Single-Valued Simply-Connected Covering Groups Permitting Well-Defined Dirac-Wu-Yang Monopoles with Fractional Charges

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

April 2, 2015

Abstract: Although the Wu-Yang derivation of the Dirac Quantization condition (DQC) leads mathematically to fractional charge solutions, a careful study of these fractional solutions using Dirac strings on a closed surface in SO(3) shows precisely why these fractional charges cannot occur without giving rise to observable singularities which of course are not permitted, and why only the standard DQC is permitted. However, SO(3) is multivalued and so is not an exact representation of the operative symmetries. When we carefully analyze simply-connected, single-valued covering groups for which the generators are the generalized m^th roots of the 2x2 identity matrix I, which covering groups do exactly represent the operative symmetries, we find that there is no such restriction and well-defined fractional charges are topologically permitted without ambiguity.

PACS: 11.15.-q; 14.80.Hv

Contents

1. Introduction

In 1931 Dirac [1] discovered that *if* magnetic charges μ were to hypothetically exist, *then* this would imply that the electric charge *e* must be quantized. The relationship he found, often written as $e\mu = 2\pi n$ where *n* is a positive or negative integer or zero, came to be known as the Dirac Quantization Condition (DQC). Dirac's derivation employed what he called a "nodal line," in modern language referred to as a "Dirac string," which Dirac introduced as "an exceptional case . . . occurring when the wave function vanishes, since then its phase does not have a meaning." This string is often visualized as a semi-infinite solenoid of singularly-thin width which shunts magnetic field lines to and from spatial infinity, and it is not and cannot be physical observable.

But as stated rather bluntly at 447 of [2]:

> "The Dirac string is a considerable embarrassment in monopole theory. It is disconcerting to find that the vector potential that describes a Dirac monopole has a string singularity along which the magnetic field is formally infinite, even though we can argue that the string is undetectable. One is therefore encouraged to discover that it is possible to eliminate the string."

To remedy the need to resort to the fiction of such strings, Wu and Yang in the mid-1970s [3], [4] developed an approach which does not at all make use of these strings. Its results are completely equivalent to Dirac's, with the only difference being that it is cast in the more-modern language of fiber bundles. In the Wu Yang approach, one uses gauge theory and particularly north and south gauge field patches to obtain the differential equation $e^{-i\Lambda} de^{i\Lambda} = (ie\mu / 2\pi) d\varphi$ where Λ is the gauge (phase) angle and φ is the geometric azimuth about the z-axis in the three dimensional physical space of the rotation group SO(3). If we define a "reduced azimuth" $\not\varphi \equiv \varphi / 2\pi$ merely for notational convenience, it is readily seen that this equation is solved by $\exp(i\Lambda) = \exp(ie\mu\not\varphi)$.

Moreover, as Dirac observes and as is well known, "the value of [the phase] at a particular point has no physical meaning and only the difference between the values of [the phase] at two different points is of any importance." So once we start to consider specific azimuth angles, first $\not\varphi = 0$, we may use $\exp(i\Lambda) = 1 = \exp(i2\pi n)$ to deduce that $\Lambda = 2\pi n$ with integer *n* will be quantized in units of 2π. If we employ a reduced gauge angle $\not\Lambda \equiv \Lambda / 2\pi = n$, then this is a quantum number representing the number of "windings" through the gauge space. Then, using $\not\varphi = 1$ and requiring that $\exp(ie\mu\not\varphi) = 1$ because $\varphi = 2\pi$ has the same orientation as $\varphi = 0$, we obtain $\Lambda = e\mu = 2\pi n$. Finally, however, if we wind one or more additional times around the z-axis, then the generalized relationship $\exp(ie\mu\not\varphi) = 1 = \exp(i2\pi n)$ yields $e\mu = 2\pi(n / \not\varphi)$, which is suggestive of fractionalized charges as well.

Now, making the Dirac string unobservable is equivalent to the requirement that the fiber bundle be well-defined. So the question is whether there are fiber bundles which can be well-defined even for $\not\varphi > 1$, and specifically whether the mathematically-permitted fractional charge solutions $\not\varphi = 2,3,4,5...$ in $e\mu = 2\pi\left(n/\not\varphi\right)$ can be given an unambiguous topological meaning, or whether the only well-defined solution is the $\not\varphi = 1$ solution $e\mu = 2\pi\not\Lambda$ of the standard Dirac condition. In other words, are we required to discard these mathematical solutions $e\mu = 2\pi\not\Lambda/\not\varphi$ with $\not\varphi = 2,3,4,5...$ of Wu and Yang's $\exp\left(i\Lambda\right) = \exp\left(ie\mu\not\varphi\right)$ which solves their differential equation $e^{-i\Lambda}de^{i\Lambda} = ie\mu d\not\varphi$ as being *ill-defined and unphysical*? Or, are there topological mapping under which these solutions are well defined and physically acceptable and do not lead to physically observable singularities?

Clearly, when we utilize the rotation group SO(3) to talk about magnetic flux in or out of a closed two-dimensional surface, and when we first wind an azimuth over $0 \le \varphi \le 2\pi$ and then begin additional cycles over which $\varphi > 2\pi$, we are covering SO(3) more than once. So the answer to these questions will have to emerge from a careful analysis of various covering groups $\tilde{G}$ which may project onto SO(3) with a many-to-one surjective homomorphism $\pi : \tilde{G} \to SO(3)$, which we shall refer to generally as *n*-tuple covering groups of the physical space of SO(3) rotations. As we shall see, although these fractional charges $e\mu = 2\pi\left(n/m\right)$ with $m = 2,3,4,5...$ do appear to be forbidden when analyzed using Dirac strings in SO(3) which is multivalued and not simply connected, these fractional charges do obtain a clear, well-defined, unambiguous topological meaning, when we consider single-valued, simply-connected covering groups which we shall denote as ${}_m\tilde{G}$ which project onto SO(3) via $\pi : {}_m\tilde{G} \to SO(3)$ for which the kernel $\ker\pi = \left\{\sqrt[m]{1}\right\}$ represents an m^{th} root of unity.

We begin by carefully reviewing magnetic fluxes in and out of closed SO(3) spatial surfaces using Dirac strings and Dirac's original approach from [1], but making use of the language and apparatus of differential forms. Then, we show how these same results may be topologically approached using $\ker\pi = \left\{\sqrt[m]{1}\right\}$ single-valued covering groups ${}_m\tilde{G}$ projecting onto SO(3) via $\pi : {}_m\tilde{G} \to SO(3)$, such that the fractional charge solutions $2\pi n/m = e\mu$ of the Wu-Yang equation $\exp\left(i\Lambda\right) = \exp\left(ie\mu\not\varphi\right)$ originating in $e^{-i\Lambda}de^{i\Lambda} = ie\mu d\not\varphi$ may indeed be topologically well-defined and unambiguous. Throughout this discussion we shall employ natural units in which $\hbar = c = 1$.

2. Differential Forms Review of Dirac's Original Monopole Derivation

To approach Dirac's original derivation in [1] using differential forms, we begin by considering a gauge transformation $eA \to eA' = eA + d\Lambda$ on the differential one-form gauge potential $A = A_\mu dx^\mu$, which we rewrite as $eA \to eA' = eA - d\Lambda$ to be consistent with the oppositely-signed convention implicitly employed by Dirac. Taking a derivative we obtain $edA \to edA' = edA - dd\Lambda$ and via $F = dA$ this becomes $eF' = eF - dd\Lambda$. In this expression, $dd\Lambda = 0$ is a closed form which is locally exact, but not need be globally exact. Applying Gauss' / Stokes' theorem enables us to write this as $e\iint F' = e\iint F - \oint d\Lambda$, or:

$$\oint d\Lambda = -e\iint F' + e\iint F \,. \tag{2.1}$$

In the foregoing, $d\Lambda$ is the non-integrable (locally not globally exact) wavefunction phase one-form which may have its key relationships summarized using the notation of Dirac's equations (3), (4) and (6) by the differential one-form $\kappa = d\beta = eA = d\Lambda$ with the electric and magnetic fields in turn related to this by $\nabla\times\boldsymbol{\kappa} = e\mathbf{H}$ and $\nabla\kappa_0 - \partial_t\boldsymbol{\kappa} = e\mathbf{E}$ as in Dirac's equation (7), or more simply consolidated, $eF = d\kappa$ which reiterates $F = dA$. When we consider the three-dimensional space only, $e\iint F$ corresponds with $e\iint \mathbf{H}\cdot\mathbf{dS}$ which appears in Dirac's equation (8) and its unnumbered antecedent on Dirac's page 67. Likewise, $\oint d\Lambda$ corresponds to what Dirac regularly refers to as a "change in phase round [small or otherwise] closed curves."

Now, at page 66, Dirac first notes "how the non-integrable derivatives $\boldsymbol{\kappa}\left[=e\mathbf{A}\right]$ of the phase of the wave function receive a natural interpretation in terms of the potentials of the electromagnetic field" and that this "gives us nothing new." Thus, "[t]he condition for an *unambiguous physical interpretation* of the theory was that the change in phase round a closed curve should be the same for all wave functions." "There is, however, 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 . . . . Evidently," Dirac concludes, "these conditions must now be relaxed. The change in phase round a closed curve may be different for different wave functions by arbitrary multiples of 2π and is thus not sufficiently definite to be interpreted immediately in terms of the electromagnetic field."

"To examine this question," Dirac says, we "consider first a very small closed curve." Because "the wave equation requires the wave function to be continuous . . . the change in phase round a small closed curve must be small [and] cannot now be different by multiples of 2π for different wave functions. It must have one definite value and may therefore be interpreted without ambiguity in terms of the flux of the 6-vector **E**, **H** [here, bivector in the two form $F = F_{\mu\nu}dx^\mu dx^\nu$] through the small closed curve, which flux must also be small."

However, he continues, "there is an exceptional case . . . occurring when the wave function vanishes, since then its phase does not have a meaning. As 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." He called "such a line a *nodal line*," which in modern terminology is the Dirac string. "If," Dirac states, "we now take a wave function having a nodal line passing through our small closed curve, considerations of continuity will no longer enable us to infer that the change in phase round the small closed curve must be small. All we shall be able to say is that the change in phase will be close to $2\pi n$ where *n* is some integer, positive or negative. This integer will be a characteristic of the nodal line." Specifically, "[t]he difference between the change in phase round the small closed curve and the nearest $2\pi n$ must . . . be interpreted in terms of the flux of the 6-vector [i.e., the bivector in *F*] through the closed curve." In the language of differential forms, Dirac's foregoing statement is that $\oint d\Lambda - 2\pi n = e\iint F$, or rearranged:

$$\oint d\Lambda = 2\pi n + e\iint F \ . \tag{2.2}$$

This is simply a restatement using differential forms and with $\hbar = c = 1$ of the unnumbered equation in the middle of Dirac's page 67, where $2\pi n$ is unchanged, where $e/\hbar c.\int(\mathbf{H},\mathbf{dS})$ is represented with $e\iint F$, and where $\oint d\Lambda$ more formally represents the "change in phase [$d\Lambda$] round the small closed curve [$\oint$]." It will be seen that (2.2) above is a variant of $\oint d\Lambda = -e\iint F' + e\iint F$ in (2.1) rooted in the gauge transformation $eA \to eA' = eA - d\Lambda$, in which $-e\iint F'$ is replaced by $2\pi n$.

Then, as Dirac states, "[we] can now treat a large closed curve by dividing it up into a network of small closed curves lying in a surface whose boundary is the large closed curve. The total change in phase round the large closed curve will equal the sum of all the changes round the small closed curves and will therefore be":

$$\oint d\Lambda = 2\pi \Sigma n + e\iint F \ , \tag{2.3}$$

which restates Dirac's equation (8) in differential forms language. Here, $\oint d\Lambda$ is simply the change in phase round *any curve large or small*, built up from the "network of small closed curves." In general, for any particular "small closed curve" in this network, if there *is* a nodal line passing through the small closed curve, then (2.2) will apply. If there *is not* a nodal line passing through, then $n = 0$ for that particular small closed curve and so (2.2) for that curve simply becomes $\oint d\Lambda = e\iint F$. For such a non-nodal small closed curve, "considerations of continuity" tell us that "the change in phase round the small closed [$\oint d\Lambda$] curve must be small" commensurate with the small $e\iint F$ to which this is equal, and will approach zero as the small closed curve reduces in area to an infinitesimal point in the calculus sense of a very small Δx approaching an infinitesimal dx. Specifically, if we take a given small closed curve to be a circle circumference, then this curve will enclose a small finite surface area $\Delta A = \pi r^2$ which in the calculus sense approaches the infinitesimal area $dA = 2\pi r dr$ for a geometric point as $r \to 0$.

In (2.2) and (2.3) above, $\iint F$ applies to any surface, open or closed. Let us then imagine as illustrated in Figure 1 below that we have a spatial surface which is a two-dimensional sphere bounding a three-dimensional volume. We further imagine that starting from the north pole and working southward, we have built up a "network of small closed curves" with localized parts of the field strength F flowing therethrough via $\iint F$ such that this network covers the entire surface of the sphere, *except for* a very small but finite opening which encircles the south pole and is bounded by one final "small closed curve" needed to complete the coverage of the entire surface. We regard this final "small closed curve" (SCC) as a very small circle bounding a very small surface area $\Delta A = \pi r^2$ which passes a very small flux $\iint F_{\text{SCC}}$. We imagine that one or more of these small closed curves in the $\iint F$ network do have $n \neq 0$ nodal lines passing through them such that they are governed by (2.2). Thus, "[t]he total change in phase . . . will equal the sum of *all* the changes round the small closed curves" in the network with "the integration being taken over the surface and the summation over all nodal lines that pass through it, the proper sign being given to each term in the sum." This entire $\iint F$ network will therefore be governed by (2.3). And we also imagine that this final small closed curve may or may not have its own nodal line passing through. If it does not, then its $n = 0$ and its (2.2) is $\oint d\Lambda = e\iint F$. If it does, and if the "characteristic" of that nodal line is designated by $n_{\text{SCC}} \neq 0$, then its (2.2) is $\oint d\Lambda = 2\pi n_{\text{SCC}} + e\iint F$.

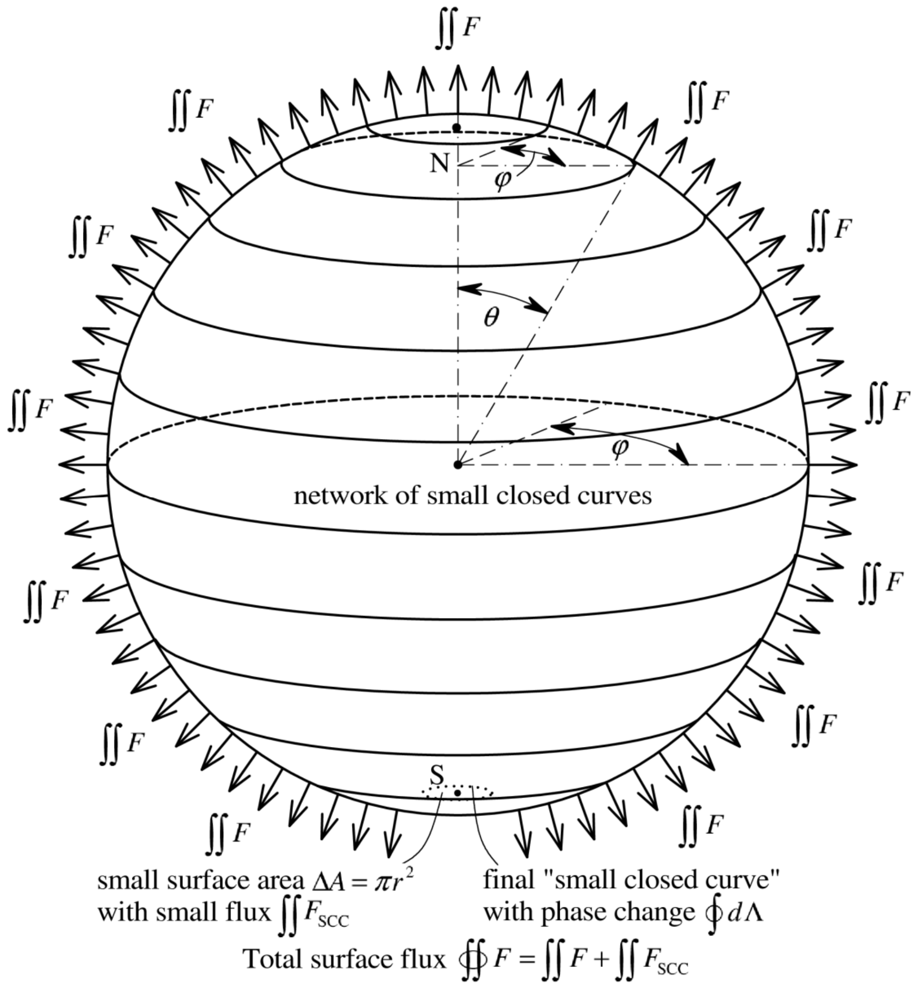

Figure 1: The Dirac Network of "Small Closed Curves"

As we build up this network from north to south there are times when the closed curve is a large closed curve (for example, near the equator), but by the time we approach the south pole the large closed curve has reduced down once again to a small closed curve. However, the total surface $\iint F$ to the north which this small closed curve now bounds with the characteristic $2\pi\Sigma n$ includes the entire sphere except for the very small area immediately about the south pole. In other words, this final small closed curve bounds and joins two surfaces: a very large surface to the north with larger flux $\iint F$, and a very small remaining surface to the south with very small balance of flux $\iint F_{\text{SCC}}$. (Indeed, any latitudinal line bounds and joins a north surface above and a south surface below.) The large surface north of the final small closed curve is *almost closed*, but remains open solely by virtue of this one last very small opening about the south pole. This means that $\iint F \cong \oiint F$ to a very close approximation, and that these differ only by the very small field flux $\iint F_{\text{SCC}}$ through the boundary of this small closed curve over the small remaining surface to the south. More precisely, we may write $\oiint F = \iint F + \iint F_{\text{SCC}}$ which patches the north and south surfaces together into the complete closed surface, where $\iint F_{\text{SCC}} \cong 0$ is very small. We may then use this notation to rewrite (2.3) as:

$$\oint d\Lambda = 2\pi\Sigma n + e\iint F = 2\pi\Sigma n + e\oiint F - e\iint F_{\text{SCC}} , \tag{2.4}$$

This represents in precise terms, exactly what is being depicted in Figure 1 above.

Now, in the calculus sense, let us take the limit in which we this final small closed curve (presumed to be a circle of radius *r*) becomes smaller and smaller such that finite area $\Delta A = \pi r^2$ bounded by this curve now becomes the infinitesimal area $dA = 2\pi r dr \to 0$ as $r \to 0$. As this area approaches zero, so too the flux $\iint F_{\text{SCC}} \to 0$ approaches zero, thus $\iint F \to \oiint F$. Likewise, the change in phase round the small closed curve $\oint d\Lambda \to 0$ as well, because the calculus has now changed the small circle into a single infinitesimal point. If the small closed curve contained a nodal line with $n_{\text{SCC}} \neq 0$, then $2\pi\Sigma n \to 2\pi(\Sigma n + n_{\text{SCC}})$ in (2.4). If not, we may still replace $2\pi\Sigma n \to 2\pi(\Sigma n + n_{\text{SCC}})$ in (2.4), simply keeping in mind that $n_{\text{SCC}} = 0$. Accounting for all of this, once these calculus limits are taken, (2.4) becomes:

$$\begin{aligned} 0 = \oint d\Lambda &= 2\pi(\Sigma n + n_{\text{SCC}}) + e\oiint F = 2\pi(\Sigma n + n_{\text{SCC}}) + e\oiint \frac{\mu}{4\pi} d\cos\theta d\varphi \\ &= 2\pi(\Sigma n + n_{\text{SCC}}) + e\mu \end{aligned} , \tag{2.5}$$

where we also now use $F = (\mu / 4\pi) d\cos\theta d\varphi$ for the field strength two-form represented in spherical coordinates. The final step in which $e\mu$ appears, occurs by evaluating the integral:

$$\oiint F = \oiint \frac{\mu}{4\pi} d\cos\theta d\varphi = \frac{\mu}{4\pi}\int_0^{\pi} d\cos\theta \int_0^{2\pi} d\varphi = \frac{\mu}{4\pi}\cos\theta\Big|_0^{\pi} \varphi\Big|_0^{2\pi} = \mu . \tag{2.6}$$

This integral is evaluated under the supposition that $d\mu = 0$ so that μ is constant-valued over the entire surface of integration, as can be seen by $\mu / 4\pi$ having been moved outside the integral after the second equal sign in (2.6). We make special note of the fact that this includes the integral $\int_0^{2\pi} d\varphi = \varphi\big|_0^{2\pi} = 2\pi$ taken over the range $0 \le \varphi \le 2\pi$. Were we to evaluate this over a larger domain, then we would start to cover this surface more than once.

Now, (2.5) readily restructures into:

$$e\mu = -2\pi\left(\Sigma n + n_{\mathrm{SCC}}\right). \tag{2.7}$$

But of course, $\Sigma n + n_{\mathrm{SCC}}$ is itself an integer which is positive or negative or zero. So if we simply rename this integer via $\Sigma n + n_{\mathrm{SCC}} \to -n$, then we obtain:

$$e\mu = 2\pi n, \tag{2.8}$$

which is the Dirac Quantization Condition.

Ordinarily, one stops right here at (2.8) and does not consider the possibility of covering the flux surface $\oiint F$ more than a single time. That is, ordinarily, one stops at the upper domain of $0 \le \varphi \le 2\pi$ in (2.6), recognizing that once we reach $\varphi = 2\pi$ we have returned to the same geometric orientation with which we started at $\varphi = 0$, and that for any periodic function such as $\exp i\varphi$ there is a one-to-many (really, one-to-infinite) discrete mapping from the range to the domain. For example, the range value 1 in $\exp i\varphi = 1$ maps into the infinite set of domain values $\varphi = 2\pi m$, or $\cancel{\varphi} = \varphi / 2\pi = m$, for all integers *m*.

However, as summarized in the introduction, the Wu-Yang differential equation $e^{-i\Lambda} de^{i\Lambda} = ie\mu d\cancel{\varphi}$ has the general solution $\exp(i\Lambda) = \exp(ie\mu\cancel{\varphi})$ which in turn is solved by $2\pi n = e\mu\cancel{\varphi}$ with $\cancel{\varphi} = m = 0,1,2,3\ldots$. This of course becomes the standard DQC $e\mu = 2\pi n$ in the specialization for which $\cancel{\varphi} = 1$, and corresponds to evaluating (2.6) over the domain $0 \le \varphi \le 2\pi$ of a single complete closed cover of the surface. Referring to (2.5) and (2.6), this means that the standard DQC (2.8) is really the solution for a *single closed covering* $\oiint_1 F$ of the flux surface, which we represent by the subscripted "1" next to the closed double integral. But the general set of Wu-Yang solutions $2\pi n = e\mu\cancel{\varphi}$ with $\cancel{\varphi} = m$ represents an *m*-tuple closed covering of the flux surface which we denote generally by $\oiint_i F$ with $1 \le i \le m$, and for which the domain of Figure 2 is now taken to be $0 \le \varphi \le 2\pi m$. The specific question we must then consider, is whether these multiple-covering solutions are allowable, which is to say, topologically speaking, whether there are covering groups with a many-to-one surjective homomorphism onto the SO(3) sphere of Figure 1 which permit a well-defined, unambiguous interpretation of $2\pi n = e\mu\cancel{\varphi}$ for $\cancel{\varphi} = m > 1$ and do not contain physical singularities which would arise were a Dirac string to become physically observable. These solutions – were they to be so-permitted – are readily restructured into $e = (n/\cancel{\varphi})(2\pi/\mu)$, and so for complete multiple closed covers $\oiint_i F$ of the field strength bivector in $F = F_{\mu\nu} dx^\mu dx^\nu$, would be fractional charge solutions.

3. The Dirac Derivation Extended to Multiple Covers on SO(3), and how Semi-Infinite Dirac Strings Preclude Fractional Charges when Analyzed without Simply-Connected, Single-Valued Covering Groups

Because the physical space that we directly experience is that of the rotation group SO(3) with fluxes as shown in Figure 1, let us now extend the original Dirac derivation reviewed in the last section to multiple complete closed coverings, and see what transpires on SO(3).

First as already mentioned, we generalize the flux to $\oiint_i F$ with $1 \le i \le m$ to represent such an *m*-tuple covering over the azimuth domain $0 \le \varphi \le 2\pi m$, i.e., $0 \le \cancel{\varphi} \le m$ where *m* represents the quantized number of complete covers. This means the integral in (2.6) will now be given by:

$$\sum_{i=1}^{m} \oiint_i F = \sum_{i=1}^{m} \oiint_m \frac{\mu}{4\pi} d\cos\theta d\varphi = \frac{\mu}{4\pi} \int_0^\pi d\cos\theta \int_0^{2\pi m} d\varphi = \frac{\mu}{4\pi} \cos\theta \Big|_0^\pi \varphi \Big|_0^{2\pi m} = m\mu \,. \tag{3.1}$$

But if we are now taking φ over the domain $0 \le \varphi \le 2\pi m$ then we also need to pay close attention to the closed curves (small and otherwise) around which we are measuring a change in phase, that is, we need to carefully attend to $\oint d\Lambda$.

To do this, let us return to Figure 1 and imagine that we have now developed the network of small closed curves such each open surface enclosed by each small closed curve has been replicated *m* times, once on each covering surface. We can give each cover *C* a number *i* such that $1 \le i \le m$, and then designate that cover as C_i. To any individual small closed curve on one of these *m* closed surfaces C_i, we can relate another small closed curve on a different one of these *m* closed surfaces. To each of the "related" small closed curves on the i^{th} surface $1 \le i \le m$, we can assign the characteristic quantum number n_i. If $n_i = 0$ then there is no string passing through the surface bounded by that small closed curve. If there is a string passing through, then the characteristic number will be $n_i \ne 0$. Roughly speaking, when we say that *m* small closed curves on the *m* different surfaces are "related," we are envisioning that each of the n_i with $1 \le i \le m$ are radially oriented relative to a "center" of $\oiint_i F$ at substantially similar spatial angles φ and θ, and that we network the small closed curves not only on any given cover C_i, but as between one cover and the next. More precisely, however, what we have in mind that that if a (non-observable) string passes through, say, the first surface with characteristic n_1, then by considerations of continuity we will want to trace the passage of that string (or its non-passage should it terminate between two surfaces) through a "related" locale on the second surface with characteristic n_2, and likewise for related closed curves on the remaining surfaces. So if we have, say, 1000 small closed curves set up on C_1, then we will have 1000 small closed curves set up on each of the remaining C_i with $i > 1$, and will relate curves from one surface to the next via 1000 different sets of n_i with $1 \le i \le m$.

In view of this network of relationships among these *m*-element sets of small closed curves from one cover to the next, equation (2.2) for a single set of small closed curves on *m* covers now becomes:

$$\sum_{i=1}^{m} \oint_i d\Lambda = 2\pi \sum_{i=1}^{m} n_i + \sum_{i=1}^{m} e \iint_i F \ , \tag{3.2}$$

Specifically, $\sum_{i=1}^{m} e \iint_i F$ tells us that we have an *m*-tuple of layers of small open surfaces bounded by small closed curves and that we will add up the field flux through each; $\sum_{i=1}^{m} n_i$ tells us to also sum up the n_i string characteristics passing through these surfaces; and $\sum_{i=1}^{m} \oint_i d\Lambda$ tells us to add up the changes in phase about each of these small closed curves on each of these small open surfaces.

Now, as we did for Figure 1, we build up a network of small closed curves all the way from the north pole to the south pole, stopping just short of the south pole, leaving open a final *m*-tuplet of small closed curves near the south pole. Working with (2.3) and (2.4) but for more than one cover, this means that with $\oiint_m F = \iint_m F + \iint_m F_{\mathrm{SCC}}$, (3.2) now becomes:

$$\sum_{i=1}^{m} \oint_m d\Lambda = 2\pi \sum \left(\sum_{i=1}^{m} n_i \right) + \sum_{i=1}^{m} e \iint_m F = \sum \left(\sum_{i=1}^{m} 2\pi n_i \right) + e \sum_{i=1}^{m} \oiint_i F - e \sum_{i=1}^{m} \iint_i F_{\mathrm{SCC}} \ . \tag{3.3}$$

Now, as in (2.5) we take the limit as $r \to 0$ for the final *m*-tuplet of small closed curves near the south pole. But, for the *m*-tuple covering, we have *m* final $n_{\mathrm{SCC}\,i}$, $1 \le i \le m$, each of which is positive or negative or zero. Now, in this calculus limit, $\sum_{i=1}^{m} \oint_m d\Lambda \to 0$ and $\sum_{i=1}^{m} \iint_i F_{\mathrm{SCC}} \to 0$, so that (3.3) becomes:

$$0 = \sum_{i=1}^{m} \oint_m d\Lambda = 2\pi \sum \left(\sum_{i=1}^{m} \left(n_i + n_{\mathrm{SCC}\,i} \right) \right) + e \sum_{i=1}^{m} \oiint_i F = 2\pi \sum \left(\sum_{i=1}^{m} \left(n_i + n_{\mathrm{SCC}\,i} \right) \right) + m e \mu \ . \tag{3.4}$$

This readily restructures into:

$$e\mu = -\frac{1}{m} 2\pi \sum \left(\sum_{i=1}^{m} \left(n_i + n_{\mathrm{SCC}\,i} \right) \right), \tag{3.5}$$

Now we reach the non-trivial of exactly how one now works with the sums $\sum \left(\sum_{i=1}^{m} \left(n_i + n_{\mathrm{SCC}\,i} \right) \right)$. Of course, there is nothing special about the $n_{\mathrm{SCC}\,i}$ over and above the other n_i; this was just a mathematical device to allow us to take a calculus limit and so we can merge these together to redefine $n_i + n_{\mathrm{SCC}\,i} \to n_i$ so that the new n_i includes the $n_{\mathrm{SCC}\,i}$. And the minus sign is just a matter of convention so we can flip that. Then (3.5) is simply:

$$e\mu = \frac{1}{m} 2\pi \sum \left(\sum_{i=1}^{m} n_i \right). \tag{3.6}$$

The real, non-trivial question arises as the *independence or lack of independence* of the n_i in any given set of "related" n_i from one cover to the next. If we *assume* these to be *independent*, i.e., that for "related" small closed curves on different covers C_i the characteristic n_i can have different values so that, for example, we might have an $n_1 \neq n_2$ even though these are "related," then we can assert that $\sum \left(\sum_{i=1}^{m} n_i \right)$ is just a sum of *independent* integers, and thus may itself be any integer whatsoever. If we do so, then as we did from (2.7) to (2.8) we may redefine $\sum \left(\sum_{i=1}^{m} n_i \right) \to n$, in which case (3.6) above becomes:

$$e\mu = 2\pi \frac{n}{m}; \quad n = 1, 2, 3...; \quad m = 1, 2, 3...; \tag{3.7}$$

and we have the exact same fractional charge solution $2\pi n = e\mu\cancel{\varphi}$ with $\cancel{\varphi} = m = 0, 1, 2, 3...$ which descends from Wu and Yang's $\exp(i\Lambda) = \exp(ie\mu\cancel{\varphi})$ which in turn solves their differential equation $e^{-i\Lambda} de^{i\Lambda} = ie\mu d\cancel{\varphi}$. But can we do this? Is it appropriate to redefine $\sum \left(\sum_{i=1}^{m} n_i \right) \to n$ to be any integer whatsoever, or are there some restrictions that we must impose?

Let us now postulate a single semi-infinite Dirac unobservable string that passes from inside the first cover through all *m* covers and terminates at spatial infinity. By virtue of this, the small closed curves on each surface through which this invisible string passes are what we have called "related." Let us suppose that at its passage through a small closed curve on the first cover this string has a non-zero characteristic n_1, which is a *characteristic of the string itself*. Then, when this string passes through the related small closed curve on the second cover it must also have the same characteristic, with $n_2 = n_1$. And due to the postulated passage of this string from inside the first cover out to spatial infinity, each of the related n_i *for this same string must all have the same value from one cover to the next*. So these n_i are *not independent* from one cover to the next; they are all the same. Therefore the sum in (3.6) must be $\sum_{i=1}^{m} n_i = mn_1$.

As a consequence, if all of the strings are semi-infinite and pass through all the covers as postulated, (3.6) will reduce to:

$$e\mu = \frac{1}{m} 2\pi \sum \left(\sum_{i=1}^{m} n_i \right) = \frac{1}{m} 2\pi \sum mn_1 = 2\pi \sum n_1 = 2\pi n, \tag{3.8}$$

with a final trivial redefinition $\Sigma n_1 \to n$. In this situation the apparent fractional denominator *m* in (3.5) through (3.7) is cancelled by $\sum_{i=1}^{m} n_i = mn_1$ deriving from the hypothesized semi-infinite string via $m/m=1$, and we arrive at the "ordinary" DQC of $e\mu = 2\pi n$, notwithstanding that we have employed multiple closed covers. So if all the strings are semi-infinite passing from inside the first cover out to spatial infinity, and if the "last word" is an analysis in the physical space of the SO(3) rotation group which is not simply connected but rather is multi-valued, then no matter how many covers we employ, the DQC will always emerge to be $e\mu = 2\pi n$, and there will be no charge fractionalization, at least as a result of Dirac's monopole analysis.

In order for these fractionalized charges to exist and be topologically well-defined, there are two approaches one might consider. First, what prevented these fractionalized charges via the $m/m=1$ cancellation in (3.8) was the postulate that the Dirac strings were semi-infinite and passed through all the covers. If we relax this postulate and permit a *finite* Dirac string to start inside the first cover but terminate before it passes through all of the covers and reaches spatial infinity, then we can allow the n_i within a "related" set of small closed curves to be independent of one another and can thereby readmit the fractional charges. But this creates other problems, because a string that does not reach all the way to spatial infinity cannot be made to have a wavefunction formally equal to zero, $\psi(r)=0$, but can only have a small but finite wavefunction $\psi(r)\cong 0$. As will be discussed in more detail toward the end of the next section, this necessarily leads to physically observable singularities, which are unacceptable. So if one is looking to find some physical meaning in these fractional Wu-Yang charges, this is not the way to do so.

The second approach which appears far more viable is to recognize that the entire analysis so far has been conducted in the physical space of the SO(3) rotation group which, again, is not simply connected but rather is multi-valued. So the removal of these fractions in (3.8) may not be a problem having to do with fractional charges themselves, but rather, one due to the projective representation of SO(3). That is, as pointed out by [5] at page 4:

> "[O]ne way of getting around the multi-valuedness of projective representations is to view this as a problem not with the representation but rather with the underlying group. From this standpoint, the true symmetry at work is not exactly G but something closely related to it, the covering group $\tilde{G}$. Technically, we require a homomorphism $\tilde{G} \to G$ which is onto and many-to-one, but in just such a way that the multi-valued or projective representations of G descend from genuine, single-valued representations of $\tilde{G}$."

This is the approach upon which we shall now embark, and as we shall demonstrate, when projections are developed onto SO(3) from kernels $\sqrt[m]{1}$ which represent the m^{th} roots of unity "without strings," well-defined fractional charges with the denominator *m* and the fractional DQC $e\mu = 2\pi n/m$ of (3.7) can indeed be unambiguously included, and the $\not{\varphi} > 1$ solutions of Wu-Yang need no longer be discarded as unphysical.

4. Roots of Unity, the Complete Symmetries of Root Covering groups, and the Missing Fractional Symmetries of SO(3) and SO(1,3)

Once we add imaginary and complex numbers to the system of real numbers, the taking of square, cubic, quartic and any other roots acquires some very deep complexities that reach into many arenas of advanced mathematics and physics. For, if we take note of the fact that the range number 1 can be readily represented in the complex plane by $1=\exp(i2\pi n)$ for the domain of the infinite set of integers *n*, then the m^{th} root of 1 is given by the Euler relation:

$$\sqrt[m]{1}=\exp i\vartheta=\exp\left(i2\pi\frac{n}{m}\right)=\cos\left(2\pi\frac{n}{m}\right)+i\sin\left(2\pi\frac{n}{m}\right). \tag{4.1}$$

For any integer *m* there thus exists a set of *m* distinct m^{th} roots with $0\le n\le m-1$ which then recycle themselves for $m\le n\le 2m-1$ and so on. For $n=0$ or for $n/m=k$ being itself an integer *k*, this yields the trivial root $\sqrt[m]{1}=1$. But otherwise these roots have a rich and non-trivial multivalued structure as complex numbers $a+bi$ defined on the unit sphere in an imaginary plane, at evenly-spaced angular dispositions $\vartheta=2\pi n/m$.

Why is this of interest here? Because when we solve the Wu-Yang differential equation $e^{-i\Lambda}de^{i\Lambda}=ie\mu d\varphi$ we obtain $\exp(i\Lambda)=\exp(ie\mu\varphi)$ which in turn is solved by $2\pi n=e\mu\varphi$ with $\varphi=m=0,1,2,3....$ which is a fractional variant $e\mu=2\pi n/m$ of the Dirac Quantization condition, *and because this fractional Wu-Yang term* $e\mu=2\pi n/m$ *is identical to this angle* $\vartheta=2\pi n/m$ *in the Euler relation for the* m^{th} *root of unity.* So if we connect this Wu-Yang solution with (4.1) by setting $\vartheta=e\mu=2\pi n/m$, then (4.1) becomes:

$$\sqrt[m]{1}=\exp i\vartheta=\exp\left(i2\pi\frac{n}{m}\right)=\cos\left(2\pi\frac{n}{m}\right)+i\sin\left(2\pi\frac{n}{m}\right)=\exp(ie\mu)=\cos(e\mu)+i\sin(e\mu). \tag{4.2}$$

In the situation where $\sqrt[m]{1}=1=\exp(i2\pi n)$ which is the trivial root of unity, (4.2) results in $e\mu=2\pi n$ which is the standard DQC. But for the other non-trivial roots, the result is something other than the standard DQC, and the Wu-Yang fractional denominator *m* becomes associated directly with the m^{th} roots of unity.

This is important because as just noted from [5], “the multi-valuedness of projective representations is . . . a problem not with the representation but rather with the underlying group. From this standpoint, the true symmetry at work is not exactly *G* but something closely related to it, the covering group $\tilde{G}$.” Specifically, if we are projecting a covering group $\pi:\tilde{G}\to SO(3)$ and this projection has the kernel $\ker\pi=\sqrt[m]{1}$ in (4.2), then what will be missing from the symmetries of SO(3) but included in the symmetries of $\tilde{G}$, are precisely the *multivalued non-trivial roots of unity* which in (4.2) are connected with Wu-Yang fractional charges except in the special case where $n/m=k$ is itself an integer *k*. So while the analysis of the last section appeared to preclude fractional charges when analyzed in terms of SO(3) because of the $m/m=1$ cancellation in (3.8), the missing SO(3) symmetries which are supplied and made exact by $\tilde{G}$ appear to provide the precise means to trump this term cancellation via the multivalued roots of unity belonging to $\tilde{G}$ and thus to project fractional charges onto SO(3). This occurs in very much the same way that $\pi: SL(2,C) \to SO(1,3)$ projects spin ½ fermions and all related physics into SO(1,3) in the Dirac theory of the electron, even though all of this physics is not apparent if – as did Schrödinger and Klein-Gordon in the early 1920 – we only consider $SO(1,3)$ alone without the view of the underlying Clifford algebra $SL(2,C)$ which projects onto $SO(1,3)$. Let us now take a few moments to examine all this in detail.

From the time of Pythagoras through when Dirac in [6], [7] developed the first-order wave equation $(i\partial\!\!\!/ - m)\psi = 0$, one always calculated squares of lengths or intervals, and then took a square root to find an invariant physical length or interval at first order. Pythagoras taught that a radial distance $r^2 = x^2 + y^2 + z^2$, so to find the radius itself one would take the square root $r = \pm\sqrt{x^2+y^2+z^2}$ which in the sense of (4.1) and (4.2) can be written as $r = \exp(i\pi n)\sqrt{x^2+y^2+z^2}$. Minkowski [8] had established that when space and time were taken together, the interval $ds^2 = dt^2 - dx^2 - dy^2 - dz^2$ between two events was Lorentz-invariant, which Einstein [9] thereafter generalized from Minkowski spacetime with $\mathrm{diag}(\eta_{\mu\nu}) = (1,-1,-1,-1)$ into non-Euclidean Riemannian spacetime $\eta_{\mu\nu} \to g_{\mu\nu}$ and formalized into the metric equation $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$. Here too, one may take the square root $ds = \pm\sqrt{g_{\mu\nu}dx^\mu dx^\nu}$, or in the sense of (4.1) and (4.2), $ds = \exp(i\pi n)\sqrt{g_{\mu\nu}dx^\mu dx^\nu}$.

So to linearize the Klein-Gordon wave equation which is also known as the relativistic Schrödinger equation, what Dirac did was to deconstruct $\eta_{\mu\nu}$ into a set of 4x4 matrices γ^μ satisfying the anticommutator relationship $\eta^{\mu\nu} = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\}$, which represented a more sophisticated way of taking square roots of identity matrices *I* or of negative identity matrices -*I*, given that $\eta^{\mu\nu}$ was now the square of γ^μ but with these γ^μ anti-commuted with suitable indexing. This is the most important example of a Clifford Algebra. In so doing, Dirac made use of the non-commuting quaternions $i^2 = j^2 = k^2 = ijk = -1$ first developed by Hamilton in 1843 which by 1925 had been represented in the spin matrices $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = -i\sigma_1\sigma_2\sigma_3 = I$ of Pauli which are summarized in index notation as $\{\sigma_i, \sigma_j\} = \delta_{ij}$ and $i\varepsilon_{ijk}\sigma_k = \frac{1}{2}[\sigma_i, \sigma_j]$, by embedding these σ_i in the space components of γ^i.

Indeed, what Pauli had effectively done following Hamilton's lead, was to go from using $1 = (-1)^2 = \exp(i\pi n)^2 = \exp(i2\pi n)$ to describe the square root of the real number 1, to using the three distinct σ_i matrices in $I = \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ to describe the square roots of the 2x2 identity matrix *I*. With these fundamental steps by Pauli and Dirac, the second-order relativistic relationship between mass and energy-momentum $\eta^{\mu\nu} p_\mu p_\nu - m^2 = 0$ which is directly tied to the metric equation $g_{\mu\nu} dx^\mu dx^\nu - ds^2 = 0$ could be written in first order, not as the $\pm\sqrt{\eta^{\mu\nu} p_\mu p_\nu - m^2} = 0$ square root of a regular number, but rather as $\left(\gamma^\mu p_\mu - m\right)u = 0$, where $\gamma^\mu p_\mu - m$ now operates on a complex spinor u which defines the eigenstates of the energy-momentum operator $\not{p} = \gamma^\mu p_\mu$ corresponding to the mass eigenvalues *m*. With these steps, $\sqrt[2]{1} = \exp\left(i\pi n\right) = \pm 1$, which is (4.1) for *m*=2, and which applies to regular numbers, became the *kernel* of these generator matrices σ_i and γ^μ which could be squared into *I* or *–I* identity or negative identity matrices.

At a superficial level, one might have taken the view that this was simply a clever mathematical exercise in taking roots of unity – in this case, square roots of unity – with no meaning beyond the mathematics. But as its turned out, these generator matrices themselves, be they Pauli's σ_i or Dirac's γ^μ and its Clifford algebra, revealed a new, deeper geometric structure *in the natural world itself* which could not be seen, for example, by merely writing $\pm\sqrt{\eta^{\mu\nu} p_\mu p_\nu - m^2} = 0$ but could very well be seen by writing $\left(\gamma^\mu p_\mu - m\right)u = 0$. Or, going back to Pythagoras, these revealed a deeper structure which could not been seen by writing $\pm\sqrt{x^2 + y^2 + z^2 - r^2} = 0$ but could have been seen by using the Pauli matrices to write $\left(\sigma_\mu x^\mu - r\right)u = 0$. Most importantly, *it turned out that all these new features of cleverly taking mathematical roots of identity matrices appear in the real, observable physical world.* They do so in the form of the spin ½ of fermions, the existence of positrons and other antimatter, chirality and parity, fine splitting in atomic spectra, and many other phenomena which nine decades later have been as firmly established as any connection between physical theory and empirical observation. So although we live in the second-order physical space of $x^2 + y^2 + z^2 - r^2 = 0$ characterized as $SO(3)$ which leaves r^2 unchanged under rotations mixing the *x*, *y* and *z* coordinates, as well as in the second order physical spacetime of $g_{\mu\nu} dx^\mu dx^\nu - ds^2 = 0$ which is tangentially that of Minkowski characterized by $SO(1,3)$ which leaves ds^2 unchanged under rotations and Lorentz transformations, there also exist in nature, certain "root spaces" that we do not directly experience, but which nonetheless "project" π real observable physics onto these spaces in which we do live.

So, for example, $\pi : SU(2) \to SO(3)$ projects the two-component spinors of Pauli which are the eigenstates of $SU(2)$ onto the $SO(3)$ physical space in which we live, while $\pi : SL(2,C) \to SO(1,3)$ projects the four-component Dirac spinor eigenstates of $SL(2,C)$ onto the spacetime theater of $SO(1,3)$ in which we directly experience nature. What happens in these root spaces does *not* stay in these root spaces; to use a pun, these spaces are not Las Vegas. Everything that happen in these (square) root spaces does get projected out onto $SO(3)$ and $SO(1,3)$ and manifests itself in the physics we observe in even these spaces, *even though none of this is apparent unless and until we become aware of and formally develop these root spaces and* *their projections into* $SO(1,3)$ *and its spatial subset* $SO(3)$. Spin ½ and all the other consequences of Dirac's equation do not originate in the real spacetime of $SO(1,3)$ which preserves $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ as an invariant. Nor do these originate in the real space of Pythagoras which preserves $r^2 = x^2 + y^2 + z^2$ as an invariant. Rather, they originate in the complex Hilbert space of $SL(2,C)$ in which $\left(\gamma^\mu p_\mu - m\right)u = 0$ and makes m an eigenvalue for the spinors u of the operator $\not{p} = \gamma^\mu p_\mu$. These features of the Clifford algebra in $SO(1,3)$ *are then projected out onto* $SO(1,3)$.

So now, let us go back to (4.2) which does embed the fractional $e\mu = 2\pi n / m$ of Wu-Yang. To use the language of kernels, $SO(3)$ is the Pythagorean space of $r^2 = x^2 + y^2 + z^2$ for which no roots have been taken and onto which no projections have been made, which is to say, it is the space for which $\ker \pi : \sqrt[m]{1} = \sqrt[1]{1} = 1$ with m=1 and $\pi : SO(3) \to SO(3)$. It is just SO(3) being itself with no further analysis. Via (4.2), m=1 means that $e\mu = 2\pi n$ which is the standard DQC, even though (4.2) embeds $e\mu = 2\pi n / m$. Why is this important? Because in (3.8) of section 3 and Figure 1, the entire analysis was carried out using Dirac strings in SO(3) with no resort to any square, cubic or other root space, and we found out that for any semi-infinite strings penetrating all of coverings, we are indeed restricted to the conventional $e\mu = 2\pi n$. But it is not that nature forbids the fractional charges of $e\mu = 2\pi n / m$; it is just that the SO(3) topology is not simply connected and does not contain enough of the richness of nature's root subspaces to project out a fractional charge onto SO(3). This is similar to how if we analyze spin in SO(3) alone, we will never be able to explain spin ½ fermions and all the attendant phenomena and will only see whole-unit spins. But of course the natural world exhibits more. The limitation is simply that SO(3) and SO(1,3), by themselves, also lack the richness to reveal the spin ½ fermions and all else with which these are connected, and that we cannot see this without going into a root space and projecting $\pi : SU(2) \to SO(3)$ or $\pi : SL(2,C) \to SO(1,3)$.

Indeed, the derivation of the standard DQC $e\mu = 2\pi n$ and the *apparent* cancellation of fractional charges in (3.8) notwithstanding multiple coverings can be further understood if now take any and all m^{th} roots as they appear in (4.2), but consider only those m^{th} roots for which $\sqrt[m]{1} = \exp\left(i2\pi n / m\right) = 1$. That is, let us now consider (4.2) for all roots, but only in the special case of the trivial root of unity $\sqrt[m]{1} = 1$. In order to ensure that $\sqrt[m]{1} = 1$ we must have $n = mk$ where k is an integer, whereby (4.2) becomes:

$$1 = \sqrt[m]{1} = \exp i\vartheta = \exp\left(i2\pi \frac{n}{m}\right) = \exp\left(i2\pi \frac{mk}{m}\right) = \exp\left(i2\pi k\right) = \exp\left(ie\mu\right). \tag{4.3}$$

This too is the standard DQC, now $e\mu = \vartheta = 2\pi k$. And here, the fractional m in the denominator is cancelled out by $n = mk$ in the numerator via the exact same $m/m = 1$ cancellation that we saw in (3.8). So now we have a question: Are these identical results $e\mu = \vartheta = 2\pi k$ derived through an identical cancellation $m/m = 1$ somehow linked, whereby (4.3) is somehow a topological restatement of the string analysis which led to (3.8)?

In section 3 we considered *m* coverings generally of SO(3), but when we mandated that all strings pass through all covers and terminate at spatial infinity, we found that for a string with characteristic n_1 passing through all of the covers it was required that $\sum_{i=1}^{m} n_i = mn_1$. This is what offset what otherwise had been a fractional Wu-Yang charge. Above, the correspondence to having *m* coverings generally is having kernels with $\ker\pi = \sqrt[m]{1}$, and the correspondence to a string running through all covers to spatial infinity corresponds to only taking the trivial roots for which $\ker\pi = \sqrt[m]{1} = 1$ as we did in (4.3).

Of course, the Dirac string is a fictive entity, so let's explore exactly what this fiction entails: First, by using Dirac strings we are not really changing the effective aterminal configuration of magnetic fields. We are simply postulating a solenoid which is infinitely thin, so that at one mouth of the solenoid postulated to be in a region of space that is finitely accessible, it will be impossible to ascribe a spatial orientation to the solenoid, and therefore the magnetic field lines will emanate from the solenoid with a complete rotational symmetry, that is, they will be mono-polar field lines. Any finite string thickness would destroy the mono-polar character of the field lines because it would establish a physically-detectable direction for the solenoid. Of course, we are simultaneously postulating that the field lines are infinitely dense inside this solenoid because of its infinitesimal width, which is another aspect of the fiction.

Second, we are postulating that the other end of the solenoid accumulates or deposits the field lines at spatial infinity. "Sweeping under the rug" at infinity is a genuine and apt metaphor for this. But, there is good reason for this: one of the benefits of using spatial infinity in physics is that we can make any field formally vanish at infinity. This includes a wavefunction field which can formally be made to have $\psi(r) \to 0$ as $r \to \infty$. Were we to situate the other end of the string at anywhere other than spatial infinity, we could not make this field formally equal to zero. We could only make it very, very small but finite, $\psi(r) \cong 0$. Since the Dirac string is constructed for the "exceptional case . . . occurring when the wave function vanishes," we need to first situate one end of the string at a locale where the wavefunction can truly be equal to zero, and that locale is $\psi(\infty) = 0$. Then we can run an infinitely-thin string from infinity with vanishing wavefunction $\psi(r) = 0$ continuing unabated toward a finitely-accessible region of space, without creating any observable singularity, which continued "vanishing will require two conditions, so that in general the points at which it vanishes will lie along a line."

Were we to put the far end of the string *anywhere but at spatial infinity*, we could could have a $\psi(r) \cong 0$ which is exceedingly small, but this could never be formally equal to zero. Therefore, the string itself would become a solenoid of *finite* width defined about some line where $\psi(r) \cong 0$ is sufficiently close to zero within some bounds, but is not formally equal to zero. Because of the small albeit finite solenoid width the string singularity would become observable which is unphysical and impermissible. Further, the magnetic monopoles would no longer be monopoles but just the regular magnetic fields emanating from a very long, very thin solenoid which approach a monopole configuration as the solenoid becomes longer and thinner.

Now, as we saw in (3.8), the $m/m=1$ offset which resulted in the standard DQC $e\mu=2\pi n$ without charge fractions $e\mu=2\pi n/m$ was a direct result of postulating that the Dirac strings are semi-infinite and thus passed through all the covers. But as we have now explained in detail, if we allow the strings to be anything other than semi-infinite to avert the $m/m=1$ cancellation of (3.8), we introduce other serious problems which are unacceptable.

All of the above is very true. But what is equally true is that all of this describes a picture in the three-dimensional $r^2=x^2+y^2+z^2$ space of SO(3) which is not simply connected. When we start to consider root spaces, this same set of circumstances are described by $1=\sqrt[m]{1}$. That is, an SO(3) space with *m* covers and fictive strings which are all semi-infinite and infinitely thin yields the same result as projecting onto SO(3) from a covering group $\tilde{G}$ using the projection $\pi:\tilde{G}\to SO(3)$ with a kernel $\ker\pi=\sqrt[m]{1}$ restricted to the trivial root of unity $1=\sqrt[m]{1}$. But here is the pivotal difference that emerges from conducting our analysis using $\tilde{G}$ rather than SO(3):

When we ask in the context of SO(3) if there is some way to restore the fractional Wu-Yang charges $e\mu=2\pi n/m$ which were offset by $\sum_{i=1}^{m} n_i = mn_1$, the answer is that we must have the string pass through only some of the covers, which yields an unacceptable, singular result for the reasons just outlined. But when we ask this same question in the context of $\pi:\tilde{G}\to SO(3)$ with $\ker\pi=\sqrt[m]{1}$, we come to understand that what we need to do is find a non-trivial $\tilde{G}$ which is a sophisticated m^{th} root $\sqrt[m]{SO(3)}$, in exactly the same spirit that Dirac sought to find a square root of the Klein-Gordon equation whereby $SL(2,C)$ is non-trivial square root $\sqrt[2]{SO(1,3)}$ of Minkowski's $SO(1,3)$, or Pauli following Hamilton defined $SU(2)$ which is a similar square root $\sqrt[2]{SO(3)}$ of the Pythagorean $SO(3)$. And then, we must lift all restriction from $\ker\pi=\sqrt[m]{1}$, and consider *all* of the non-trivial roots of unity, and not just the trivial $1=\sqrt[m]{1}$.

Can this sort of approach be made to work? As pointed out at the end of the last section from [5], "the true symmetry at work is not exactly *G* [here, SO(3)] but something closely related to it, the covering group $\tilde{G}$." So when we analyzed the fractional charge solution $e\mu=2\pi n/m$ via an analysis using only SO(3), we were not fully accounting for *all of the symmetries* at work. We were only using an approximate symmetry. The exact symmetries are to be found in the covering groups $\tilde{G}$. So if we can develop what can informally be described as $\sqrt[m]{SO(3)}$ and is more formally described as $\pi:\tilde{G}\to SO(3)$ with $\ker\pi=\sqrt[m]{1}$, and if we can develop this for *all* of the primitive roots $\sqrt[m]{1}$ and not only the trivial root $1=\sqrt[m]{1}$, and if the fractional charges $e\mu=2\pi n/m$ can be well-defined without ambiguity in this way, then the fractional charges can become real, observable entities projected from a root space into SO(3), in the same way that spin $1/2$ is projected into $SO(1,3)$ from $SL(2,C)$.

Even more to the point: SO(3) which precludes fractional charges for the reasons laid out in section 3 and in the above, is not an exact symmetry. Something is missing. What is missing is the exact symmetry of the covering group $\tilde{G}$. The symmetry that a covering group $\tilde{G}$ with $\ker\pi = \sqrt[m]{1}$ contains, which SO(3) does not contain, are all m^{th} roots of unity which sit on the unit circle in the complex plane at evenly-spaced angles of $\vartheta = 2\pi n / m$ running from $0 \le \vartheta < 2\pi$ before recycling at $\vartheta = 2\pi$. SO(3) only admits one root at $\vartheta = 2\pi n$. So the symmetry missing from SO(3) which is not missing from groups with $\ker\pi = \sqrt[m]{1}$, is the very symmetry which admits rather than precludes the fractional Wu-Yang charges $e\mu = 2\pi n / m$. Thus, even though SO(3) misses the fractional charges just as $SO(1,3)$ alone missed spin ½, there is nothing to prevent a covering group $\tilde{G}$ which has a more complete symmetry from projecting these fractional charges onto SO(3) any more than $SL(2,C)$ can be prevented from projecting spin ½ and all of its associated physics onto $SO(1,3)$.

Even more concisely: if one is challenged to show observable physics for spin ½ using $SO(1,3)$ alone, the answer is that this cannot be done without also having access to $SL(2,C)$. If one is likewise challenged to show observable physics for fractional W-Yang charges $e\mu = 2\pi n / m$ using Dirac strings in SO(3) alone, the answer is that this also cannot be done, here, without having access to some covering group $\tilde{G}$ with an m^{th} root $\ker\pi = \sqrt[m]{1}$.

That is the theory; now we need to turn to the mechanics of this theory. For while it is one thing to suggest constructing these $\sqrt[m]{SO(3)}$ groups informally, and to more formally suggest constructing $\pi : \tilde{G} \to SO(3)$ with $\ker\pi = \sqrt[m]{1}$ with all roots trivial and non-trivial appearing alike, it is another thing to demonstrate that such groups actually exist and can indeed be unambiguously constructed. So now, we shall demonstrate precisely how to construct these generalized root covering groups by considering the easiest example of the cubed root covering group ${}^3\tilde{G}$ defined with the projection $\pi : {}^3\tilde{G} \to SO(3)$ with $\ker\pi = \sqrt[3]{1}$ and a DQC $e\mu = 2\pi n / 3$. In the course of constructing ${}^3\tilde{G}$, it will become apparent how to generally construct $\pi : {}^m\tilde{G} \to SO(3)$ which $\ker\pi = \sqrt[m]{1}$ for a generalized fractional DQC with a DQC $e\mu = 2\pi n / m$, notwithstanding the fact that SO(3) misses this symmetry. It is on this basis that the Wu-Yang fractional charges in $e\mu = 2\pi n / m$ can lead to observable physics that is well-defined and non-singular.

5. Roots of the Identity matrix, and their Relationship to the Spin Matrices of SU(2)

As discussed in the last section, the Pauli's matrices $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = I$ effectively represent square m=2 roots of unity for a 2x2 identity matrix. Using projective language, and also using (4.1) with m=2, we may say that these σ_i matrices are used to generate the projection:

$$\pi : {}^2\tilde{G} = SU(2) \to SO(3) : \ker\pi = \sqrt[2]{1} = \exp i\vartheta = \exp\left(i2\pi n / 2\right) = \exp\left(i\pi n\right) = \pm 1. \tag{5.1}$$

Here, the covering group $\tilde{G}$ is $SU(2)$. The fact that the generators $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = I$ square to the unit matrix *I* is tied to the kernel of this projection being $\ker \pi = \sqrt[2]{1} = \pm 1$, that is, to the fact that $(\pm 1)^2 = 1$. The left superscript in ${}^2\tilde{G}$ is used to designate that the kernel is the *m*=2 square root of unity and the generators are *m*=2 square roots of the identity matrix. Obviously, one of these kernel roots is trivial, namely, the one for which $\sqrt[2]{1} = +1$. So corresponding to this trivial root one has the "three" generators $I_1^2 = I_2^2 = I_3^2 = I$ which are all unit matrices. The $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = I$ are the non-trivial generators corresponding to the non-trivial root $\sqrt[2]{1} = -1$.

Now we wish to do is generalize (5.1) to any and all m^{th} roots of unity, that is, to:

$$\pi : {}_m\tilde{G}(2) \rightarrow SO(3) : \ker \pi = \sqrt[m]{1} = \exp i\vartheta = \exp\left(i2\pi n / m\right) \tag{5.2}$$

for any and all *m*. Here, the left subscript in ${}_m\tilde{G}(2)$ is used to designate that the kernel is the m^{th} root of unity generally. We add the parenthetical (2) to $\tilde{G}(2)$ to designate that the generators of ${}_m\tilde{G}(2)$, which we shall designate as τ_i, must be 2x2 matrices just like the σ_i generators of SU(2). As we shall see, for a given *m*, this will lead to there being a total of *m* sets of these τ_i, each set corresponding to one of the m^{th} roots, and each set containing precisely three generators τ_i with $i = 1, 2, 3$. For any given *m*, these generators must be constructed such that ${\tau_1}^m = {\tau_2}^m = {\tau_3}^m = I$ to mirror the kernel $\ker \pi = \sqrt[m]{1}$. Out of these, *m*-1 sets of these τ_i generators are non-trivial, while the final set contains the trivial $\tau_i = I_i$ which yield $I_1^m = I_2^m = I_3^m = I$ corresponding to the trivial kernel for which $\sqrt[m]{1} = +1$.

Were we to use the Wu-Yang solution to set $e\mu = 2\pi n / m$ above, then as we saw in (4.3), the standard DQC $e\mu = 2\pi n$ would result from the trivial generators $\tau_i = I_i$ through the $m / m = 1$ cancellation which appeared in both (3.8) and (4.3). But for all of the non-trivial τ_i, this cancellation would not occur, and it would be possible to then project these fractional charges onto $SO(3)$ by taking advantage of the multiple roots which do appear in the exact symmetry of the ${}_m\tilde{G}(2)$ level, but do not appear in the close-but-inexact symmetry of the $SO(3)$. So, our goal is to be able to use the Wu-Yang solution $e\mu = 2\pi n / m$ to extend (5.2) by setting $\vartheta = e\mu = 2\pi n / m$ and thus writing:

$$\pi : {}_m\tilde{G}(2) \rightarrow SO(3) : \ker \pi = \sqrt[m]{1} = \exp i\vartheta = \exp\left(i2\pi n / m\right) = \exp\left(ie\mu\right). \tag{5.3}$$

But to be able to support this proposed extension of (5.2) to admit Wu-Yang fractional charges, we must show that these roots groups ${}_m\tilde{G}(2)$ do exist for any and all *m*, can be given definite, and unambiguous representations, and can be unambiguously projected onto $SO(3)$. Again from [5], "[t]echnically, we require a homomorphism $\tilde{G} \rightarrow G$ which is onto and many-to-one, but in just such a way that the multi-valued or projective representations of *G* descend from genuine, single-valued representations of $\tilde{G}$." Each single valued representation of $\tilde{G}$ needs to be a matrix set τ_i which corresponds to one of the roots of unity in $\ker\pi = \sqrt[m]{1}$ such that $\tau_i^{\,m} = I$. So let us see how to form these root generators τ_i in general for any *m*, and let's see the way in which ${}_2\tilde{G}(2) = SU(2)$ contains the simplest special case of these τ_i which for SU(2) are equal to the Pauli matrices σ_i.

We start with the Pauli matrices σ_i themselves, posit three associated angles θ_i in physical space, and form the unitary matrices $U_i = \exp(i\sigma_i\theta_i)$ for SO(3) rotations through respective angles $\theta_i = \theta_x, \theta_y, \theta_z$ about each of the x, y, z axes. It is well-known how to use the series $e^{ix} = 1 + ix - \frac{1}{2!}x^2 - i\frac{1}{3!}x^3 + \frac{1}{4!}x^4 ...$ together with the fact that $\sigma_i^{\,2n} = 1$ and $\sigma_i^{\,2n+1} = \sigma_i$ to find these unitary matrices, namely:

$$\begin{aligned} U_1 &= \exp(i\sigma_1\theta_1) = \begin{pmatrix} \cos\theta_1 & i\sin\theta_1 \\ i\sin\theta_1 & \cos\theta_1 \end{pmatrix} \\ U_2 &= \exp(i\sigma_2\theta_2) = \begin{pmatrix} \cos\theta_2 & \sin\theta_2 \\ -\sin\theta_2 & \cos\theta_2 \end{pmatrix} \\ U_3 &= \exp(i\sigma_3\theta_3) = \begin{pmatrix} \cos\theta_3 + i\sin\theta_3 & 0 \\ 0 & \cos\theta_3 - i\sin\theta_3 \end{pmatrix} \end{aligned} . \qquad (5.4)$$

Now, it happens that with a judicious choice of the angles θ_i we can cause each of these U_i to be identical to the corresponding σ_i up to an overall constant factor. Specifically, if we choose each of these angles such that $\theta_i = \pi/2$, we readily see that:

$$\begin{aligned} U_1 &= \exp\left(i\sigma_1\frac{\pi}{2}\right) = \begin{pmatrix} \cos(\pi/2) & i\sin(\pi/2) \\ i\sin(\pi/2) & \cos(\pi/2) \end{pmatrix} = i\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = i\sigma_1 \\ U_2 &= \exp\left(i\sigma_2\frac{\pi}{2}\right) = \begin{pmatrix} \cos(\pi/2) & \sin(\pi/2) \\ -\sin(\pi/2) & \cos(\pi/2) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = i\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = i\sigma_2 \\ U_3 &= \exp\left(i\sigma_3\frac{\pi}{2}\right) = \begin{pmatrix} \cos(\pi/2) + i\sin(\pi/2) & 0 \\ 0 & \cos(\pi/2) - i\sin(\pi/2) \end{pmatrix} = i\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = i\sigma_3 \end{aligned} . \qquad (5.5)$$

Consolidating, in general we see that $U_i = \exp(i\sigma_i\pi/2) = i\sigma_i$, which we rewrite as:

$$\sigma_i = -i\exp\left(i\sigma_i\frac{\pi}{2}\right). \qquad (5.6)$$

We then square this to write $\sigma_i^2 = I_i = (-i)^2 \exp(i\sigma_i \pi)$ and deliberately do *not* turn $(-i)^2 \to -1$ because when we later take square roots of this, we want to recover –*i* and not extraneously introduce a two-valued $\pm i$. Of course, the identity matrix to any integer power *n* is still the identity matrix $I_i^n = I_i$, so that all of this finally yields:

$$I_i = I_i^n = (-i)^{2n} \exp(i\sigma_i \pi n). \tag{5.7}$$

Now, we may again use $\vartheta = 2\pi n / m$ thus $\vartheta / 2 = \pi n / m$ from the Euler relation (4.1) to write the m^{th} root of these 2x2 identity matrices (5.7) as:

$$\begin{aligned} \sqrt[m]{I_i} &= (-i)^{\frac{2n}{m}} \exp\left(i\sigma_i \frac{\vartheta}{2}\right) = (-i)^{\frac{2n}{m}} \left[\cos\left(\sigma_i \frac{\vartheta}{2}\right) + i \sin\left(\sigma_i \frac{\vartheta}{2}\right)\right] \\ &= (-i)^{\frac{2n}{m}} \exp\left(i\sigma_i \pi \frac{n}{m}\right) = (-i)^{\frac{2n}{m}} \left[\cos\left(\sigma_i \pi \frac{n}{m}\right) + i \sin\left(\sigma_i \pi \frac{n}{m}\right)\right] \end{aligned}. \tag{5.8}$$

In this form we see an overall coefficient $(-i)^{2n/m}$ that we need to write in a more generally usable form. So we turn again to the Euler expression in the form $-i = \exp(i3\pi/2)$ to write:

$$(-i)^{\frac{2n}{m}} = \exp\left(i3\pi \frac{n}{m}\right), \tag{5.9}$$

which we then place into (5.8) to obtain:

$${}_m^n\tau_i \equiv \sqrt[m]{I_i} = \exp\left(i3\pi \frac{n}{m}\right) \exp\left(i\sigma_i \frac{\vartheta}{2}\right) = \exp\left(i3\pi \frac{n}{m}\right) \exp\left(i\sigma_i \pi \frac{n}{m}\right) = \exp\left(i\pi \frac{n}{m}(\sigma_i + 3)\right). \tag{5.10}$$

In the final expression we see what is effectively a term $\sigma_i + 3I_i$, which does have a trace, so this uses the generators of U(2), not of the special unitary group SU(2). Above, because we will use τ_i to generally designate the m^{th} roots $\sqrt[m]{I_i}$ of the identity matrix, we have introduced a notational aid to help keep track of any given set of τ_i. Specifically, we represent these τ_i as ${}_m^n\tau_i$ where the *m* left subscript tells us the general root $\sqrt[m]{I_i}$ which the τ_i represent and the *n* left superscript tells us the *n* value $0 \le n < m$ of the particular m^{th} root being represented.

Now, to illustrate the use of (5.10) and to check the calculations which led to (5.10), let us take the *m*=2 square root of the identity matrices I_i. We simply set *m*=2 in (5.10) to obtain:

$${}_2^n\tau_i = \sqrt[2]{I_i} = \exp\left(i3\pi \frac{n}{2}\right) \exp\left(i\sigma_i \frac{\vartheta}{2}\right) = \exp\left(i3\pi \frac{n}{2}\right) \exp\left(i\sigma_i \pi \frac{n}{2}\right). \tag{5.11}$$

There are two solutions to this. For $n=0$ we of course have ${}_{2}^{0}\tau_i = \sqrt[2]{I_i} = I_i$ and $\vartheta = 0$ which is the trivial square root. For $n=1$ this becomes:

$$ {}_{2}^{1}\tau_i = \sqrt[2]{I_i} = -i\exp\left(i\sigma_i \frac{\vartheta}{2}\right) = -i\exp\left(i\sigma_i \frac{\pi}{2}\right) = \sigma_i , \tag{5.12} $$

which recovers (5.6) with $\vartheta = \pi = 180°$. These are the very two angles $\vartheta = 0$ and $\vartheta = \pi$ which we expect to have when we take a square root using the Euler relation, and confirms that the covering group ${}_{2}^{1}\tilde{G}(2) = SU(2)$, where we now add a superscript to the left of $\tilde{G}(2)$ to indicate that this correspondence (5.12) occurs when *n*=1 and *m*=2. In general, this means that we shall now represent the covering group as ${}_{m}^{n}\tilde{G}(2)$.

Now, in (5.10), we have everything we need to generally calculate all the non-trivial 2x2 root generator matrices ${}_{m}^{n}\tau_i$ for any m^{th} root of the 2x2 identity matrix, and have shown how this works for the Pauli matrices themselves. These ${}_{m}^{n}\tau_i$ will become the m sets of $i = 1,2,3$ generators for the covering groups ${}_{m}^{n}\tilde{G}(2)$ which we will then wish to project onto $SO(3)$ via $\pi : {}_{m}^{n}\tilde{G}(2) \to SO(3)$ with $\ker \pi = \sqrt[m]{1}$. If we can create these ${}_{m}^{n}\tilde{G}(2)$ and then carry out the projection onto $SO(3)$ without ambiguity, then the *m*-to-1 nature of this projection will validate projecting fractional charges with $e\mu = 2\pi n / m$ onto SO(3), with the regular DQC $e\mu = 2\pi k$ being the $n = mk$ solutions of $e\mu = 2\pi n / m$ that were found in (3.8) and (4.3).

6. The Homomorphic Mapping of Root-of-Unity Covering Groups onto the Rotation Group SO(3) through SU(2)

To establish an unambiguous the projection $\pi : {}_{m}^{n}\tilde{G}(2) \to SO(3)$ we start by finding a general expression for the commutator of any two ${}_{m}^{n}\tau_i$, which we denote by ${}_{m}^{n}\left[\tau_i, \tau_j\right] = {}_{m}^{n}\left[\tau_i\tau_j - \tau_j\tau_i\right]$ with the *n*, *m* designations moved outside the commutator to avoid visual clutter. Working from (5.10) we construct:

$$ {}_{m}^{n}\left[\tau_i, \tau_j\right] = \exp\left(i6\pi \frac{n}{m}\right)\left[\exp\left(i\sigma_i\pi \frac{n}{m}\right), \exp\left(i\sigma_j\pi \frac{n}{m}\right)\right]. \tag{6.1} $$

To evaluate this, it helps to also construct the commutators $\left[U_i, U_j\right]$ of the unitary matrices (5.4). This exercise is straightforward and yields:

$$\begin{aligned}
\left[U_1,U_2\right]&=\left[\exp\left(i\sigma_1\theta_1\right),\exp\left(i\sigma_2\theta_2\right)\right]=-2i\sin\theta_1\sin\theta_2\begin{pmatrix}1&0\\0&-1\end{pmatrix}=-2i\sin\theta_1\sin\theta_2\sigma_3\\
\left[U_2,U_3\right]&=\left[\exp\left(i\sigma_2\theta_2\right),\exp\left(i\sigma_3\theta_3\right)\right]=-2i\sin\theta_2\sin\theta_3\begin{pmatrix}0&1\\1&0\end{pmatrix}=-2i\sin\theta_2\sin\theta_3\sigma_1\ .\\
\left[U_3,U_1\right]&=\left[\exp\left(i\sigma_3\theta_3\right),\exp\left(i\sigma_1\theta_1\right)\right]=-2i\sin\theta_3\sin\theta_1\begin{pmatrix}0&-i\\i&0\end{pmatrix}=-2i\sin\theta_2\sin\theta_3\sigma_2
\end{aligned}\tag{6.2}$$

In the circumstance where $\theta\equiv\theta_1=\theta_2=\theta_3$ this consolidates to:

$$\left[U_i,U_j\right]=\left[\exp\left(i\sigma_i\theta\right),\exp\left(i\sigma_j\theta\right)\right]=-2i\sin^2\theta\varepsilon_{ijk}\sigma_k\,.\tag{6.3}$$

Thus, if we set $\theta=\pi n/m$ (6.3) becomes:

$$\left[\exp\left(i\sigma_i\pi\frac{n}{m}\right),\exp\left(i\sigma_j\pi\frac{n}{m}\right)\right]=-2i\sin^2\left(\pi\frac{n}{m}\right)\varepsilon_{ijk}\sigma_k\,.\tag{6.4}$$

Finally, inserting (6.4) into (6.1) and also applying $i\varepsilon_{ijk}\sigma_k=\tfrac{1}{2}\left[\sigma_i,\sigma_j\right]$, we finally obtain:

$${}^n_m\left[\tau_i,\tau_j\right]=-2i\exp\left(i6\pi\frac{n}{m}\right)\sin^2\left(\pi\frac{n}{m}\right)\varepsilon_{ijk}\sigma_k=-\exp\left(i6\pi\frac{n}{m}\right)\sin^2\left(\pi\frac{n}{m}\right)\left[\sigma_i,\sigma_j\right].\tag{6.5}$$

Alternatively, isolating σ_k with some simple re-indexing, this may be written as:

$$\sigma_i=\frac{1}{4}i\exp\left(-i6\pi\frac{n}{m}\right)\csc^2\left(\pi\frac{n}{m}\right)\varepsilon_{ijk}\,{}^n_m\left[\tau_j,\tau_k\right].\tag{6.6}$$

This means that the Pauli spin matrices σ_i and thus their commutator $\left[\sigma_i,\sigma_j\right]=2i\varepsilon_{ijk}\sigma_k$ can always be expressed as the commutator $\varepsilon_{ijk}\,{}^n_m\left[\tau_j,\tau_k\right]$ of the root of identity matrices ${}^n_m\tau_i\equiv\sqrt[m]{I_i}$, times an overall multiplying factor which is a defined function of *n* and *m*.

As a check we may set *m*=2 and *n*=1 in (6.5) to find that:

$${}^1_2\left[\tau_i,\tau_j\right]=-\exp\left(i3\pi\right)\sin^2\left(\pi/2\right)\left[\sigma_i,\sigma_j\right]=\left[\sigma_i,\sigma_j\right],\tag{6.7}$$

which is similarly a consequence of and thus compatible with the result ${}^1_2\tau_i=\sigma_i$ from (5.12).

The result in (6.5) is very important in making the general projection $\pi:{}^n_m\tilde{G}(2)\rightarrow SO(3)$. First, this is easily re-indexed into:

$$\frac{1}{4}\varepsilon_{ijk}\ {}^{n}_{m}\left[\tau_j,\tau_k\right]=-i\exp\left(i6\pi\frac{n}{m}\right)\sin^2\left(\pi\frac{n}{m}\right)\sigma_i\,. \tag{6.8}$$

Now, using the space coordinates $x^i=\left(x,y,z\right)$ and forming $\sigma_i x^i$ we can use (6.8) to write:

$$\frac{1}{4}\varepsilon_{ijk}\ {}^{n}_{m}\left[\tau_j,\tau_k\right]x^i=-i\exp\left(i6\pi\frac{n}{m}\right)\sin^2\left(\pi\frac{n}{m}\right)\sigma_i x^i=-i\exp\left(i6\pi\frac{n}{m}\right)\sin^2\left(\pi\frac{n}{m}\right)\begin{pmatrix} z & x-iy \\ x+iy & -z\end{pmatrix}. \tag{6.9}$$

Therefore, restructuring to isolate $\sigma_i x^i$, and also making use of the spinor relationships:

$$x=\tfrac{1}{2}\left(\xi_2{}^2-\xi_1{}^2\right);\quad y=\tfrac{1}{2i}\left(\xi_1{}^2+\xi_2{}^2\right);\quad z=\xi_1\xi_2 \tag{6.10}$$

as well as the cross product:

$$\varepsilon_{ijk}\ {}^{n}_{m}\left[\tau_j,\tau_k\right]x^i=2\ {}^{n}_{m}\left[\boldsymbol{\tau}\times\boldsymbol{\tau}\right]\cdot\mathbf{x}\,, \tag{6.11}$$

we may alternatively represent (6.9) as:

$$\begin{aligned}\boldsymbol{\sigma}\cdot\mathbf{x}=\sigma_i x^i&=\begin{pmatrix} z & x-iy \\ x+iy & -z\end{pmatrix}=\begin{pmatrix} \xi_1\xi_2 & -\xi_1{}^2 \\ \xi_2{}^2 & -\xi_1\xi_2\end{pmatrix}=\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}\left(\xi_2\quad -\xi_1\right)=-\xi\xi^\dagger \\ &=\frac{1}{4}i\exp\left(-i6\pi\frac{n}{m}\right)\csc^2\left(\pi\frac{n}{m}\right)\varepsilon_{ijk}\ {}^{n}_{m}\left[\tau_j,\tau_k\right]x^i=\frac{1}{2}i\exp\left(-i6\pi\frac{n}{m}\right)\csc^2\left(\pi\frac{n}{m}\right){}^{n}_{m}\left[\boldsymbol{\tau}\times\boldsymbol{\tau}\right]\cdot\mathbf{x}\end{aligned} \tag{6.12}$$

Of course, the determinant $\left|\boldsymbol{\sigma}\cdot\mathbf{x}\right|=x^2+y^2+z^2=r^2$ is the Pythagorean invariant of rotation under SO(3) transformations, which are equivalent to an SU(2) transformation on the transposed complex spinor doublet $\xi^T=\left(\xi_1,\xi_2\right)^T$. And we have already seen in (5.12) and (6.7) that for the square root m=2 and the non-trivial solution n=1, the covering group ${}^{1}_{2}\tilde{G}(2)=SU(2)$. But we now see that for any higher root $m>2$ there will be an additional projection onto SO(3) which gets routed through $SU(2)={}^{1}_{2}\tilde{G}(2)$, by virtue of (6.6) as exemplified in (6.12). So to update (5.2) and characterize fully what is occurring in (6.9) and (6.12), we now write:

$$\pi:{}^{n}_{m}\tilde{G}(2)\rightarrow{}^{1}_{2}\tilde{G}(2)=SU(2)\rightarrow SO(3):\ker\pi=\sqrt[m]{1}=\exp i\vartheta=\exp\left(i2\pi n/m\right). \tag{6.13}$$

Additionally, (6.6) and (6.12) show us how to embed this result into Dirac theory. $\left(i\gamma^\mu\partial_\mu-m\right)\psi=0$. Using the Dirac representation for γ^μ and with $\partial_\mu=\left(\partial_t,\nabla\right)$, and particularly substituting $\mathbf{x}\rightarrow\nabla$ in (6.12), and thus writing:

$$\boldsymbol{\sigma}\cdot\nabla=\frac{1}{2}i\exp\left(-i6\pi\frac{n}{m}\right)\csc^2\left(\pi\frac{n}{m}\right){}^n_m[\boldsymbol{\tau}\times\boldsymbol{\tau}]\cdot\nabla\,, \tag{6.14}$$

we may represent Dirac's equation to include the ${}^n_m[\boldsymbol{\tau}\times\boldsymbol{\tau}]$ cross product, via

$$\begin{aligned}0&=\left(i\gamma^\sigma\partial_\sigma-m\right)\psi=i\begin{pmatrix}\partial_t-m & \boldsymbol{\sigma}\cdot\nabla\\ -\boldsymbol{\sigma}\cdot\nabla & -\partial_t-m\end{pmatrix}\begin{pmatrix}\psi_A\\ \psi_B\end{pmatrix}\\ &=i\begin{pmatrix}\partial_t-m & 0\\ 0 & -\partial_t-m\end{pmatrix}\begin{pmatrix}\psi_A\\ \psi_B\end{pmatrix}-\frac{1}{2}\exp\left(-i6\pi\frac{n}{m}\right)\csc^2\left(\pi\frac{n}{m}\right)\begin{pmatrix}0 & {}^n_m[\boldsymbol{\tau}\times\boldsymbol{\tau}]\cdot\nabla\\ -{}^n_m[\boldsymbol{\tau}\times\boldsymbol{\tau}]\cdot\nabla & 0\end{pmatrix}\begin{pmatrix}\psi_A\\ \psi_B\end{pmatrix}\end{aligned} .\tag{6.15}$$

The projection for this is now characterized by:

$$\pi:{}^n_m\tilde{G}(2)\to{}^1_2\tilde{G}L(2,C)=SL(2,C)\to SO(1,3):\ker\pi=\sqrt[m]{1}=\exp i\vartheta=\exp\left(i2\pi n/m\right). \tag{6.16}$$

Note that the covering group is still the same ${}^n_m\tilde{G}(2)$. But it is now projecting onto the $SO(1,3)$ of spacetime with rotation and boost, via $SL(2,C)$ which we have equated to what we have denoted as ${}^1_2\tilde{G}L(2,C)$. This is because in (6.15) we have taken ${}^1_2\tilde{G}(2)$ of (6.13), replicated it twice for particle and antiparticle (C), and used that in the usual linear (L) combination with the time-dependent term $\gamma^0\partial_t-m$ of Dirac's equation.

7. Conclusion: Fractional Wu-Yang Charges Projected onto SO(3) from Root-of-Unity Covering Groups

Everything we have found leading to (6.13) and (6.16) appears to provide a well-defined, unambiguous projection for fractional charges onto SO(3) and SO(1,3). Specifically, for any m^{th} root of unity, there will be precisely m different covering groups ${}^n_m\tilde{G}(2)$ with associated generators ${}^n_m\tau_i$ as defined in (5.10), such that $0\le n<m$, which then will recycle starting with n=m. The n=0 solutions are all trivial ${}^0_m\tau_i=I_i$, but the remaining m-1 solutions each define a unique, single-valued, simply-connected covering group ${}^n_m\tilde{G}(2)$. So, for example, for m=5 there will be five ${}^n_5\tau_i$ generators and five associated ${}^n_5\tilde{G}(2)$ each of which is *single valued and simply connected*. When projected onto SO(3) as in (6.13), this will provide a quintuple cover of SO(3). In section 3 this meant that the azimuth domain is $0\le\varphi\le10\pi$ and there could be no fractional charges because of the $m/m=1$ cancellation in (3.8). But all of this was due to SO(3) being a multi-valued representation which is not simply-connected.

Now, as summarized in (6.13), each of these five ${}^n_5\tilde{G}(2)$ will map onto SO(3) via a five-to-one surjective homomorphism, with the exact symmetry carried by each single-valued ${}^n_5\tilde{G}(2)$ with $0\le n<m$, and with SO(3) being inexact to the extent it has an inherent five-valued

ambiguity. Put another way, from the viewpoint of the number 1, there are five distinct fifth roots, and it is impossible to state that one or the other is the actual root from whence 1 was obtained upon rising one of these roots to its fifth power. But from the viewpoint of each of these five roots, we can state very clearly what each root is. But then, when we take each of these five roots and raise them to the fifth power, we will map all five of these roots onto the number 1, which is to say we will quintuple cover 1, and the number 1 will have no way of knowing from which of the five roots it came. Answering the question from whence came the 1, from the viewpoint of 1, will be ambiguous and ill-defined.

But in the context of the many-to-one homomorphic mapping of (6.13), there is no longer any $m/m=1$ cancellation, and we can retain the Wu-Yang fractionalization solution $\vartheta = e\mu = 2\pi n/m$, see (5.3), in a well-defined and unambiguous manner. This means that we can now use $\vartheta/2 = \pi n/m = e\mu/2$ to update (5.10) for the ${}^{n}_{m}\tilde{G}(2)$ generators and write this as:

$$ {}^{n}_{m}\tau_i \equiv \sqrt[m]{I_i} = \exp\left(i3\pi\frac{n}{m}\right)\exp\left(i\sigma_i\frac{\vartheta}{2}\right) = \exp\left(i3\pi\frac{n}{m}\right)\exp\left(i\sigma_i\pi\frac{n}{m}\right) = \exp\left(i\frac{3}{2}e\mu\right)\exp\left(i\frac{1}{2}\sigma_i e\mu\right). \tag{7.1} $$

It is of interest to note that the argument $\frac{1}{2}\sigma_i e\mu$ naturally emerges to contain the half-spin generators $\frac{1}{2}\sigma_i\hbar$ for fermions multiplied by the electric and magnetic charge product $e\mu$.

Likewise, (6.5) and (6.6) which relate the Pauli spin matrices to these fractional charge generators respectively become:

$$ \begin{aligned} {}^{n}_{m}\left[\tau_i,\tau_j\right] &= -2i\exp\left(i3\vartheta\right)\sin^2\left(\frac{\vartheta}{2}\right)\varepsilon_{ijk}\sigma_k = -2i\exp\left(i6\pi\frac{n}{m}\right)\sin^2\left(\pi\frac{n}{m}\right)\varepsilon_{ijk}\sigma_k \\ &= -2i\exp\left(i3e\mu\right)\sin^2\left(\frac{e\mu}{2}\right)\varepsilon_{ijk}\sigma_k \end{aligned}, \tag{7.2} $$

$$ \begin{aligned} \sigma_i &= \frac{1}{4}i\exp\left(-i3\vartheta\right)\csc^2\left(\frac{\vartheta}{2}\right)\varepsilon_{ijk}\,{}^{n}_{m}\left[\tau_j,\tau_k\right] = \frac{1}{4}i\exp\left(-i6\pi\frac{n}{m}\right)\csc^2\left(\pi\frac{n}{m}\right)\varepsilon_{ijk}\,{}^{n}_{m}\left[\tau_j,\tau_k\right] \\ &= \frac{1}{4}i\exp\left(-i3e\mu\right)\csc^2\left(\frac{e\mu}{2}\right)\varepsilon_{ijk}\,{}^{n}_{m}\left[\tau_j,\tau_k\right] \end{aligned}. \tag{7.3} $$

This also means that (6.12) becomes:

$$\begin{aligned}\boldsymbol{\sigma}\cdot\mathbf{x}=\sigma_i x^i &= \begin{pmatrix} z & x-iy \\ x+iy & -z\end{pmatrix} = \begin{pmatrix} \xi_1\xi_2 & -\xi_1^{\ 2} \\ \xi_2^{\ 2} & -\xi_1\xi_2\end{pmatrix} = \begin{pmatrix}\xi_1 \\ \xi_2\end{pmatrix}\begin{pmatrix}\xi_2 & -\xi_1\end{pmatrix} = -\xi\xi^\dagger \\ &= \frac{1}{4} i \exp(-i3\vartheta)\csc^2\left(\frac{\vartheta}{2}\right)\varepsilon_{ijk}\ {}^n_m\left[\tau_j,\tau_k\right]x^i = \frac{1}{2} i\exp(-i3\vartheta)\csc^2\left(\frac{\vartheta}{2}\right){}^n_m\left[\boldsymbol{\tau}\times\boldsymbol{\tau}\right]\cdot\mathbf{x} \\ &= \frac{1}{4} i \exp\left(-i6\pi\frac{n}{m}\right)\csc^2\left(\pi\frac{n}{m}\right)\varepsilon_{ijk}\ {}^n_m\left[\tau_j,\tau_k\right]x^i = \frac{1}{2} i\exp\left(-i6\pi\frac{n}{m}\right)\csc^2\left(\pi\frac{n}{m}\right){}^n_m\left[\boldsymbol{\tau}\times\boldsymbol{\tau}\right]\cdot\mathbf{x} \\ &= \frac{1}{4} i \exp(-i3e\mu)\csc^2\left(\frac{e\mu}{2}\right)\varepsilon_{ijk}\ {}^n_m\left[\tau_j,\tau_k\right]x^i = \frac{1}{2} i\exp(-i3e\mu)\csc^2\left(\frac{e\mu}{2}\right){}^n_m\left[\boldsymbol{\tau}\times\boldsymbol{\tau}\right]\cdot\mathbf{x}\end{aligned} \quad ,(7.4)$$

which can readily be adapted for use in the Dirac equation as in (6.15).

Now, it is important to note that $\vartheta/2=\pi n/m$ as used in (7.1) through (7.4) has *nothing to do with the Wu-Yang fractionalization. This is entirely about the geometry of* ${}^n_m\tilde{G}(2)$ and this is all built up by identity from the Euler relationship $\exp i\vartheta=\exp\left(i2\pi n/m\right)$ of (4.1). Then, *independently*, the Wu-Yang differential equation $e^{-i\Lambda}de^{i\Lambda}=ie\mu d\not{\varphi}$ is solved by $\exp\left(i\Lambda\right)=\exp\left(ie\mu\not{\varphi}\right)$ which in turn has the solution $2\pi n=e\mu\not{\varphi}$ with $\not{\varphi}=m=0,1,2,3...$, which may therefore be written as $e\mu=2\pi n/m$. There is nothing which *a priori* eliminates this as a solution and requires us to set $m=1$, or, more precisely, to set $n=mk$ such that $e\mu=2\pi n/m=2\pi mk/m=2\pi k$, where k is itself an integer. This only happens at (3.8) when we consider Dirac strings in SO(3) which strings are fictive, and more importantly, which SO(3) is multivalued and so only represents an approximate, not an exact symmetry. That is, like the number 1, SO(3) has no idea from which roots it has been projected.

But when we instead turn to m^{th} root geometric spaces with generators defined by ${}^n_m\tau_i\equiv\sqrt[m]{I_i}$ as the m^{th} roots of 2x2 identity matrices, we find that there are a whole host of generator relationships and homomorphic projections which contain the fraction $\vartheta/2=\pi n/m$ based not on Wu-Yang, but based on the Euler relation $\exp i\vartheta=\exp\left(i2\pi n/m\right)$. In the geometries with covering groups ${}^n_m\tilde{G}(2)$ we start at the roots and then raise them to the m^{th} power to return to the identity matrix. From the view of ${}^n_m\tilde{G}(2)$ everything is well-defined and single-valued, and there is nothing which requires m to be limited 1 or requires us to set $n=mk$ as we were required to do at (3.8). The geometry contains all of these roots, well-defined, naturally and unambiguously. As a result, we are enabled to connect the Wu-Yang solution to the Euler relation via $\vartheta=e\mu=2\pi n/m$. Once this is done, (7.1) through (7.4) provide a well-defined picture of how Wu-Yang fractional charges can and do find accommodation in SO(3) and SO(1,3), and we find that the product $e\mu$ of the electric and magnetic charge strengths is identical to the Euler angle, $\vartheta=e\mu$. In all cases, ${}^n_m\tilde{G}(2)$ and ${}^n_m\tau_i$ then come to represent the covering group and generators which project a $\nu=n/m$ fractional charge onto SO(3).

If we take the determinant of (7.4) and boil this all down to its essence, we obtain:

$$r^2 = x^2 + y^2 + z^2 = \left|\boldsymbol{\sigma}\cdot\mathbf{x}\right| = \left|\begin{pmatrix} z & x-iy \\ x+iy & -z \end{pmatrix}\right| = \left|\frac{1}{2}i\exp\left(-i3e\mu\right)\csc^2\left(\frac{e\mu}{2}\right){}_m^n\left[\boldsymbol{\tau}\times\boldsymbol{\tau}\right]\cdot\mathbf{x}\right|. \tag{7.5}$$

This supersedes the result in (3.8) which precluded fractional charges when the multi-valued SO(3) was considered alone without projection form any single-valued covering groups. But this relationship (7.5) is now well-defined and unambiguous and it preserves lengths under SO(3) rotations. If this relationship is true, then the mathematics of the Euler relationship $\exp i\vartheta = \exp\left(i2\pi n/m\right)$ will necessarily produce and permit well-defined fractional charges with $\vartheta = e\mu = 2\pi n/m$.

References

[1] Dirac, P.A.M., *Quantized Singularities in the Electromagnetic Field*, Proceedings of the Royal Society A 133 (821): 60–72, (September 1931)

[2] http://www.theory.caltech.edu/~preskill/pubs/preskill-1984-monopoles.pdf

[3] 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

[4] Wu, T. T. and Yang, C. N., *Dirac Monopole without Strings: Classical Lagrangian theory*, Phys. Rev. D 14 (1976) 437

[5] *PQM Supplementary Notes: Spin, topology, SU(2) → SO(3) etc*, http://www.damtp.cam.ac.uk/user/examples/D18S.pdf (2014)

[6] Dirac, P. A. M., *The Quantum Theory of the Electron*, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 117 (778): 610 (1928)

[7] P.A.M. Dirac, *The Quantum Theory of the Electron. Part II*, Proc. Roy. Soc. Lon. A118, 351 (1928)

[8] Minkowski, H., *Space and Time*, Jahresberichte der Deutschen Mathematiker-Vereinigung 10: 75–88 (1909)

[9] A. Einstein, *The Foundation of the General Theory of Relativity*, Annalen der Physik (ser. 4), 49, 769–822 (1916)