitly highlight the azimuthal dependency, and also include (2.2), we now find that:

$$\varepsilon_\varphi = d\tau_\varphi = -\mathcal{A}\frac{2\pi}{e^2}de - \varphi d\mu = -\mathcal{A}\frac{1}{2\alpha}de - \varphi d\mu = A'_{N\varphi} - A_{S\varphi} = \frac{1}{ie}e^{-i\Lambda}de^{i\Lambda} - \frac{\mu}{2\pi}d\varphi. \qquad (2.28)$$

Because $\varepsilon_\varphi = d\tau_\varphi$ defined prior to (2.2) as $\varepsilon \equiv A'_N - A_S$ represents an observable distinctness between the north and south gauge field patches, we now see that *any such observable distinctness vanishes under one of two conditions*: First, when $de = d\mu = 0$, that is, when the charge strengths are constant and not running. Second, when charge strengths do run, but are constrained to one another according to the differential equation $\varepsilon = -\mathcal{A}(2\pi/e^2)de - \varphi d\mu = 0$. In (2.7) we found that whenever $\varepsilon = 0$, the solution will be $e\mu/2\pi = \mathcal{A}/\varphi$, which is the $\varphi = 1$ DQC generalized to all other azimuth windings of like orientation. Combining this with (2.28) tells us that:

$$\frac{\mathcal{A}}{\varphi} = -\frac{e^2}{2\pi}\frac{d\mu}{de} = \frac{e\mu}{2\pi}, \qquad (2.29)$$

which reduces from the latter two expressions to:

$$\frac{d\mu}{\mu} = -\frac{de}{e}. \qquad (2.30)$$

This differential equation is solved by all $e\mu = K$ with constant *K*, and so *includes* $e\mu = 2\pi\mathcal{A}/\varphi$ for the φ-generalized DQC of (2.7). So the second condition under which $\varepsilon_\varphi = 0$ is when the charge strengths are running, but are related by $e\mu = K$ which includes (2.7). In all other cases, there is an observable distinctness between the north and south patches, and as we see from (2.28), *it is the very fact that either* $de \neq 0$ *or* $d\mu \neq 0$ *or both are running in some relationship other than* $-\mathcal{A}(2\pi/e^2)de - \varphi d\mu = 0$, *which is responsible for any observable distinctness* $\varepsilon = A'_N - A_S \neq 0$ *between the north and south gauge patches.*

This is what is meant when we say that there is an observable distinctness $\varepsilon_\varphi = d\tau_\varphi \neq 0$ between the two gauge patches when $de \neq 0$ or $d\mu \neq 0$ and there is no *a priori* constraint between *de* and $d\mu$. When $-\mathcal{A}(2\pi/e^2)de - \varphi d\mu = 0$, this <u>is</u> an *a priori* constraint. When $-\mathcal{A}(2\pi/e^2)de - \varphi d\mu = \varepsilon_\varphi \neq 0$ there is <u>not</u> an *a priori* constraint, because ε_φ does nothing

more than quantify the independence of de from $d\mu$. It is also worth noting that if $\varepsilon_{\varphi}=0$, then because $e\mu/2\pi=\cancel{\text{A}}/\cancel{\varphi}$ in (2.7) is a consequence of $\varepsilon_{\varphi}=0$, and because $e\mu/2\pi=\cancel{\text{A}}/\cancel{\varphi}$ can be rewritten as $\cancel{\text{A}}2\pi/e=\cancel{\varphi}\mu=0$, a comparison with (2.20) tells us that in any situation where $\varepsilon_{\varphi}=0$ applies, we also have $\tau_{\varphi}=0$. In other words, the condition $\varepsilon_{\varphi}=d\tau_{\varphi}=0$ also implies the integrated condition $\tau_{\varphi}=0$. So $\tau_{\varphi}\neq 0$, like $\varepsilon_{\varphi}\neq 0$, is how we may generally characterize and quantify the absence of *a priori* constraint between *e* and μ. Further, because the fractional charges in (2.7) are a $\varepsilon_{\varphi}=0$ thus $\tau_{\varphi}=0$ solution to (2.2), and because the only fractional charges physically observed in nature reveal themselves near a temperature $T=0$, we keep in the back of our mind (for when we turn from mathematical solutions to physical interpretations) the possibility that ε_{φ} and τ_{φ} which measure the observable distinctness of the gauge patches and the absence of *a priori* constraint between the charge strengths might gain its observability by being temperature-dependent, such that $\varepsilon_{\varphi}(T=0)=\tau_{\varphi}(T=0)=0$ when the temperature approaches absolute zero, and $\varepsilon_{\varphi}(T>0)\neq 0$ and $\tau_{\varphi}(T>0)\neq 0$ otherwise.

Therefore, we may combine the result in (2.28) together with (2.2) and the original definition $\varepsilon_{\varphi}\equiv A'_{N\varphi}-A_{S\varphi}$, and explicitly rewrite the most general form of the Wu-Yang differential equation (2.2), which is a restructured (2.28), as:

$$\frac{1}{ie}e^{-i\Lambda}de^{i\Lambda}=\frac{\mu}{2\pi}d\varphi+\varepsilon_{\varphi}=\frac{\mu}{2\pi}d\varphi-\cancel{\text{A}}\frac{2\pi}{e^2}de-\cancel{\varphi}d\mu=\frac{\mu}{2\pi}d\varphi+\varepsilon_{\varphi}=\frac{\mu}{2\pi}d\varphi+A'_{N\varphi}-A_{S\varphi}. \quad (2.31)$$

We see that the commonly-used $e^{-i\Lambda}de^{i\Lambda}/ie=(\mu/2\pi)d\varphi$ with no observable gauge patch distinctness is a special case where $de=d\mu=0$ or where $-\cancel{\text{A}}\left(2\pi/e^2\right)de-\cancel{\varphi}d\mu=0$. As Dirac points out at 67 of [2], because "the wave function is complex, its vanishing will require two conditions, so that in general the points at which it vanishes will lie along a line," which "nodal line" in x^{μ} along which $\exp\left(i\Lambda\left(x^{\mu}\right)\right)\psi=\left(\cos\Lambda\left(x^{\mu}\right)+i\sin\Lambda\left(x^{\mu}\right)\right)(A+iB)=0$ subsequently became known as the "Dirac String." In (2.31), it is the presence of the transition function containing $e^{-i\Lambda}de^{i\Lambda}/ie=d\Lambda/e$ which ensures the non-observability of this Dirac string.

Now, we come to one other important feature of the mathematical solutions to the Wu-Yang differential equation (2.2), namely, energy quantization. We start with equation (2.28) for the observable distinctness between two gauge patches, which becomes non-zero due to the running of the charge strengths *e* and / or μ. If we restore the fundamental constants $\hbar$ and *c* to dimension everything explicitly in terms of energy, and then extract the four-vectors from within the closed, exact differential forms, (2.28) yields:

$$\varepsilon_{\varphi\mu}=d\tau_{\varphi\mu}=\cancel{\text{A}}\left(-\frac{2\pi(\hbar c)^{1.5}}{e^2}\partial_{\mu}e\right)+\cancel{\varphi}\left(-(\hbar c)^{1.5}\partial_{\mu}\mu\right)=\cancel{\text{A}}h\nu_{\cancel{\text{A}}\mu}+\cancel{\varphi}h\nu_{\varphi\mu}=E_{\cancel{\text{A}}\mu}+E_{\varphi\mu}. \quad (2.32)$$

Above, we define a pair of energy vectors:

$$\begin{cases} E_{\not{A}\mu} = \not{A}h\nu_{\not{A}\mu} = nh\nu_{\not{A}\mu} \equiv \not{A}\left(-\dfrac{2\pi(\hbar c)^{1.5}}{e^2}\partial_\mu e \right) \\ E_{\not{\varphi}\mu} = \not{\varphi}h\nu_{\not{\varphi}\mu} = mh\nu_{\not{\varphi}\mu} \equiv \not{\varphi}\left(-(\hbar c)^{1.5}\partial_\mu \mu \right) \end{cases} \tag{2.33}$$

which relate to frequency / wave vectors $h\nu^\mu = hck^\mu = h(\nu, c/\boldsymbol{\lambda}) = h(\nu, c\mathbf{k})$ in the usual way. What we uncover in (2.32) is the *quantization of energy* in the form of Planck's $E = nh\nu$ and deBroglie's $\mathbf{p} = nh/\boldsymbol{\lambda}$.

Per (2.32), when there are both running electric and magnetic charge strengths $\partial_\mu e$ and $\partial_\mu \mu$, we have a superposition $\not{A}h\nu_{\not{A}\mu} + \not{\varphi}h\nu_{\not{\varphi}\mu}$ of two quantized energy harmonics, one being $E_{\not{A}\mu} = \not{A}h\nu_{\not{A}\mu}$ with energy packets $h\nu_{\not{A}\mu} \equiv -2\pi(\hbar c)^{1.5}\partial_\mu e/e$ arising from the $\partial_\mu e$ running and the quantization $\not{A}h\nu_{\not{A}\mu}$ corresponding to the harmonic series of these packets, and the other being $E_{\not{\varphi}\mu} = \not{\varphi}h\nu_{\not{\varphi}\mu}$ with energy packets $h\nu_{\not{\varphi}\mu} = -(\hbar c)^{1.5}\partial_\mu \mu$ arising from the $\partial_\mu \mu$ running and $\not{\varphi}h\nu_{\not{\varphi}\mu}$ representing the harmonic series of these packets. This superposed form of energy quantization is topologically established by both the gauge winding number $\not{A} = n$ and the azimuth winding number $\not{\varphi} = m$. In the specialization of (2.19) where the magnetic charge strength $\mu = 0$ is set to zero and where we also set $d\mu = 0$, (2.32) reduces to:

$$\varepsilon_{\not{\varphi}\mu} = d\tau_{\not{\varphi}\mu} = \not{A}\left(-\frac{2\pi(\hbar c)^{1.5}}{e^2}\partial_\mu e \right) = \not{A}h\nu_{\not{A}\mu} = E_{\not{A}\mu}, \tag{2.34}$$

and the energy quantization is established by the gauge winding number $\not{A} = n$ alone.

What is fascinating about this is what appears to be a transmutation of $\not{A} = n$ from an electric *charge* quantum number in (2.7) and (2.17) and (2.21), into an *energy* quantum number in (2.32) and (2.34); and a similar transmutation of $\not{\varphi} = m$ from a magnetic charge *fractionalization* quantum number in (2.7) and (2.22), into an *energy* quantum number for a second harmonic in (2.32) when the monopole charge strength is not zeroed out.

3. Summary and Conclusion

The complete set of mathematical solutions to the Wu-Yang differential equations (2.1) and (2.2), in Dirac's words, enables us to "perfect and generalise the mathematical formalism that forms the existing basis of theoretical physics" for Dirac monopoles which are the asymptotic solution for many important variants of the 't Hooft-Polyakov monopoles. Let us now summarize these results.

What we have found is that there are three sets of solutions to the Wu-Yang differential equation (2.1) generalized to observable differences $\varepsilon = A'_N - A_S$ between the gauge patches, which differences we have shown are synonymous with local circumstances under which $d\mu \neq 0$ and / or $de \neq 0$ and $-\cancel{\text{A}}\left(2\pi / e^2\right)de - \cancel{\varphi} d\mu \neq 0$. The first set has $\mu \neq 0$ and $\varepsilon = 0$; the second set has $\mu = 0$ and $\varepsilon \neq 0$; and the third has both $\mu \neq 0$ and $\varepsilon \neq 0$.

This <u>first set of solutions</u> are for $\mu \neq 0$ and $\varepsilon = 0$, the latter of which we now know means also that $d\mu = de = 0$, <u>*or*</u> that the charge strengths are running but are constrained by $-\cancel{\text{A}}\left(2\pi / e^2\right)de - \cancel{\varphi} d\mu = 0$. These solutions are generally given by $\Lambda = e\mu\cancel{\varphi} = 2\pi n$ in (2.5) which yields the quantized, fractionalized charges $e = \left(\cancel{\text{A}} / \cancel{\varphi}\right)\left(2\pi / \mu\right) = \left(\cancel{\text{A}} / \cancel{\varphi}\right)e_{\text{u}}$ of (2.7). In the special case where $\cancel{\varphi} = 1$ this recovers the DQC, and in the further special case where $n = \cancel{\text{A}} = 1$ this describes a unit electron charge. Among the things we learn from this is that the DQC proper is *a specific solution to (2.2)* when $\varepsilon = 0$, and when $\cancel{\varphi} = 1$. If we restrict consideration to azimuths which have the same orientation and entanglement a.k.a. version as $\cancel{\varphi} = 1$, then the permitted fractional denominators $\cancel{\varphi} = 2\left(l + \frac{1}{2}\right) = 2\left(l + s\right) = 2j = 1, 3, 5, 7...$ become odd integers only, which a) happen to be what we observe in the FQHE with the exception of $\cancel{\varphi} = 2$, and b) also happen to have the exact same numerical pattern as the Casimir numbers obtained in the operations $\mathbf{L}^2\left|\xi\right\rangle = l\left(l+1\right)\left|\xi\right\rangle$, $\mathbf{S}^2\left|\xi\right\rangle = s\left(s+1\right)\left|\xi\right\rangle$ and $\mathbf{J}^2\left|\xi\right\rangle = j\left(j+1\right)\left|\xi\right\rangle$ for orbital, spin and total angular momentum of electrons in atomic shells.

The <u>second set of solutions</u> occur when $\mu = 0$ (and $d\mu = 0$) and $\varepsilon = A'_N - A_S \neq 0$, the latter of which we now know means that $de \neq 0$, which is what yields an observable difference between the gauge patches. These solutions are generally given by $\Lambda = e\tau_{\cancel{\varphi}} = 2\pi n$ in (2.16), which in turn yields electric charge quantization *without fractionalization*, as in $e = \cancel{\text{A}}\left(2\pi / \tau_{\cancel{\varphi}}\right) = \cancel{\text{A}}e_{\text{u}}$ of (2.18). The related expression $\tau_{\cancel{\varphi}} = \cancel{\text{A}}\left(2\pi / e\right)$, taken together with $\varepsilon_{\cancel{\varphi}} = d\tau_{\cancel{\varphi}}$ as first found in (2.11), leads in (2.34) to the finding that $\varepsilon_{\cancel{\varphi}\mu} = \cancel{\text{A}}h\nu_{\cancel{\text{A}}\mu}$, which means that the energy-dimensioned four-vector $\varepsilon_{\cancel{\varphi}\mu}$ is quantized in units of the energy packet $h\nu_{\cancel{\text{A}}\mu}$, with $\cancel{\text{A}} = n$ also serving as the quantum number for this energy quantization. That is, $\cancel{\text{A}} = n$ which describes charge quantization in $e = \cancel{\text{A}}e_{\text{u}}$, transmutes into also describing energy quantization in $\varepsilon_{\cancel{\varphi}\mu} = \cancel{\text{A}}h\nu_{\cancel{\text{A}}\mu}$.

The final, <u>third set of solutions</u> occur when $\mu \neq 0$ <u>*and*</u> when $\varepsilon = A'_N - A_S \neq 0$, so that $d\mu \neq 0$ and $de \neq 0$ and $-\cancel{\text{A}}\left(2\pi / e^2\right)de - \cancel{\varphi} d\mu \neq 0$. The general solution in this circumstance is $e\mu\cancel{\varphi} + e\tau_{\cancel{\varphi}} = 2\pi\cancel{\text{A}}$ in (2.15), which is rewritten as $\tau_{\cancel{\varphi}} = \cancel{\text{A}}\left(2\pi / e\right) - \cancel{\varphi}\mu$ in (2.20), and via $\varepsilon_{\cancel{\varphi}\mu} = d\tau_{\cancel{\varphi}\mu}$ found in (2.11), leads us to (2.32). Here too, the energy is quantized, but in a superposition of quantum states $\cancel{\text{A}}h\nu_{\cancel{\text{A}}\mu} + \cancel{\varphi}h\nu_{\cancel{\varphi}\mu}$. The gauge angle winding number $\cancel{\text{A}} = n$ contributes an energy quantization based on $\partial_\mu e \neq 0$, and the azimuth angle winding number

$\not{\varphi} = m$ of the azimuth further superimposes an energy quantization based on the $\partial_\mu \mu \neq 0$. As in the second solution set, $\not{\Lambda} = n$ appears to describe electric charge quantization in (2.21) as well as the energy quantization $\varepsilon_{\not{\varphi}\mu} = \not{\Lambda} h\nu_{\not{\Lambda}\mu}$ in (2.34). Further, the azimuthal $\not{\varphi} = m$ appears via its role in (2.22) to describe magnetic charge fractionalization, and via $\varepsilon_{\not{\varphi}\mu} = \not{\Lambda} h\nu_{\not{\Lambda}\mu} + \not{\varphi} h\nu_{\not{\varphi}\mu}$ in (2.32), one of the two superimposed harmonic series of energy quantization. In the mathematical circumstance where $e = \mu$ (which does *not* mean that $de = d\mu$), the charge strengths are related to $\tau_{\not{\varphi}}$ and to the winding quanta $\not{\Lambda}$ and $\not{\varphi}$ by (2.26) and the electromagnetic coupling strength is related to these by (2.27). We again note that physically, a relationship such as $e = \mu$ would only be observed under GUT conditions. Because the Dirac monopoles describe the long-distance asymptotic behavior of some TP monopoles, it may be worthwhile to see if these relationships can be shown to correspond to (some) of the *non-asymptotic* TP monopole solutions at the closer range to be expected in any GUT.

Perhaps the most important feature of all three of these sets of solutions, is that the Wu-Yang differential equation $e^{-i\Lambda} de^{i\Lambda} / ie = (\mu / 2\pi) d\varphi + \varepsilon$ of (2.2) – although it contains e, μ, Λ and φ all of which have smooth continuous values in spacetime – ends up yielding solutions in which Λ and φ become naturally quantized such that the reduced $\not{\Lambda} = n$ and $\not{\varphi} = m$ are topological quantum numbers respectively representing the number of gauge space windings and azimuth windings. And it is also clear that depending on circumstance, these two topological quantum numbers participate in solutions which exhibit charge quantization, charge fractionalization, and energy quantization.

The task from here is to embark on the second half of Dirac's program and "try to interpret the new mathematical features in terms of physical entities." It should be very clear that in so doing, a primary focus will be on whether the types of topological quantization which appear in these solutions bear a clear mapping to the types of quantization observed in the natural world. The most striking correspondence that needs to be studied is that of the fractionalized charges of (2.9) for electron states that have the same orientation and entanglement as a unit unfractionalized electron charge.

Although the FQHE fractionalization has been approached to date in terms of the collective behaviors of electrons and "quasi-particles" in conductive materials near 0K [19], it is difficult to dismiss the thought that the charge fractions 1,2,3,5,7,9… associated with this effect evidence anything less than another quantum number that exists in the natural world. If this is so, then either this is a new quantum number never before seen until the FQHE was discovered in which case nature has been a bit more extravagant than may have been anticipated, or it is a quantum number that is already known from elsewhere and simply takes on a new manifestation in the FQHE in which case nature has been economical. The fact that the odd-integer fractions line up perfectly via $\not{\varphi} = 2\left(l + \frac{1}{2}\right) = 2(l + s) = 2j = 1,3,5,7...$ with the Casimir numbers of electrons in atomic shells at the least requires us to study an interpretation in which nature has indeed economized. If this is the case, then experimental validation of these results may be obtainable by a close study correlating each fractional denominator of the FQHE to its observed angular momentum states.

Appendix A: Magnetic Surface Flux Calculation with a Varying Magnetic Charge Strength

The underlying calculation which best illustrates connection between local charge strength μ and a total charge / magnetic flux also found to be μ is:

$$\iiint dF = \oiint F = \oiint \frac{\mu}{4\pi} d\cos\theta d\varphi = \frac{\mu}{4\pi}\int_0^{\pi} d\cos\theta \int_0^{2\pi} d\varphi = \frac{\mu}{4\pi}\cos\theta\Big|_0^{\pi}\varphi\Big|_0^{2\pi} = \mu . \tag{A1}$$

The implied supposition above is that μ is constant, which is highlighted by the fact that $\mu/4\pi$ is moved *outside the integral* following the third equal sign in the above. But if we now allow μ to be a *running* function $\mu(\varphi,\theta)$ of φ and / or θ, then we are no longer allowed to move $\mu/4\pi$ outside the integral, but must keep this term inside with the rest of the integrand. Let us now examine how this works, so we can tie all of the mathematics together consistently for the situation where the magnetic charge does vary, $d\mu \neq 0$.

We start with $A_N = (\mu/4\pi)(\cos\theta - 1)d\varphi$ and $A_S = (\mu/4\pi)(\cos\theta + 1)d\varphi$ for the north and south gauge field patches, which as always differ by $A_S = A_N + (\mu/2\pi)d\varphi$. But we no longer limit ourselves to $d\mu = 0$, but rather, allow $d\mu \neq 0$. The field strength $F = dA$ as always, and also, dd=0 as always. But now, we obtain non-zero terms with $d\mu$, which leads to two *seemingly-different* results, for each of a north and south field strength, namely:

$$\begin{cases} F_N = dA_N = \dfrac{\mu}{4\pi} d\cos\theta d\varphi + \dfrac{d\mu}{4\pi}\cos\theta d\varphi - \dfrac{d\mu}{4\pi} d\varphi \\ F_S = dA_S = \dfrac{\mu}{4\pi} d\cos\theta d\varphi + \dfrac{d\mu}{4\pi}\cos\theta d\varphi + \dfrac{d\mu}{4\pi} d\varphi \end{cases} . \tag{A2}$$

This *apparent* difference arises because of the sign difference in the final terms with $\pm(d\mu/4\pi)d\varphi$. But let us see how this develops, mathematically, when we use each of these to extend the calculation (A1) to $d\mu \neq 0$. As we shall see, this apparent difference is an unobservable artifact related to the unobservable Dirac strings.

The mathematics is simplified if we first set $\mu \equiv \mu_0 + \tilde{\mu}$ in each of the local equations $A_N = (\mu/4\pi)(\cos\theta - 1)d\varphi$ and $A_S = (\mu/4\pi)(\cos\theta + 1)d\varphi$, where μ_0 is defined to be an arbitrary constant, $d\mu_0 = 0$, and $\tilde{\mu}$ contains the "spread" (hence the overhead "$\leftrightarrow$") about this constant μ_0, such that $d\tilde{\mu} \neq 0$. So in effect we segregate all running of the charge strength into $\tilde{\mu}$. Using this definition, we can rewrite (A2) as:

$$\begin{cases} F_N = dA_N = \dfrac{\mu_0}{4\pi}\cos\theta d\varphi + \dfrac{\vec{\mu}}{4\pi} d\cos\theta d\varphi + \dfrac{d\vec{\mu}}{4\pi}\cos\theta d\varphi - \dfrac{d\vec{\mu}}{4\pi} d\varphi \\ F_S = dA_S = \dfrac{\mu_0}{4\pi}\cos\theta d\varphi + \dfrac{\vec{\mu}}{4\pi} d\cos\theta d\varphi + \dfrac{d\vec{\mu}}{4\pi}\cos\theta d\varphi + \dfrac{d\vec{\mu}}{4\pi} d\varphi \end{cases} . \tag{A3}$$

These are still local field equations, but let's take their integrals $\oiiint dF = \oint\!\!\oint F$ and see what results globally.

First, since we are now allowing $\vec{\mu}(\varphi,\theta)$ to run with φ and θ, the total derivative:

$$d\vec{\mu}(\varphi,\theta) = \frac{\partial\vec{\mu}}{\partial\varphi} d\varphi + \frac{\partial\vec{\mu}}{\partial\theta} d\theta . \tag{A4}$$

Now we follow (A1), and also use (A4), as well as the fact that $d\varphi d\varphi = 0$ because in differential form geometry $dx^\mu dx^\nu = -dx^\nu dx^\mu$. The only difference between F_N and F_S in (A3) is in the final term $\mp(d\vec{\mu}/4\pi)d\varphi$, so rather than doing the same calculation twice with just a sign variation in the last term, we simply use F_S^N to represent the calculation with a respective $\mp$ sign for the final term. Thus, we may calculate:

$$\begin{aligned} \Phi_{\mu S}^{\ N} &= \iiint dF_S^N = \oint\!\!\oint F_S^N = \oint\!\!\oint \frac{\mu_0}{4\pi}\cos\theta d\varphi + \oint\!\!\oint \frac{\vec{\mu}}{4\pi} d\cos\theta d\varphi + \oint\!\!\oint \frac{d\vec{\mu}}{4\pi}\cos\theta d\varphi \mp \oint\!\!\oint \frac{d\vec{\mu}}{4\pi} d\varphi \\ &= \mu_0 + \oint\!\!\oint \frac{\vec{\mu}}{4\pi} d\cos\theta d\varphi + \frac{1}{4\pi}\int_0^\pi \frac{\partial\vec{\mu}}{\partial\theta} d\theta \cos\theta \int_0^{2\pi} d\varphi \mp \frac{1}{4\pi}\int_0^\pi \frac{\partial\vec{\mu}}{\partial\theta} d\theta \int_0^{2\pi} d\varphi \\ &= \mu_0 + \oint\!\!\oint \frac{\vec{\mu}}{4\pi} d\cos\theta d\varphi + \frac{1}{2}\int_{\theta=0}^{\pi} d\vec{\mu}\cos\theta \mp \frac{1}{2}\int_{\theta=0}^{\pi} d\vec{\mu} \\ &= \mu_0 + \oint\!\!\oint \frac{\vec{\mu}}{4\pi} d\cos\theta d\varphi + \frac{1}{2}\left(\vec{\mu}(\theta)\cos\theta\right)\Big|_{\theta=0}^{\pi} \mp \frac{1}{2}\vec{\mu}(\theta)\Big|_{\theta=0}^{\pi} \\ &= \mu_0 + \oint\!\!\oint \frac{\vec{\mu}}{4\pi} d\cos\theta d\varphi - \begin{cases} \vec{\mu}(\theta=\pi) \\ \vec{\mu}(\theta=0) \end{cases} \end{aligned} \tag{A5}$$

To simplify this result further, we exploit the fact that μ_0 in $\mu \equiv \mu_0 + \vec{\mu}$ is a completely arbitrary constant used merely to help us do the integrals in (A5). What is important is not the value of μ_0, *but the fact that this is a constant*. So let us now select this arbitrary constant to be $\mu_0 = 0$ so that $\mu = \vec{\mu}$. Then (A5), separated again into north and south, will simplify to:

$$\begin{cases} \Phi_{\mu N} = \iiint dF_N = \oint\!\!\oint F_N = \oint\!\!\oint \dfrac{\mu}{4\pi} d\cos\theta d\varphi - \mu(\theta=\pi) \\ \Phi_{\mu S} = \iiint dF_N = \oint\!\!\oint F_N = \oint\!\!\oint \dfrac{\mu}{4\pi} d\cos\theta d\varphi - \mu(\theta=0) \end{cases} . \tag{A6}$$

Superficially, it looks like the total fluxes are different because the former $\Phi_{\mu N}$ contains $\mu(\theta=\pi)$ while the latter $\Phi_{\mu S}$ contains $\mu(\theta=0)$. But $\theta=\pi$ defines a semi-infinite line from the origin through the south pole of the closed surface, while $\theta=0$ defines a second semi-infinite line from the origin through the north pole. So the north-defined flux $\Phi_{\mu N}$ contains $\mu(\theta=\pi)$ which is taken at the only place over A_N which is undefined, namely the south pole, while the south-defined flux $\Phi_{\mu S}$ contains $\mu(\theta=0)$ which is likewise undefined because it is taken at the north pole. These are precisely the Dirac strings, which are unphysical and unobservable. These extra terms are unphysical artifacts of our coordinate system and are not observable, so they can be removed from (A6) with the result that the fluxes are the same irrespective of whether they are obtained using the $F=dA_N$ or $F=dA_S$. It is of interest that Dirac strings make an appearance in this way, when one attempts to relate the local charge strength to the global field flux for running magnetic charges with $d\mu\neq 0$.

Consequently, we may now consolidate both of (A6) into a single equation:

$$\Phi_\mu=\iiint dF=\oiint F=\oiint\frac{\mu}{4\pi}d\cos\theta d\varphi \tag{A7}$$

for the total flux across the close surface of monopole, in the situation where μ in the gauge patches is allowed to run, $d\mu\neq 0$. Of course $(4\pi)^{-1}\oiint\mu\, d\cos\theta d\varphi$ is not expressly calculated, because this depends on the specific variable character of $\mu(\varphi,\theta)$. But the form of this equation is exactly the same whether $d\mu=0$ or $d\mu\neq 0$. The difference is that when $d\mu=0$, we also have $\Phi_\mu=\mu$ even though Φ_μ is globally defined and μ is locally defined. But if $d\mu\neq 0$, then $\Phi_\mu\neq\mu$, and the global Φ_μ is not the same as the local μ.

References

[1] Misner, C. W., Thorne, K. S., and Wheeler, J. A., *Gravitation*, Freeman (1973)
[2] Dirac, P.A.M., *Quantized Singularities in the Electromagnetic Field*, Proceedings of the Royal Society A 133 (821): 60–72, (September 1931)
[3] Thomson, J. J., *Cathode Rays*, The Electrician 39, 104 (1897)
[4] Millikan, R. A., *On the Elementary Electric charge and the Avogadro Constant*, Physical Review, series II, **2**, pp. 109-143 (1913)
[5] Witten, E., *Duality, Spacetime and Quantum Mechanics*, Physics Today, 28 to 33 (1997)
[6] 't Hooft, G. (1974). "Magnetic monopoles in unified gauge theories". *Nuclear Physics B* **79** (2): 276–284
[7] A.M. Polyakov, Zh. Eksp. Teor. Fiz. Pis'ma. Red. 20, 430, JETP Lett. 20, 194 (1974)
[8] M. Georgi and S.L. Glashow, *Unified Weak and Electromagnetic Interactions without Neutral Currents*, Phys. Rev. Lett. 28 1494 (1972)
[9] Weinberg, S., *The Quantum Theory of Fields, Volume II, Modern Applications*, Cambridge (1996)
[10] H. Weyl, *Gravitation and Electricity*, Sitzungsber. Preuss. Akad.Wiss., 465-480. (1918).
[11] H. Weyl, Electron und Gravitation, Zeit. f. Physik, 56 (1929), 330

[12] Yang, C. N.; Mills, R. (1954). *Conservation of Isotopic Spin and Isotopic Gauge Invariance*. *Physical Review* 96 (1), 191–195 (1954)
[13] W. Heisenberg, "Quantum-mechanical re-interpretation of kinematic and mechanical relations," Z. Phys. 33, 879-893 (1925)
[14] A. Einstein, *The Foundation of the General Theory of Relativity*, Annalen der Physik (ser. 4), 49, 769–822 (1916)
[15] Wu, T. T. and Yang, C.N., *Concept of non-integrable phase factors and global formulation of gauge fields*, Phys. Rev. D 12 (1975) 3845
[16] Wu, T. T. and Yang, C. N., *Dirac Monopole without Strings: Classical Lagrangian theory*, Phys. Rev. D 14 (1976) 437
[17] Zee, A., *Quantum Field Theory in a Nutshell*, Princeton (2003)
[18] http://pdg.lbl.gov/2014/reviews/rpp2014-rev-phys-constants.pdf
[19] Laughlin, R. B., *Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations*, *Physical Review Letters* 50 18 (1983)