Filling the Mass Gap and Unifying Classical Gauge Theory with Gravitation: Chromodynamic Symmetries, Confinement Properties, and Short-Range Interactions of Yang-Mills Gauge Theory

Jay R. Yablon
Schenectady, New York
jyablon@nycap.rr.com

August 9, 2013 Draft

Abstract: This is the second partial draft of a paper under development to further elaborate the author's thesis presented in several earlier-published papers, that baryons including protons and neutrons are Yang-Mills magnetic monopoles. This paper fully develops the non-linear aspects of Yang-Mills gauge theory and applies these to the inverses used to populate the Yang-Mills magnetic monopolies with quarks and turn them into baryons and give rise to QCD. We also show how the perturbations in these inverses, which arise from the non-linear theory, are responsible for the short-range of the nuclear interaction, notwithstanding the zero-mass gluon gauge fields. This solves the mass gap problem and demonstrates how strong interactions may have a short range notwithstanding their massless gluon gauge fields. Additionally, sections 7 and 8 develop a classical field equation which fully unifies gauge theory with gravitational theory.

Contents

1. Introduction

Quantum Chromodynamics (QCD) is a highly-successful theory of strong interactions. Yet despite its success, QCD has a number of limitations which owe less to any endemic problems with QCD, than to our own human inability to heretofore perform exact analytical calculations which fully and completely take advantage of the non-linear gauge field interactions of Yang-Mills theory. Specifically, there are several distinct shortcomings – not of QCD itself but in our present understanding of QCD and Yang-Mills gauge theory and how to calculate with these – which must be fairly acknowledged and remedied.

First, while it is abundantly clear that baryons including protons and neutrons comprise exactly three quarks, QCD, which is premised upon the Yang-Mills gauge group $SU(3)_C$, does not explain the "three-ness" of these nucleons and other baryons. It essentially postulates (with solid support from empirical evidence) three quarks per baryon without explaining on a theoretical footing *why* this number must be 3 rather than 4 or 7 or 11 or some other number, and thus, why SU(3) rather than SU(4) or SU(7) or SU(11) or some other group is the gauge group that works to accurately reproduce the empirical evidence.

Second, as Jaffe and Witten point out on page 3 of [1], it has not yet proven possible to use QCD 1) "to explain why the nuclear force is strong but shortranged," 2) to explain "'quark confinement,' that is, even though the theory is described in terms of elementary fields, such as the quark fields, that transform non-trivially under SU(3), the physical particle states – such as the proton, neutron, and pion – are SU(3)-invariant," and 3) to explain "'chiral symmetry breaking'." Jaffe and Witten continue (emphasis added):

> "Both experiment – since QCD has numerous successes in confrontation with experiment – and computer simulations . . . have given strong encouragement that QCD does have the properties [short range, confinement and chiral symmetry breaking] cited above. These properties can be seen, to some extent, in theoretical calculations carried out in a variety of highly oversimplified models (like strongly coupled lattice gauge theory, see, for example, [3]). *But they are not fully understood theoretically; there does not exist a convincing, whether or not mathematically complete, theoretical computation demonstrating any of the three properties in QCD, as opposed to a severely simplified truncation of it.*"

It remains vital to demonstrate a convincing theoretical basis for nuclear short range notwithstanding zero-mass gauge fields (gluons), quark and gauge field confinement, and chiral symmetry breaking, without truncation or oversimplification.

Third, as Zee points out in section VII.1 of [2], present methods used to calculate in Yang-Mills theory, such as perturbation theory or lattice gauge theory, are among the severely truncated methods used at present which must eventually be replaced by more complete and exact ways of doing *analytical* (as opposed to numerical) calculations with Yang-Mills theory. Perturbation theory in Zee's description is "an unnatural act as it involves brutally splitting [the Lagrangian density] $\mathcal{L}$ into two parts: a part quadratic in the fields and the rest," while lattice gauge theory [3], in contrast, "does violence to Lorentz invariance rather than to gauge invariance." This is not an adverse reflection on Yang-Mills or QCD, but only on our ability to calculate with them, analytically. Better methods and approaches are needed which does violence to neither.

Fourth and finally, as pointed out in [4]:

> "the European Muon Scattering Collaboration showed that the quark structure of the nucleus was different than the quark structure of the nuclei. And that offered a completely different perspective. I think of it as a paradigm shift. Some people think that's too strong. But at that time, we had major problems in nuclear physics. There were fundamental aspects of the nucleus that we were calculating and we were getting the wrong answer. And if you say that the quark structure of a proton changes inside a nucleus, then our conventional way of doing nuclear calculations might, in fact, not be right. And so these QCD effects then could explain why we have trouble getting the binding energy of helium-4, the binding energy in nuclear matter, or problems with other nuclear structure problems."

So while we have to date been able to use our present understanding of QCD to explain many things about strong interactions, there is still a disconnect between QCD as a *particle* theory and QCD as a *nuclear* theory. We are able with strong supporting evidence to model baryons and nucleons as systems of three quarks but we do not know *a priori* as a theoretical matter *why* these should be based upon three quarks. We develop QCD using an $SU(3)_C$ color group to represent a color triplet of quarks transforming non-trivially under SU(3) (particle physics) but do not presently understand how and why "the physical particle states [of nuclear physics] – such as the proton, neutron, and pion [more generally mesons] – are SU(3)-invariant." Our methods of calculating, such as perturbation and lattice gauge theory, are severely truncated, at the cost of either gauge or Lorentz symmetry. And as a result of all of this, we are not yet able to make use of QCD to connect the elementary particles – quarks – to the observed physical particle states – protons and neutrons and other baryons, as well as mesons – in such a way as to even calculate such long-known but not-yet-explained things as nuclear binding energies and mass defects, not to mention proton and neutron and other hadron masses themselves.

Two additional points need to be made in order to avoid misunderstanding of the scope and intent of this paper and more generally of the author's research in this field. First, although lattice gauge theory [3] is the method most-commonly employed to do calculations in QCD, it is not the author's goal or intent to use lattice gauge theory in any way in this paper or in his research. Lattice gauge theory is an approximation scheme devised precisely because it was found to be troublesome to do exact calculations with Yang-Mills theory. Essentially, one takes the infinitesimal differential element *dx* and approximates continuum spacetime to a finitely-spaced lattice so as to reduce "the functional integration to a finite-dimensional integral. One must then verify the existence of limits of appropriate expectations of gauge-invariant observables as the lattice spacing tends to zero and as the volume tends to infinity." [1] at page 11. The author's goal is to do direct, exact analytical calculations using Yang-Mills theory while preserving both gauge symmetry and Lorentz symmetry without any truncation or approximation. Thus, there will be no consideration of particular ways to normalize or regularize a lattice gauge calculation, because the whole intent of this work is to go beyond having to take the trouble to calculate in such an analytically-limited, albeit effective, manner.

Second, Jaffe and Witten in [1] at pages 1-2 point out that for non-Abelian (non-commuting) gauge theory:

> "At the classical level one replaces the gauge group U(1) of electromagnetism by a compact gauge group G. The definition of the curvature arising from the connection must be modified to $F = dA + gA \wedge A$, and Maxwell's equations are replaced by the Yang–Mills equations, $0 = d_A F = d_A {}^*F$, where d_A is the gauge-covariant extension of the exterior derivative."

They later proceed to survey a wide variety of methods used "to show the existence of quantum fields on non-compact configuration space" and specifically to demonstrate that "relativistic, nonlinear quantum field theories exist." On page 12 of [1], they then observe that (embedded references renumbered here):

> "One view of the mass gap in Yang–Mills theory suggests that it could arise from the quartic potential $(A \wedge A)^2$ in the action, where $F = dA + gA \wedge A$, see [5], and may be tied to curvature in the space of connections, see [6]."

This is the view upon which the author has based his work, in particular, because it is in accord Occam's razor as restated by Einstein [7], that "the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience." In the specific context under consideration here, all of the other methods enumerated in section 6 of [1] appear to entail supplementing pure Yang-Mills theory with other devices or suppositions or making truncated approximations in order to be able to explain a nuclear short range coincident with massless gauge fields, quark and gauge field confinement, and chiral symmetry breaking. *If*, however, these can be *fully* explained using no more than the Yang-Mills fields strength $\mathrm{F} = \mathrm{dA} + \mathrm{gA} \wedge \mathrm{A}$ via the quartic action terms $(\mathrm{A} \wedge \mathrm{A})^2$ (and perhaps also the cubic terms), then this would place the mass gap and confinement and chiral solutions *entirely* on the shoulders of Yang-Mills theory without any supplement, and this would undoubtedly be the simplest view one can take. Furthermore, because the classical Yang-Mills equations are simply those of Maxwell extended into the non-Abelian domain, this would entirely explain nuclear short range, quark and gauge field confinement, and chiral symmetry breaking on the basis of Maxwell's theory in non-Abelian form. *This would then reveal Maxwell's theory with non-commuting gauge fields and nothing more, to be the governing theory of nuclear physics.* A simpler result – so long as it is not too simple to explain the datum of experience – can scarcely be imagined.

2. Thesis and Methodology

In a recent paper [8], the author presented the thesis that the non-vanishing magnetic monopoles of Yang-Mills theory are in fact synonymous with baryons. That is, magnetic monopoles, long-pursued since the time of James Clerk Maxwell have, in Yang-Mills incarnation, always been hiding in plain sight as baryons, and most importantly, as the protons and neutrons which rest at the center of the material universe.

More to be added.

3. Classical Yang-Mills Theory: Three Equivalent Viewpoints

Yang-Mills gauge theories, developed in 1954 [9], rest mathematically upon the generalization of the 2x2 Pauli matrices of SU(2) into SU(N) matrices of any NxN dimensionality. These Pauli matrices in turn embody the quaternions developed in 1843 by William Rowan Hamilton, famously carved into the Brougham Bridge in Dublin, Ireland. Normalized such that $Tr\left(\lambda^i\lambda^j\right)=\frac{1}{2}\delta^{ij}$, the N^2-1 generators $\lambda^i; i=1,2,3...N^2-1$ of any Yang-Mills gauge group SU(N) maintain the commutator relationship $\left[\lambda_i,\lambda_j\right]=if_{ijk}\lambda_k$, where f_{ijk} are the group structure constants. This generalizes the Pauli relationship which is $\left[\sigma_i,\sigma_j\right]=i\varepsilon_{ijk}\sigma_k$ for the normalization $Tr\left(\sigma^i\sigma^j\right)=\frac{1}{2}\delta^{ij}$. Each generator is an NxN matrix and so can be written $\lambda^i{}_{AB}; A,B=1,2,3...N$, but in general it is simpler to suppress these A,B indexes and simply keep in mind at all times that these indexes are implicitly there.

Physically, an SU(N) gauge theory extending Maxwell's electrodynamics into non-Abelian domains is developed from these generators in the following way: first, one posits a set of N^2-1 vector potentials (gauge fields) $G^{i\mu};\ i=1,2,3...N^2-1$. Next, one sums these with the generators to form $G^\mu{}_{AB}\equiv\lambda^i{}_{AB}G^{i\mu}$ which with A,B indexes implicit is normally written as $G^\mu\equiv\lambda^i G^{i\mu}$. This is an NxN matrix of spacetime 4-vector potentials. Similarly, one forms a set of N^2-1 field strength tensors $F^{i\mu\nu}$, each of which is a bivector with a "chromo-electric" field $\mathbf{E_i}$ and a chromo-magnetic field $\mathbf{B_i}$. We then use these to form $F^{\mu\nu}_{AB}\equiv\lambda^i_{AB}F^{i\mu\nu}$ which is an NxN Yang-Mills matrix of 4x4 antisymmetric second rank tensor bivectors. Finally, *in very important contrast to the electrodynamic field strength* $F^{\mu\nu}=\partial^\mu G^\nu-\partial^\nu G^\mu$, we specify the NxN field strength matrix $F^{\mu\nu}$ in terms of the NxN gauge field matrix G^μ as (see, e.g., [2], equation IV.5(16)):

$$F^{\mu\nu}=\partial^\mu G^\nu-\partial^\nu G^\mu-i\left[G^\mu,G^\nu\right]=\partial^{[\mu}G^{\nu]}-i\left[G^\mu,G^\nu\right] \tag{3.1}$$

This commutator $\left[G^\mu,G^\nu\right]$ is non-vanishing, $\left[G^\mu,G^\nu\right]\neq 0$. *Everything that differentiates Yang-Mills gauge theory from an Abelian gauge theory such as QED, originates solely and exclusively from the fact that these gauge field / vector potential matrices* $G^\mu\equiv\lambda^i G^{i\mu}$ *do not commute, i.e., from the fact that* $\left[G^\mu,G^\nu\right]\neq 0$.

Starting with (3.1), there are several different, fully equivalent ways in which one can think about Yang-Mills gauge theories. Because of the difficulties surveyed in the introduction that have been encountered to date doing calculations with Yang-Mills theory, the way one chooses to think about Yang-Mills, depending on circumstance, can make a big difference in whether a calculation or conceptualization is reasonably clean and simple, or messy and obtuse. *The first way to think about Yang-Mills is that of (3.1), as a theory in which the gauge fields do not commute.* As we shall review momentarily, this leads very directly to non-vanishing magnetic monopole source charges that will be central to the development here.

For a second way to think about Yang-Mills, it is worth being reminded how to expand out (3.1) using $F^{\mu\nu} = \lambda^i F^{i\mu\nu}$, $G^\mu = \lambda^i G^{i\mu}$ and $\left[\lambda_i, \lambda_j\right] = if_{ijk}\lambda_k$. We find while renaming summed indexes as needed that:

$$\begin{aligned}\lambda^i F^{i\mu\nu} &= \partial^\mu \lambda^i G^{i\nu} - \partial^\nu \lambda^i G^{i\mu} - i\left[\lambda^i G^{i\mu}, \lambda^j G^{j\nu}\right] \\ &= \lambda^i \partial^\mu G^{i\nu} - \lambda^i \partial^\nu G^{i\mu} - i\left[\lambda^i, \lambda^j\right] G^{i\mu} G^{j\nu} \\ &= \lambda^i \partial^\mu G^{i\nu} - \lambda^i \partial^\nu G^{i\mu} + f^{kji} \lambda^i G^{k\mu} G^{j\nu}\end{aligned} \tag{3.2}$$

The λ^i is then factored out from all terms, leaving, after more renaming, the perhaps more-familiar expression:

$$F^{i\mu\nu} = \partial^\mu G^{i\nu} - \partial^\nu G^{i\mu} + f^{ijk} G^{j\mu} G^{k\nu} = \partial^{[\mu} G^{i\nu]} + f^{ijk} G^{j\mu} G^{k\nu} \tag{3.3}$$

If we now use (3.3) to form a Lagrangian density akin to the pure field terms in (2.21), we obtain the also familiar:

$$\begin{aligned}\mathfrak{L} = -\tfrac{1}{4} F^{i\mu\nu} F_{i\mu\nu} &= -\tfrac{1}{4}\left(\partial^{[\mu} G^{i\nu]} + f^{ijk} G^{j\mu} G^{k\nu}\right)\left(\partial_{[\mu} G_{i\nu]} + f_{ilm} G_{l\mu} G_{m\nu}\right) \\ &= -\tfrac{1}{4}\partial^{[\mu} G^{i\nu]} \partial_{[\mu} G_{i\nu]} - \tfrac{1}{2} f_{ijk} \partial^{[\mu} G^{i\nu]} G_{j\mu} G_{k\nu} - \tfrac{1}{4} f^{ijk} f_{ilm} G^{j\mu} G^{k\nu} G_{l\mu} G_{m\nu}\end{aligned} \tag{3.4}$$

The first term, $-\tfrac{1}{4}\partial^{[\mu} G^{i\nu]} \partial_{[\mu} G_{i\nu]}$, a "harmonic oscillator" term, is quadratic in the gauge fields, and is fully analogous and indeed identical in form to the term $-\tfrac{1}{4} F^{\mu\nu} F_{\mu\nu} = -\tfrac{1}{4}\partial^{[\mu} G^{\nu]} \partial_{[\mu} G_{\nu]}$ in the Lagrangian density of electrodynamics. But the remaining terms $-\tfrac{1}{2} f_{ijk} \partial^{[\mu} G^{i\nu]} G_{j\mu} G_{k\nu}$ and $-\tfrac{1}{4} f^{ijk} f_{ilm} G^{j\mu} G^{k\nu} G_{l\mu} G_{m\nu}$, the "perturbation" terms, represent vertices with three and four interacting gauge fields. This is unprecedented in electrodynamics, and makes Yang-Mills a *non-linear* theory. *So the second way to think about Yang-Mills theory is that of (3.4), in which the gauge fields do <u>not</u> act like photons and forego interactions one another like ships passing in the night. Rather, the Yang-Mills gauge fields fully interact with one another as well as with their fermion (current) sources.* Unfortunately, doing exact calculations with (3.4) is difficult, and in general we will find it unhelpful to split (3.4) into harmonic and perturbative parts as is done in perturbative gauge theory, or to spoil the Lorentz invariance as in lattice gauge theory, again, see the discussion in the introduction. Another approach is needed.

A third way to think about Yang-Mills gauge theory is to expand the commutator in (3.1) and then reconsolidate using gauge covariant derivatives $D^\mu \equiv \partial^\mu - iG^\mu$, as such: (In general we shall scale the interaction charge strength g into the gauge field via $gG^\mu \to G^\mu$. This g can always be extracted back out when explicitly needed.):

$$F^{\mu\nu} = \partial^\mu G^\nu - \partial^\nu G^\mu - iG^\mu G^\nu + iG^\nu G^\mu = \left(\partial^\mu - iG^\mu\right)G^\nu - \left(\partial^\nu - iG^\nu\right)G^\mu = D^\mu G^\nu - D^\nu G^\mu = D^{[\mu}G^{\nu]} \quad (3.5)$$

We compare $F^{\mu\nu} = D^{[\mu}G^{\nu]}$ above to the Abelian field strength $F^{\mu\nu} = \partial^{[\mu}G^{\nu]}$ and see that the *only* difference is that the ordinary derivative is replaced by $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$. *This is actually a very pedagogically-useful observation*: Consider that gauge theory first originates when one has a field equation or a Lagrangian for a scalar ϕ or fermion ψ field which includes a term $\partial_\mu \phi$ or $\partial_\mu \psi$. One then subjects the field to the *local* gauge (phase) transformation $\phi \to e^{i\,\theta(x)}\phi$ or $\psi \to e^{i\,\theta(x)}\psi$ and insists that the field equation or Lagrangian remain invariant under this transformation. What does one do to ensure such invariance? Make the replacement $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$. So now, one changes $\partial_\mu \phi \to D_\mu \phi$ and $\partial_\mu \psi \to D_\mu \psi$ with the consequence that ϕ or ψ now acquires an interaction with the gauge field G^μ.

So if we start with an Abelian gauge theory such as QED for which $F^{\mu\nu} = \partial^{[\mu}G^{\nu]}$, we can easily turn it into a non-Abelian gauge theory by replacing $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$ so that $F^{\mu\nu} = D^{[\mu}G^{\nu]}$,which is (3.5). As a consequence, the gauge field G^ν acquires an interaction with the gauge field G^μ, i.e., the gauge field now starts to interact non-linearly with itself! This says exactly the same thing as (3.4), with the exception that in the form of (3.5), the pure gauge term in the Lagrangian is the much cleaner (the ½ rather than ¼ owes to the $Tr\left(\lambda^i \lambda^j\right) = \frac{1}{2}\delta^{ij}$ normalization):

$$\mathfrak{L} = -\tfrac{1}{2}\,\mathrm{Tr}F^{\mu\nu}F_{\mu\nu} = -\tfrac{1}{2}\,\mathrm{Tr}D^{[\mu}G^{\nu]}D_{[\mu}G_{\nu]}\,. \quad (3.6)$$

Given that (3.4) and (3.6) state *exactly the same physics*, it should be clear that (3.6) is a much easier expression to work with than (3.4). *This is a third way to think about Yang-Mills theories: A non-Abelian gauge theory is simply an Abelian gauge theory for which gauge theory has been applied to gauge theory. Or, perhaps with a bit more color (pun intended), Yang-Mills gauge theory is <u>gauge theory on steroids</u>.*

Specifically, in gravitational theory, the principle of minimal coupling suggests that we merely replace the ordinary derivatives $\partial_\mu G^\nu$ of a vector G^ν with covariant derivatives $\partial_{;\mu} G^\nu \equiv \partial_\mu G^\nu + \Gamma^\nu_{\mu\sigma}G^\sigma$ simultaneously with replacing the Minkowski metric tensor $\eta_{\mu\nu}$ with the generalized metric tensor $g_{\mu\nu}$ for the gravitational field, to migrate from a flat spacetime to curved one in which $\Gamma^\nu_{\mu\sigma}G^\sigma$ represents the curvature discerned under parallel transport. (See, e.g., [10] page 259) In gauge theory, this steroidal replacement of $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$ represents an analogous principle of minimal coupling for gauge theory, in which the $-iG^\mu$ represents the gauge (really, phase) curvature based on a relative comparison of non-observable phases.

We shall find that for chromo-magnetic sources $P^{\sigma\mu\nu}$ with a classical field equation $P^{\sigma\mu\nu} = \partial^\sigma F^{\mu\nu} + \partial^\mu F^{\nu\sigma} + \partial^\nu F^{\sigma\mu}$, it is most helpful to view Yang-Mills theory in the form of (3.1), as a theory on which the gauge field does not self-commute, that is, to think about the "non-Abelian" view of Yang-Mills theory. But, when it comes to chromo-electric sources with the classical equation $J^\nu = \partial_\mu F^{\mu\nu}$, the more convenient view is that of (3.6), in which we view Yang-Mills as gauge theory on steroids. In the notations to be used here, the gauge-covariant derivatives d_A referenced on page 2 of [1] will be denoted simply with the uppercase D, and when written in vector form will be represented as $D^\mu \equiv \partial^\mu - iG^\mu$.

4. The Field Equations and Configuration Space Operator of Classical Yang-Mills Theory

Now we turn to Yang-Mills theory at the level of the classical field equations $0 = d_A F = d_A{}^*F$ discussed on pages 1 and 2 of [1]. Using D rather than d_A, these are written *in vacuo* as $0 = DF = D^*F$. And, for non-vanishing electric and magnetic sources J (one-form) and P (three-form), these are respectively written as $^*J=D^*F$ and $P=DF$. Expanded into tensor notation, these classical Yang-Mills equations, with sources, are:

$$J^\nu = D_{;\mu} F^{\mu\nu}, \tag{4.1}$$

$$P^{\sigma\mu\nu} = D^{;\sigma} F^{\mu\nu} + D^{;\mu} F^{\nu\sigma} + D^{;\nu} F^{\sigma\mu} \equiv D^{;(\sigma} F^{\mu\nu)} = \partial^{;(\sigma} F^{\mu\nu)} - iG^{(\sigma} F^{\mu\nu)}. \tag{4.2}$$

In (4.2), we have also defined a "cyclator" notation $(\sigma\mu\nu)$ to represent the cycling of three indexes over three terms, as shown, which will be useful for compacting the somewhat lengthy expressions we shall soon be deriving for $P^{\sigma\mu\nu}$, and we have also regarded the spacetime as curved and so have included the gravitationally-covariant derivatives $\partial_{;\mu} G^\nu \equiv \partial_\mu G^\nu + \Gamma^\nu_{\mu\sigma} G^\sigma$. Here in (4.1) and (4.2) too, we see a "steroidal" minimal coupling in which the spacetime derivatives of the classical Maxwell equations are replaced with gauge-covariant derivatives $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu \to D^{;u} = \partial^{;\mu} - iG^\mu$ where we also apply the minimal coupling principle from gravitational theory $\partial_\mu G^\nu \to \partial_{;\mu} G^\nu \equiv \partial_\mu G^\nu + \Gamma^\nu_{\mu\sigma} G^\sigma$ as reviewed in the previous section.

As a first step, taking the "gauge theory on steroids" view of Yang-Mills, now employing spacetime-covariant derivatives, we substitute the field strength represented as $F^{\mu\nu} = D^{;[\mu} G^{\nu]}$ from (3.5) into (4.1), while taking the "gauge fields that do not commute" view of Yang-Mills, we substitute the entirely equivalent $F^{\mu\nu} = \partial^{;[\mu} G^{\nu]} - i\left[G^\mu, G^\nu\right]$ of (3.1) into (4.2).

For (4.1), using $D^{;\mu} \equiv \partial^{;\mu} - iG^\mu$ and some well-known index gymnastics, we obtain:

$$\begin{aligned} J^\nu &= D_{;\mu} F^{\mu\nu} = D_{;\mu} D^{;[\mu} G^{\nu]} = D_{;\mu} D^{;\mu} G^\nu - D_{;\mu} D^{;\nu} G^\mu = \left(g^{\mu\nu} D_{;\sigma} D^{;\sigma} - D^{;\mu} D^{;\nu} \right) G_\mu \\ &\overset{+m^2}{\Rightarrow} \left(g^{\mu\nu} \left(D_{;\sigma} D^{;\sigma} + m^2 \right) - D^{;\mu} D^{;\nu} \right) G_\mu \end{aligned}. \tag{4.3}$$

In the final line, we introduce a "Proca mass" *m* for the gauge field, by hand, in the usual way, using $\partial_\sigma \partial^\sigma \to \partial_\sigma \partial^\sigma + m^2$. The Proca mass serves three purposes. First, in circumstances where one is *not* concerned with gauge symmetry and renormalizability and simply wants to know the effect of mass *m* on the field equation (4.3), this tells us what that effect will be. Second, for circumstances where one *is* concerned with preserving gauge symmetry, and wants to be able to generate masses from a Lagrangian with gauge symmetry via spontaneous symmetry breaking, the Proca mass *m* operates as a "red flag" to tell us which masses we want to be able to introduce not by hand, but by symmetry breaking. In other words, terms with Proca masses eventually need to be zeroed out and replaced with mass terms hidden in the gauge symmetry, in more complete theories. This will be very important for the filling mass gap in section 9, where we shall eventually set this mass to zero and show how even with this mass going to zero there will be non-zero gauge boson mass eigenstates remaining behind in the Yang-Mills inverses. Third, with $m = 0$, the configuration space operator of electrodynamics, $g^{\mu\nu}\partial_\sigma \partial^\sigma - \partial^\mu \partial^\nu$ in flat spacetime, has no inverse, which requires gauge fixing, see, e.g., [2], chapter III.4. But $g^{\mu\nu}\left(\partial_\sigma \partial^\sigma + m^2\right) - \partial^\mu \partial^\nu$ with the Proca mass is easily invertible as we shall see in section 9.

The above (4.3) should be contrasted to $J^\nu = \left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\mu}\partial^{;\nu}\right)G_\mu$, which is the analogous classical equation for Maxwell's electrodynamics, in curved as well as flat spacetime because we are including the spacetime-covariant derivatives. We see the gauge theory "minimal coupling principle" at work here: each ordinary spacetime-covariant derivative $\partial_{;\sigma}$ is replaced by the steroidal $D_{;\sigma}$ which is covariant in both spacetime and in the gauge (phase) space. The configuration space operator in (4.3) is $g^{\mu\nu}\left(D_{;\sigma}D^{;\sigma} + m^2\right) - D^{;\mu}D^{;\nu}$, in contrast to the analogous electrodynamic operator $g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\mu}\partial^{;\nu}$. These operators will play an important role in the development here, and in section 9 we shall be obtaining their inverses.

For (4.2), it will help to first review how the monopole density (4.2) behaves in an Abelian gauge theory for which the field strength is simply $F^{\mu\nu} = \partial^{;[\mu}G^{\nu]}$. In doing so, we keep in mind that the Riemann curvature tensor $R^\sigma{}_{\alpha\mu\nu}$ maybe *defined* via $\left[\partial_{;\mu}, \partial_{;\nu}\right]G_\alpha \equiv R^\sigma{}_{\alpha\mu\nu}G_\sigma$ as a direct measure of the degree to which spacetime derivatives are non-commuting. This can be explicitly expanded to show the Christoffel symbols via the expression $\partial_{;\mu}G^\nu = \partial_\mu G^\nu + \Gamma^\nu_{\mu\sigma}G^\sigma$ for the covariant (;) derivative of a vector field. We also keep in mind that one of the important geometric identities satisfied by the Riemann tensor is the first Bianchi identity $R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma} = 0$, with a cycling of indexes identical to that which obtains in the magnetic monopole field equation (4.2). Writing (4.2) in the Abelian form $P^{\sigma\mu\nu} = \partial^{;\sigma}F^{\mu\nu} + \partial^{;\mu}F^{\nu\sigma} + \partial^{;\nu}F^{\sigma\mu}$ and combining with the Abelian field strength $F^{\mu\nu} = \partial^{;[\mu}G^{\nu]}$, this well-known calculation is as follows:

$$\begin{aligned} P^{\sigma\mu\nu} &= \partial^{;\sigma} F^{\mu\nu} + \partial^{;\mu} F^{\nu\sigma} + \partial^{;\nu} F^{\sigma\mu} \\ &= \partial^{;\sigma}\left(\partial^{;\mu} G^{\nu} - \partial^{;\nu} G^{\mu}\right) + \partial^{;\mu}\left(\partial^{;\nu} G^{\sigma} - \partial^{;\sigma} G^{;\nu}\right) + \partial^{;\nu}\left(\partial^{;\sigma} G^{\mu} - \partial^{;\mu} G^{\sigma}\right) \\ &= \left[\partial^{;\sigma}, \partial^{;\mu}\right] G^{\nu} + \left[\partial^{;\mu}, \partial^{;\nu}\right] G^{\sigma} + \left[\partial^{;\nu}, \partial^{;\sigma}\right] G^{\mu} \\ &= \left(R_{\tau}{}^{\nu\sigma\mu} + R_{\tau}{}^{\sigma\mu\nu} + R_{\tau}{}^{\mu\nu\sigma}\right) G^{\tau} = \mathbf{0} \end{aligned} . \tag{4.4}$$

This is a very important result, because it tells us that the vanishing of magnetic monopoles in Maxwell's theory (and to be discussed later, the confinement of quarks in QCD, see Section 1 of [8]) is brought about not only via the trivial relationship $\left[\partial^{\mu}, \partial^{\nu}\right] = 0$ for the commuting of derivatives in flat spacetime, but even in curved spacetime, *by the very nature of the spacetime geometry itself*. That is, the non-existence of magnetic monopoles in Maxwell's electrodynamics is a direct consequence of spacetime geometry, such that $P^{\sigma\mu\nu} = 0$ is a geometrically-rooted relationship. In the language of "differential forms," (4.4) for $P^{\sigma\mu\nu} = 0$ is expressed compactly as $P = dF = ddG = 0$, and is discussed in geometric terms by saying that "the exterior derivative of an exterior derivative is zero," $dd = 0$, see, e.g., [11] §4.6.

It will also be of interest here to consider the monopole equation (4.4) and its non-Abelian counterparts in integral form. Differential forms provide a very helpful way to take volume and surface integrals while easily applying Gauss' / Stokes theorem, which theorem we write generally for any differential form *X*, as $\iint dX = \oint X$. Specifically, to express in integral form the absence of magnetic monopole densities specified in (4.4), one writes $P = dF = ddG = 0$ as: (wedge products $\wedge$ in $\frac{1}{2!} F^{\mu\nu} dx_{\mu} \wedge dx_{\nu} = F^{\mu\nu} dx_{\mu} dx_{\nu}$ are considered to already have been summed)

$$\iiint P = \iiint dF = \iiint ddG = \oiint F = \oiint F^{\mu\nu} dx_{\mu} dx_{\nu} = \oiint dG = 0 . \tag{4.5}$$

One may extract Maxwell's magnetic charge equation in integral form, $\oiint \vec{B} \cdot d\vec{A} = 0$, from the space-space *ij* bivector components of $\oiint F^{\mu\nu} dx_{\mu} dx_{\nu} = 0$. While magnetic fields may flow across some surfaces, there is never a *net* flux of a magnetic field through any *closed* two dimensional surface. In non-Abelian theory, this will tell us that there is no net color passing through any closed two dimensional surface surrounding a Yang-Mills monopole, and will thus be at the root of how quarks and gluons become confined. Faraday's inductive law $\oint \vec{E} \cdot d\vec{l} = -\iint (\partial \vec{B} / \partial t) \cdot d\vec{A}$ is extracted from the time-space 0*k* bivector components. While magnetic fields are often referred to as dipole fields, it is probably better to think of them as *aterminal* fields, i.e., as fields for which the field lines never end at any terminal locale.

We now turn back to the non-Abelian $F^{\mu\nu} = \partial^{;[\mu} G^{\nu]} - i\left[G^{\mu}, G^{\nu}\right]$ of (3.1). Using this in the non-Abelian (4.2), also making use of $D^{;\mu} = \partial^{;\mu} - iG^{\mu}$, noting as just reviewed in (4.4) that $\left(R_{\tau}{}^{\nu\sigma\mu} + R_{\tau}{}^{\sigma\mu\nu} + R_{\tau}{}^{\mu\nu\sigma}\right) G^{\tau} = 0$, and condensing with the cyclator $(\sigma\mu\nu)$, we obtain:

$$\begin{aligned}
P^{\sigma\mu\nu} &= D^{;\sigma}F^{\mu\nu} + D^{;\mu}F^{\nu\sigma} + D^{;\nu}F^{\sigma\mu} \\
&= D^{;\sigma}\left(\partial^{;[\mu}G^{\nu]} - i\left[G^{\mu},G^{\nu}\right]\right) + D^{;\mu}\left(\partial^{;[\nu}G^{\sigma]} - i\left[G^{\nu},G^{\sigma}\right]\right) + D^{;\nu}\left(\partial^{;[\sigma}G^{\mu]} - i\left[G^{\sigma},G^{\mu}\right]\right) \\
&= \left(R_{\tau}{}^{\nu\sigma\mu} + R_{\tau}{}^{\sigma\mu\nu} + R_{\tau}{}^{\mu\nu\sigma}\right)G^{\tau} - i\left(\partial^{;\sigma}\left[G^{\mu},G^{\nu}\right] + \partial^{;\mu}\left[G^{\nu},G^{\sigma}\right] + \partial^{;\nu}\left[G^{\sigma},G^{\mu}\right]\right) \\
&\quad - i\left(G^{\sigma}\partial^{;[\mu}G^{\nu]} + G^{\mu}\partial^{;[\nu}G^{\sigma]} + G^{\nu}\partial^{;[\sigma}G^{\mu]}\right) - \left(G^{\sigma}\left[G^{\mu},G^{\nu}\right] + G^{\mu}\left[G^{\nu},G^{\sigma}\right] + G^{\nu}\left[G^{\sigma},G^{\mu}\right]\right) \\
&= \mathbf{0} - i\left(\partial^{;\sigma}\left[G^{\mu},G^{\nu}\right] + \partial^{;\mu}\left[G^{\nu},G^{\sigma}\right] + \partial^{;\nu}\left[G^{\sigma},G^{\mu}\right] + G^{\sigma}\partial^{;[\mu}G^{\nu]} + G^{\mu}\partial^{;[\nu}G^{\sigma]} + G^{\nu}\partial^{;[\sigma}G^{\mu]}\right) \\
&\qquad - \left(G^{\sigma}\left[G^{\mu},G^{\nu}\right] + G^{\mu}\left[G^{\nu},G^{\sigma}\right] + G^{\nu}\left[G^{\sigma},G^{\mu}\right]\right) \\
&= \mathbf{0} - i\left(\partial^{;(\sigma}\left[G^{\mu},G^{\nu)}\right] + G^{(\sigma}\partial^{;[\mu}G^{\nu])}\right) - G^{(\sigma}\left[G^{\mu},G^{\nu)}\right] \\
&= \mathbf{0} - i\left(\partial^{;(\sigma}\left[G^{\mu},G^{\nu)}\right] + G^{(\sigma}D^{;[\mu}G^{\nu])}\right)
\end{aligned} \quad .(4.6)$$

It can be shown that $\partial^{;(\sigma}\left[G^{\mu},G^{\nu)}\right] + G^{(\sigma}\partial^{;[\mu}G^{\nu])} = \partial^{;(\sigma}G^{[\mu}G^{\nu])}$ by fully expanding the commutators, reducing, and reconsolidating. This is actually a form of product rule when recast as $\partial^{;(\sigma}\left(G^{[\mu}G^{\nu])}\right) = \partial^{;(\sigma}G^{[\mu}G^{\nu])} + G^{(\mu}\partial^{;[\sigma}G^{\nu])}$ and closely examine spacetime indexes which are fully antisymmetric in σ,μ,ν. But we shall not use this here because we want to maintain the ability to apply Gauss'/Stokes' theorem to the above, and having the term $\partial^{;(\sigma}\left[G^{\mu},G^{\nu)}\right]$ explicitly appear in (4.6) gives us this ability.

So, in sum, (4.3) is the classical chromo-electric field equation of Yang-Mills gauge theory corresponding to Maxwell's equation $J^{\nu} = \partial_{;\mu}F^{\mu\nu}$ for electric charges, and (4.6) is the classical chromo-magnetic field equation of Yang-Mills gauge theory corresponding to Maxwell's magnetic equation $0 = \partial^{;\sigma}F^{\mu\nu} + \partial^{;\mu}F^{\nu\sigma} + \partial^{;\nu}F^{\sigma\mu}$ for magnetic charges.

5. The Chromo-Magnetic Field Equation of Classical Yang-Mills Theory, and its Apparent Confinement Properties

The first point to be observed as regards these Yang-Mills monopoles (4.6) is that the term $\left(R_{\tau}{}^{\nu\sigma\mu} + R_{\tau}{}^{\sigma\mu\nu} + R_{\tau}{}^{\mu\nu\sigma}\right)G^{\tau}$ once again vanishes as in QED with the able assistance of the spacetime geometry itself. As discussed in section 4 above, this is why there are no magnetic monopoles in QED. But, solely and directly as a result of the fact that $\left[G^{\mu},G^{\nu}\right] \neq 0$, due to the remaining terms $-i\left(\partial^{;(\sigma}\left[G^{\mu},G^{\nu)}\right] + G^{(\sigma}\partial^{;[\mu}G^{\nu])}\right) - G^{(\sigma}\left[G^{\mu},G^{\nu)}\right]$, *these magnetic monopoles are non-vanishing.* So if one believes in Yang-Mills gauge theory, one must also believe that the magnetic monopoles (4.6) exist somewhere, in some form, in the physical universe. What form they exist in is an open question. Whether they are topologically unstable objects that can only be observed for a small fraction of a second in a high energy accelerator; whether they can be made stable via spontaneous symmetry breaking and are hiding in plain sight as baryons and most notably as protons and neutrons (which the author contends is the case); or whether they are something else, is an open question at this point. But the non-commuting nature of the Yang-Mills gauge fields compels us to take these monopoles (4.6) seriously and ask: what are they, physically, and where and how can we find them, physically?

Second, the above gets even more interesting when considered in differential forms language. The relationship (3.1) now takes on the compacted form $F = dG - iG^2 = DG$. As a result, (4.6) is written compactly with $D = d - iG$ as:

$$P = DF = D\left(dG - iG^2\right) = \left(d - iG\right)\left(dG - iG^2\right) = \mathbf{0} - i\left(dG^2 + GdG\right) - G^3 . \tag{5.1}$$

where $\left(R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma}\right)G^\tau$ is again responsible for $dd = 0$, "the exterior derivative of an exterior derivative is zero." So that term drops out as in Abelian gauge theory, but the remaining terms are non-vanishing. The correspondences between the non-zero terms in (4.6) and (5.1) are $dG^2 \Leftrightarrow \partial^{;(\sigma}\left[G^\mu, G^{\nu)}\right]$, $GdG \Leftrightarrow G^{(\sigma}\partial^{;[\mu}G^{\nu])}$ and $G^3 \Leftrightarrow G^{(\sigma}\left[G^\mu, G^{\nu)}\right]$. So now, via (5.1) and the use of Gauss'/Stokes' theorem $\iint dX = \oint X$ in differential forms, the Yang-Mills magnetic monopole equation in integral form is:

$$\begin{aligned} \iiint P = \iiint dF = \oiint F = \iiint\left(ddG - i\left(dG^2 + GdG\right) - G^3\right) = \iiint\left(-i\left(dG^2 + GdG\right) - G^3\right) \\ = \oiint dG - i\oiint G^2 - \iiint\left(iGdG + G^3\right) = \mathbf{0} - i\oiint G^2 - \iiint\left(iGdG + G^3\right) \end{aligned} . \tag{5.2}$$

Importantly, we are able to apply Gauss'/Stokes' theorem to $dG^2 \Leftrightarrow \partial^{;(\sigma}\left[G^\mu, G^{\nu)}\right]$ but not to $GdG \Leftrightarrow G^{(\sigma}\partial^{;[\mu}G^{\nu])}$ or $G^3 \Leftrightarrow G^{(\sigma}\left[G^\mu, G^{\nu)}\right]$ which is why we kept $\partial^{;(\sigma}\left[G^\mu, G^{\nu)}\right] + G^{(\sigma}\partial^{;[\mu}G^{\nu])}$ rather than converting over to $\partial^{;(\sigma}G^{[\mu}G^{\nu)]}$ as mentioned after (4.6). Note from the bottom line of (5.2), that we may deduce $\oiint dG = 0$, which in (4.5) for electrodynamics tells us that there is no *net* magnetic field flux across any *closed* two-dimensional surface.

Now, focusing on the correspondence $dG^2 \Leftrightarrow \partial^{;(\sigma}\left[G^\mu, G^{\nu)}\right]$, let us expand the differential form to formally write (antisymmetric wedge products $\frac{1}{3!}dx_\sigma \wedge dx_\mu \wedge dx_\nu$ are considered to have already been summed):

$$\begin{aligned} -i\iiint dG^2 = -i\partial^{;(\sigma}\left[G^\mu, G^{\nu)}\right] = -i\iiint\left(\partial^{;\sigma}\left[G^\mu, G^\nu\right] + \partial^{;\mu}\left[G^\nu, G^\sigma\right] + \partial^{;\nu}\left[G^\sigma, G^\mu\right]\right)dx_\sigma dx_\mu dx_\nu \\ = -3i\oiint\left[G^\mu, G^\nu\right]dx_\mu dx_\nu = -i\oiint G^2 \end{aligned} . \tag{5.3}$$

Then let us use this with (4.6) to expand some key terms in (5.2), and thereafter consolidate using $D^{;\mu} = \partial^{;\mu} - iG^\mu$ as follows:

$$\begin{aligned}\iiint P &= \iiint P^{\sigma\mu\nu} dx_\sigma dx_\mu dx_\nu \\ &= \iiint \left(R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma} \right) G^\tau dx_\sigma dx_\mu dx_\nu \\ &- i\iiint \left(\partial^{;\sigma}\left[G^\mu, G^\nu\right] + \partial^{;\mu}\left[G^\nu, G^\sigma\right] + \partial^{;\nu}\left[G^\sigma, G^\mu\right] \right) dx_\sigma dx_\mu dx_\nu \\ &- i\iiint \left(G^\sigma \partial^{;[\mu} G^{\nu]} + G^\mu \partial^{;[\nu} G^{\sigma]} + G^\nu \partial^{;[\sigma} G^{\mu]} \right) dx_\sigma dx_\mu dx_\nu \\ &- \iiint \left(G^\sigma\left[G^\mu, G^\nu\right] + G^\mu\left[G^\nu, G^\sigma\right] + G^\nu\left[G^\sigma, G^\mu\right] \right) dx_\sigma dx_\mu dx_\nu \\ &= \mathbf{0} - 3i \oiint \left[G^\mu, G^\nu\right] dx_\mu dx_\nu - 3i \iiint G^{(\sigma} D^{;[\mu} G^{\nu])} dx_\sigma dx_\mu dx_\nu \\ &= \oiint dG - i\oiint G^2 - i\iiint G\partial G - \iiint G^3 = \oiint dG - i \oiint G^2 - i\iiint GDG \end{aligned} \quad . \tag{5.4}$$

So we see that *inside* the monopole volume, $\iiint \left(R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma} \right) G^\tau dx_\sigma dx_\mu dx_\nu$ describes the coupling of individual the $N^2 - 1$ gauge fields $G^{i\tau}$ of $G^\tau = \lambda^i G^{i\tau}$ to the spacetime geometry, and that this coupling via $R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma} = 0$ conspires to result in $\oiint dG = 0$. Thus the geometry couples to the gauge fields in a manner that prevents gauge fields from *net* flowing in and out across *closed* surfaces enclosing the monopole for exactly the same reasons that there are no magnetic monopoles at all in Abelian gauge theory. What also does not *net* flow across any *closed* surface, but is nonetheless clearly contained within the overall volume represented by the triple integral, is $\iiint GDG = \iiint \left(GdG - iG^3 \right) = \iiint G^{(\sigma} D^{;[\mu} G^{\nu])} dx_\sigma dx_\mu dx_\nu$, whatever this represents. This expression simply is not integrable with $\iint dX = \oint X$. But whatever $\oiint G^2 = 3\oiint \left[G^\mu, G^\nu\right] dx_\mu dx_\nu$ represents, <u>does</u> net flow across a closed two-dimensional surface. We shall soon demonstrate that this term represents the flow of mesons.

Third, making (4.6) even more interesting, as detailed in section 1 of [8], if we perform a local transformation $F \to F' = F - dG$ on the field strength F, which in expanded form is written as $F^{\mu\nu} \to F^{\mu\nu}{}' = F^{\mu\nu} - \partial^{[\nu} G^{\mu]}$, then we find from (5.2) as a direct result of $\oiint dG = 0$, that:

$$\iiint P = \oiint F \to \oiint F' = \oiint \left(F - dG \right) = \oiint F \quad . \tag{5.5}$$

This means that the flow of the field strength $\oiint F = -i\oiint G^2$ across a two dimensional surface is invariant under the local gauge-like transformation $F^{\mu\nu} \to F^{\mu\nu}{}' = F^{\mu\nu} - \partial^{[\nu} G^{\mu]}$.

Fourth, we see from (5.4) that $\iiint G^3 = 3\iiint G^\sigma \left[G^\mu, G^\nu\right] dx_\sigma dx_\mu dx_\nu$ is one of the non-integrable terms. This involves pure antisymmetric three-field cubic interactions $G^\sigma \wedge G^\mu \wedge G^\nu$ among the gauge fields. While we shall avoid the use of the term "glueball" to describe this because this term already has certain technical meanings for which its use here might cause confusion, certainly this term contained within the monopole volume is an amalgam of pure interaction gauge fields which nicely displays the non-linearity of Yang-Mills gauge theory.

Now, as much as the MIT Bag Model reviewed in, e.g., [12] section 18 has certain inelegant features such as the *ad hoc* introduction of backpressures to force confinement, this model very correctly makes one very important point that deserves utmost attention beyond the specifics of any particular model of confinement: *focus carefully on what flows and does not flow across any closed two-dimensional surface.* This is why the integral form of Maxwell's equations is so vital to any sensible discussion of confinement. The confinement of gauge fields (which in SU(3) QCD are represented by the eight gluons of $G^\tau = \lambda^i G^{i\tau}$ with $i = 1, 2, 3...8$) is symbolically specified by $\oint\!\!\oint \text{Gluons} = 0$. Similarly, the confinement of individual quarks (which are represented by the SU(3) Dirac wavefunction ψ_A; $A = 1, 2, 3$ with three color eigenstates *R*, *G*, *B*) is specified symbolically by $\oint\!\!\oint \text{Quarks} = 0$. Different theories may have different ways to achieve these two symbolic confinements, but in the end, one should pay close attention to the two-dimensional closed surface integrals and carefully examine what does and does not flow across these closed surfaces. Equations (5.2) through (5.5) contain a lot of information about what does and does not flow across the closed $\oint\!\!\oint$ surface of a Yang-Mills monopole, so as taught by the MIT Bag Model, we should study these equations carefully to see if these magnetic monopoles exhibit any attributes of confined gluons and quarks, or interactions via mesons.

A first point is made by $\iiint \left(R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma} \right) G^\tau dx_\sigma dx_\mu dx_\nu$ which leads to $\oint\!\!\oint dG = 0$ in (5.4) and is the exact same expression which yields the absence of magnetic monopoles entirely, in Abelian electrodynamics, review (4.4). This $\iiint \left(R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma} \right) G^\tau dx_\sigma dx_\mu dx_\nu$ term contains an *individual* gauge field $G^\tau = \lambda^i G^{i\tau}$, zeroed out as a direct result of its coupling through the Riemannian geometry in the configuration of the first Bianchi identity, and upon Gauss' / Stokes' integration yields $\oint\!\!\oint dG = 0$. So the question, in the context of the MIT bag model, is whether this term is to be interpreted as telling us that gauge fields (gluons in SU(3) QCD) are confined, which means that there is never a *net* flow of gauge fields across any *closed* surface surrounding a Yang-Mills magnetic monopole. As is the case with electrodynamics, Yang-Mills magnetic fields (and gluon fields in QCD) can and do flow, in net, through *open* surfaces, but because magnetic fields are aterminal fields, an outward flux over one portion of a closed surface is always cancelled by an inward flux across another portion of the closed surface. This is strengthened by the fact displayed in (5.5) that $\oint\!\!\oint F \to \oint\!\!\oint F' = \oint\!\!\oint F$ is invariant under the transformation $F \to F' = F - dG$, i.e., $F^{\mu\nu} \to F^{\mu\nu}{}' = F^{\mu\nu} - \partial^{[\nu} G^{\mu]}$ which renders the gauge fields (gluons in QCD) *not observable* with respect to net flux through the closed surface. *This would mean as argued in section 1 of [8] that gauge fields are confined in Yang-Mills theory <u>for the exact same geometric reasons</u> that magnetic monopoles <u>do not exist at all</u> in Abelian gauge theory.*

A second point is made by the fact that $\oint\!\!\oint G^2 = 3\oint\!\!\oint \left[G^\mu, G^\nu \right] dx_\mu dx_\nu$ which is the integrable term in (5.4), is really the telling us the crux of what *does* net flow across closed surfaces of a Yang-Mills magnetic monopole. The only thing that does net flow, are these

$3\left[G^\mu, G^\nu\right]$ entities. While we still must determine, physically, what these $3\left[G^\mu, G^\nu\right]$ entities represent, we do know that $\left[G^\mu, G^\nu\right] \neq 0$ is at the heart of the non-Abelian character of Yang-Mills theories. If these $3\left[G^\mu, G^\nu\right]$ do *not* turn out to represent individual quarks, then what (5.4) would be telling us, in the sense of the MIT bag model, is that neither individual gluons nor individual quarks net flow across the closed surface of a Yang-Mills magnetic monopole, $\oiint \text{Gluons} = 0$ and $\oiint \text{Quarks} = 0$. But what we also know is that baryons interact via meson exchange, and that mesons have a color wavefunction of the form $\overline{R}R + \overline{G}G + \overline{B}B$. So mesons *should* be permitted to flow in and out of baryons, that is, we should also have $\oiint \text{Mesons} \neq 0$. So if we can show that $\oiint G^2 = 3\oiint\left[G^\mu, G^\nu\right] dx_\mu dx_\nu$ represents meson flow, as we shall shortly do, then these magnetic monopoles would forbid net quark and gluon flows but permit net meson flow, and we would have some very strong formal reasons for identifying Yang-Mills magnetic monopoles with baryons.

Additionally, the factors of "3" which also emerge in $\oiint G^2 = 3\oiint\left[G^\mu, G^\nu\right] dx_\mu dx_\nu$ and in $\iiint GDG = 3\iiint G^{(\sigma} D^{;[\mu} G^{\nu])} dx_\sigma dx_\mu dx_\nu$ in (5.3) and (5.4), although it comes from the three additive terms in the various expressions in (5.4), also signifies the number of colors of quark in QCD, the number of quarks in a baryon, and the number of terms in the meson color wavefunction $\overline{R}R + \overline{G}G + \overline{B}B$. So this "3" is a very strong hint – on top of the fact that $P^{\sigma\mu\nu}$ itself has three totally-antisymmetric spacetime indexes each capable of accommodating one of three vector current densities, and contains three additive terms – that there is some very definitive "three-ness" associated with these Yang-Mills monopoles. This "three-ness" could save us having to *postulate* three quarks per baryon as is presently done in QCD, and would instead *require* us to have three quarks per baryon upon after which we would then impose QCD as an Exclusion Principle. In other words, if this "three-ness" is telling us that a Yang-Mills monopole contains three quarks and has all the other required symmetries of a baryons, then *postulating Yang-Mills theory would be synonymous with postulating QCD and postulating baryons and postulating that the baryons contain three colored quarks. This would make QCD itself an unavoidable, purely deductive consequence of Yang-Mills gauge theory, and would greatly strengthen the roots of QCD!* It would at the same time answer the unanswered question as to why baryons contain three quarks and not some other number. These symmetry relationships are what led the author in April 2005 to begin taking seriously, the thesis that these non-vanishing magnetic monopoles originating from the non-commuting gauge fields of Yang-Mills gauge theory might be baryons.

But so far, beyond this number "3," there is no hint in this present development of any quarks in the Yang-Mills monopole (5.4). So we need to now see if there is some way to "populate" these magnetic monopoles with quarks. This brings us back to (4.3), which is the field equation relating Yang-Mills *electric* charge densities J^ν to the gauge fields G_μ, and which we shall study more closely in section 9. But at this point, it will be helpful to first explore two more views of Yang-Mills theory, namely the "perturbative" view to now be developed in section 6, and the "curvature" view to be developed in section 7. Not only are these two views helpful as to how we conceptualize Yang-Mills theory, but they also simplify the mathematical development of Yang-Mills theory.

6. The Yang-Mills Perturbation Tensor: A Fourth View of Yang-Mills

In section 4, we described three equivalent "views" of Yang-Mills gauge theory: as a field theory of non-commuting gauge fields (3.1); as a theory of non-linear interactions among the gauge fields (3.4); and as a minimally-coupled gauge theory on steroids (3.6), (4.1), (4.2) in which ordinary derivatives are made gauge-covariant $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$. Now, we introduce yet a fourth view of Yang-Mills gauge theory, the "perturbative view," which is motivated by the field equations (4.1), (4.2) when the field strength is expressed as $F^{\mu\nu} = D^{;[\mu}G^{\nu]}$ in the steroidal view of (3.5). This view is rooted in the Klein-Gordon equation

$$\begin{aligned} 0 &= \left(D_\sigma D^\sigma + m^2\right)\phi = \left(\left(\partial_\sigma - iG_\sigma\right)\left(\partial^\sigma - iG^\sigma\right) + m^2\right)\phi = \left(\partial_\sigma\partial^\sigma + m^2 - i\partial_\sigma G^\sigma - iG_\sigma\partial^\sigma - G_\sigma G^\sigma\right)\phi \\ &= \left(\partial_\sigma\partial^\sigma + m^2 + V\right)\phi \end{aligned} \tag{6.1}$$

for an interacting scalar field, where in the final line one identifies and defines an electromagnetic *perturbation* spacetime scalar:

$$V \equiv -i\partial_\sigma G^\sigma - iG_\sigma\partial^\sigma - G_\sigma G^\sigma \tag{6.2}$$

In virtually identical fashion, we may use (3.5) to rewrite the Yang-Mills chromo-electric field equation (4.3) as:

$$\begin{aligned} J^\nu &= \left(g^{\mu\nu}\left(\left(\partial_{;\sigma}\partial^{;\sigma} - i\left(\partial_{;\sigma}G^\sigma + G_\sigma\partial^{;\sigma}\right) - G_\sigma G^\sigma\right) + m^2\right) - \left(\partial^{;\mu}\partial^{;\nu} - i\left(\partial^{;\mu}G^\nu + G^\mu\partial^{;\nu}\right) - G^\mu G^\nu\right)\right)G_\mu \\ &= \left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + V + m^2\right) - \left(\partial^{;\mu}\partial^{;\nu} + V^{\mu\nu}\right)\right)G_\mu \end{aligned} \quad ,(6.3)$$

where in the final line, we have defined a "perturbation tensor" and its trace scalar:

$$V^{\mu\nu} \equiv -i\left(\partial^{;\mu}G^\nu + G^\mu\partial^{;\nu}\right) - G^\mu G^\nu \tag{6.4}$$

$$V = V_\sigma{}^\sigma = -i\partial_\sigma G^\sigma - iG_\sigma\partial^\sigma - G_\sigma G^\sigma = -i\partial_\sigma G^\sigma_{AB} - iG_{AB\sigma}\partial^\sigma - G_{AC\sigma}G^\sigma_{CB}. \tag{6.5}$$

The perturbation scalar is *identical* in form to (6.2), but in Yang-Mills theory, it is a 3x3 Yang-Mills matrix of spacetime scalars, as we are reminded about by the explicit showing of Yang-Mills indexes in (6.5).

Noting that for any two successive gave-covariant derivatives:

$$D^{;\mu}D^\nu = \left(\partial^{;\mu} - iG^\mu\right)\left(\partial^{;\nu} - iG^\nu\right) = \partial^{;\mu}\partial^{;\nu} - i\partial^{;\mu}G^\nu - iG^\mu\partial^{;\nu} - G^\mu G^\nu = \partial^{;\mu}\partial^{;\nu} + V^{\mu\nu}, \tag{6.6}$$

we see that in flat spacetime where $\left[\partial^{;\mu},\partial^{;\nu}\right]=\left[\partial^{\mu},\partial^{\nu}\right]=0$, the antisymmetric combination:

$$V^{[\mu\nu]}=V^{\mu\nu}-V^{\nu\mu}=\left[D^{;\mu},D^{;\nu}\right]. \tag{6.7}$$

So $V^{[\mu\nu]}$ is synonymous with the commutator of the Yang-Mills covariant derivatives. In curved spacetime, using (6.7) to operate on a vector field A^{σ} and applying the Riemann curvature definition $\left[\partial_{;\mu},\partial_{;\nu}\right]G_{\alpha}\equiv R^{\sigma}{}_{\alpha\mu\nu}G_{\sigma}$, we obtain:

$$\left[D^{;\mu},D^{;\nu}\right]A^{\sigma}=\left[\partial^{;\mu},\partial^{;\nu}\right]A^{\sigma}+V^{[\mu\nu]}A^{\sigma}=\left(R_{\tau}{}^{\sigma\mu\nu}+\delta_{\tau}{}^{\sigma}V^{[\mu\nu]}\right)A^{\tau}. \tag{6.8}$$

Applying (6.8) to the magnetic monopole (4.6), the curvature terms vanish as in (4.4) via $R_{\tau}{}^{\nu\sigma\mu}+R_{\tau}{}^{\sigma\mu\nu}+R_{\tau}{}^{\mu\nu\sigma}=0$, and so we obtain simply, *in both curved and flat spacetime*:

$$\begin{aligned}P^{\sigma\mu\nu}&=D^{;\sigma}D^{;[\mu}G^{\nu]}+D^{;\mu}D^{;[\nu}G^{\sigma]}+D^{;\nu}D^{;[\sigma}G^{\mu]}\\&=\left[D^{;\sigma},D^{;\mu}\right]G^{\nu}+\left[D^{;\mu},D^{;\nu}\right]G^{\sigma}+\left[D^{;\nu},D^{;\sigma}\right]G^{\mu}.\\&=V^{[\sigma\mu]}G^{\nu}+V^{[\mu\nu]}G^{\sigma}+V^{[\nu\sigma]}G^{\mu}\end{aligned} \tag{6.9}$$

The chromo-electric and chromo-magnetic field equations expressed in the form of (6.3) and (6.9), illustrate this fourth, "perturbative" view of Yang-Mills theory. In fact, it is a very useful exercise, to ask about the *difference* between the physics of Yang-Mills theory and that of ordinary Abelian gauge theory, which difference is wholly measured by the perturbation and functions of the perturbation. *It is this fourth view of Yang-Mills – the perturbative view – that will enable us to fill the "mass gap."*

To better understand the perturbative view, we introduce the labels "P" to denote "Perturbative," "YM" to denote the complete, holistic physics encompassing all features of Yang-Mills, and "L" to denote the "Linear" expressions of Abelian gauge theories, most notably electrodynamics. Schematically, YM=L+P, that is, the complete physics of Yang Mills YM theory may be thought of and analyzed as the sum of a perturbative portion P and a linear portion L. Thus, from (6.3), we can deduce that the perturbative-only portion of the current density, J_{P}^{ν}, which is the difference $J_{YM}^{\nu}-J_{L}^{\nu}$ between the complete Yang-Mills current density J_{YM}^{ν} of (6.3) and the linear density $J_{L}^{\nu}=\left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma}+m^{2}\right)-\partial^{;\mu}\partial^{;\nu}\right)G_{\mu}$ of Abelian theory. This is given by:

$$\begin{aligned}J_{P}^{\nu}\equiv J_{YM}^{\nu}-J_{L}^{\nu}&=\left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma}+V+m^{2}\right)-\left(\partial^{;\mu}\partial^{;\nu}+V^{\mu\nu}\right)\right)G_{\mu}-\left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma}+m^{2}\right)-\partial^{;\mu}\partial^{;\nu}\right)G_{\mu}\\&=\left(g^{\mu\nu}V-V^{\mu\nu}\right)G_{\mu}\end{aligned} \tag{6.10}$$

In other words, $J_{P}^{\nu}=\left(g^{\mu\nu}V-V^{\mu\nu}\right)G_{\mu}$ summarizes all of the effects which are added to an Abelian gauge theory, by the non-linear perturbations of Yang-Mills theory.

For the magnetic monopoles, of course, $P_P^{\sigma\mu\nu} \equiv P_{YM}^{\sigma\mu\nu}$, because the Abelian monopole densities of Abelian gauge theory are zero, $P_L^{\sigma\mu\nu} = 0$. We know this of course from (4.4), but we also see this by inspection from (6.9) in which the non-vanishing magnetic monopole arises completely via the index-cyclical application of the antisymmetrized perturbation operator $V^{[\mu\nu]}$ to Yang-Mills gauge fields G^σ. If $V^{\mu\nu} \to 0$, the monopole densities $P^{\sigma\mu\nu} \to 0$ go to zero. Yang-Mills monopoles are entirely a creature of perturbation, as they equivalently are creatures of non-Abelian gauge fields, of non-linear gauge interactions, and of gauge theory on steroids.

7. Hermann Weyl's Gauge Theory and Gravitational Curvature: A Fifth Geometric View of Yang-Mills

Hermann Weyl in 1918 [13], [14] first conceived the idea that that electrodynamics might be unified with gravitation by analyzing the "twisting" of vectors under parallel transport to measure the geometric curvature of the space. While Weyl first conceived of this as a local "gauge" symmetry, in 1929 [15] he corrected his original misconception into the modern view of a local "phase" symmetry. Notwithstanding, the original misnomer "gauge" is still used to name his theory, perhaps as a reminder to posterity that the most bedrock physical theories are sometimes properly-conceived in the abstract but misconceived in some details that need to be worked out over time. While gravitation operates via the curvature of a physical, non-compact configuration space $\mathfrak{R}^4$ first pioneered by Minkowski [16] based on Einstein's 1905 development of Lorentz invariance into Special Relativity [17], Weyl's theory operates along the circle of an abstract phase space centered around the local phase $\exp i\theta(x)$ for Abelian theory, and $\exp i\theta(x) = \exp i\lambda^i\theta^i(x)$ with $i = 1,2,3...N^2-1$ for an SU(N) Yang-Mills theory.

The relationship (6.8), illustrates Weyl's curvature idea very clearly. We see that the anti-symmetrized $\delta_\tau{}^\sigma V^{[\mu\nu]}$ plays a role in Yang-Mills theory very similar to that played by the Riemann tensor $R_\tau{}^{\sigma\mu\nu}$ in gravitational theory: each is a "curvature" measuring of the degree to which the spacetime derivatives do or do not commute. In fact, lowering all of the indexes on the Riemann tensor in (6.8), we see that in going from an Abelian gauge theory in curved spacetime to a Yang-Mills theory in curved spacetime, we make the *operator* replacement $R_{\tau\sigma\mu\nu} \to R_{\tau\sigma\mu\nu} + g_{\tau\sigma}V_{[\mu\nu]}$ when operating on any vector A^τ. Thus:

$$g_{\tau\sigma}\left[D_{;\mu}, D_{;\nu}\right]A^\tau = \left(R_{\tau\sigma\mu\nu} + g_{\tau\sigma}V_{[\mu\nu]}\right)A^\tau . \tag{7.1}$$

So just as $R_{\tau\sigma\mu\nu}$ represents curvature in spacetime, $g_{\tau\sigma}V_{[\mu\nu]}$ represents Weyl's gauge curvature. We note the leading role of the anti-symmetrized perturbation $V_{[\mu\nu]}$ in this curvature connection. It is also interesting to note the superposition of the symmetric metric tensor $g_{\tau\sigma}$ against the antisymmetric $\tau\sigma$ indexes in the first two positions of the Riemann tensor, which means that the resulting operator $R_{\tau\sigma\mu\nu} + g_{\tau\sigma}V_{[\mu\nu]}$ is non-symmetric. But this is absorbed in the operation on A^τ which sums out the τ index.

In fact, we can and should apply the same curvature analysis to the gauge-covariant derivative in curved spacetime, which we now write as:

$$D_{;\mu}A_\nu = \partial_{;\mu}A_\nu - iG_\mu A_\nu = \partial_\mu A_\nu - \Gamma^\alpha{}_{\mu\nu}A_\alpha - iG_\mu A_\nu . \tag{7.2}$$

With minor manipulation, and using $\Gamma_{\alpha\,\mu\nu} = \frac{1}{2}\left(g_{\nu\alpha,\mu} + g_{\alpha\mu,\nu} - g_{\mu\nu,\alpha}\right)$, we can reframe this as:

$$g_{\alpha\nu}D_{;\mu}A^\alpha = \left(g_{\alpha\nu}\partial_\mu - \Gamma_{\alpha\,\mu\nu} - ig_{\alpha\nu}G_\mu\right)A^\alpha . \tag{7.3}$$

So here, the curvature view is highlighted by the fact that when going from Abelian to Yang-Mills gauge theory in curved spacetime, we make the operator replacement $\Gamma_{\alpha\,\mu\nu} \to \Gamma_{\alpha\,\mu\nu} + ig_{\alpha\nu}G_\mu$ when operating on the vector A^α. Because $\Gamma_{\alpha\,\mu\nu}$ captures the effects of parallel transport in curved spacetime, we see that $ig_{\alpha\nu}G_\mu$ represents Weyl's parallel transport in gauge (phase) space. As with (7.1), the combined operator $\Gamma_{\alpha\,\mu\nu} + ig_{\alpha\nu}G_\mu$ is non-symmetric, because $\Gamma_{\alpha\,\mu\nu}$ is symmetric in μ,ν while $ig_{\alpha\nu}G_\mu$ is symmetric in α,ν. And as with (7.1), this is absorbed in the operation on A^α which sums out the α index. In contrast to (7.1), however, the curvature operator $R_{\tau\sigma\mu\nu} + g_{\tau\sigma}V_{[\mu\nu]}$ is a tensor, but the parallel transport operator $\Gamma_{\alpha\,\mu\nu} + ig_{\alpha\nu}G_\mu$ is not because $\Gamma_{\alpha\,\mu\nu}$ is not a tensor. Only the entire $g_{\alpha\nu}\partial_\mu - \Gamma_{\alpha\,\mu\nu} - ig_{\alpha\nu}G_\mu$ is a tensor operator.

Given this curvature view of Yang-Mills, and especially (7.1), we now note the two geometric Bianchi identities $R_{\tau\sigma\mu\nu} + R_{\tau\mu\nu\sigma} + R_{\tau\nu\sigma\mu} = 0$ and $\partial_{;\alpha}R_{\tau\sigma\mu\nu} + \partial_{;\mu}R_{\tau\sigma\nu\alpha} + \partial_{;\nu}R_{\tau\sigma\alpha\mu} = 0$. The former was already employed in (4.4) to yield vanishing magnetic monopoles in Abelian gauge theory and a vanishing term $\left(R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma}\right)G^\tau = 0$ in the non-vanishing magnetic monopole (4.6) of Yang-Mills theory, which "0" is responsible for the confinement of gauge fields with respect to any closed surface, as was discussed at length in section 5. The latter, when manipulated into the contracted form $\partial_{;\nu}\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right) = 0$ and then connected to a conserved energy tensor $\partial_{;\nu}T^{\mu\nu} = 0$, is at the center of classical gravitational field theory. So we certainly want to inject these identities into Yang-Mills theory to the greatest degree possible because they are at the center of both the magnetic monopoles and gravitational theory.

First, let's take $R_{\tau\sigma\mu\nu} + R_{\tau\mu\nu\sigma} + R_{\tau\nu\sigma\mu} = 0$. Because (4.1) contains $R_{\tau\sigma\mu\nu}$ which is the first term of this identity, let use rewrite (4.1) two more times with a simple renaming of indexes to match the other two terms in $R_{\tau\sigma\mu\nu} + R_{\tau\mu\nu\sigma} + R_{\tau\nu\sigma\mu} = 0$. Then, let's add these all together to write:

$$\begin{aligned}
&\left(g_{\tau\sigma}\left[D_{;\mu},D_{;\nu}\right]+g_{\tau\mu}\left[D_{;\nu},D_{;\sigma}\right]+g_{\tau\nu}\left[D_{;\sigma},D_{;\mu}\right]\right)A^\tau \\
&=\left(R_{\tau\sigma\mu\nu}+R_{\tau\mu\nu\sigma}+R_{\tau\nu\sigma\mu}+g_{\tau\sigma}V_{[\mu\nu]}+g_{\tau\mu}V_{[\nu\sigma]}+g_{\tau\nu}V_{[\sigma\mu]}\right)A^\tau \\
&=\left(g_{\tau\sigma}V_{[\mu\nu]}+g_{\tau\mu}V_{[\nu\sigma]}+g_{\tau\nu}V_{[\sigma\mu]}\right)A^\tau \\
&=g_{\tau(\sigma}\left[D_{;\mu},D_{;\nu)}\right]A^\tau=g_{\tau(\sigma}V_{[\mu\nu])}A^\tau
\end{aligned}. \tag{7.4}$$

In the final line, we have applied $R_{\tau\sigma\mu\nu}+R_{\tau\mu\nu\sigma}+R_{\tau\nu\sigma\mu}=0$ to zero out the terms that contain the Riemann tensor. Once again the perturbation and the curvature views converge together. In fact, here, in contrast to (7.1) and (7.3), we *can* slice off the A^τ operand, and simply write the operator equation:

$$g_{\tau(\sigma}\left[D_{;\mu},D_{;\nu)}\right]=g_{\tau(\sigma}V_{[\mu\nu])}\,. \tag{7.5}$$

This is allowed because the spacetime index symmetries on the left and right side of the above match, and so we do not need to sum out index the τ index to obtain matching spacetime symmetries.

Let us now absorb the spacetime indexes to lower the indexes on the generalized vector A^τ, and then rename this into the *specific* vector $A_\mu \to G_\mu = \lambda^i G^i_\mu$ with represents the Yang-Mills gauge field. Now, (7.4) becomes:

$$\begin{aligned}
P_{\mu\nu\sigma}&=\left[D_{;\mu},D_{;\nu}\right]G_\sigma+\left[D_{;\nu},D_{;\sigma}\right]G_\mu+\left[D_{;\sigma},D_{;\mu}\right]G_\nu=V_{[\mu\nu]}G_\sigma+V_{[\nu\sigma]}G_\mu+V_{[\sigma\mu]}G_\nu \\
&=\left[D_{;(\mu},D_{;\nu}\right]G_{\sigma)}=V_{([\mu\nu]}G_{\sigma)}
\end{aligned}. \tag{7.6}$$

Contrasting, this is *totally identical* to equation (6.9) for the Yang-Mills monopole, simply with covariant rather than contravariant indexes. Again the perturbative and curvature views converge: The Yang-Mills monopole density is no more and no less than the geometric operator identity $g_{\tau(\sigma}\left[D_{;\mu},D_{;\nu)}\right]=g_{\tau(\sigma}V_{[\mu\nu])}$ of (7.5) – which is the Yang-Mills version of $R_{\tau\sigma\mu\nu}+R_{\tau\mu\nu\sigma}+R_{\tau\nu\sigma\mu}=0$ – applied to the Yang-Mills gauge field G_μ.

Next, because (7.5) is valid standing alone as an operator equation, let us now multiply this (in the expanded form of (7.4)) from the *left* by a general vector A^τ. Thus we now write:

$$A^\tau\left(g_{\tau\sigma}\left[D_{;\mu},D_{;\nu}\right]+g_{\tau\mu}\left[D_{;\nu},D_{;\sigma}\right]+g_{\tau\nu}\left[D_{;\sigma},D_{;\mu}\right]\right)=A^\tau\left(g_{\tau\sigma}V_{[\mu\nu]}+g_{\tau\mu}V_{[\nu\sigma]}+g_{\tau\nu}V_{[\sigma\mu]}\right). \tag{7.7}$$

Upon lowering indexes this becomes:

$$\begin{aligned}
&A_\sigma\left[D_{;\mu},D_{;\nu}\right]+A_\mu\left[D_{;\nu},D_{;\sigma}\right]+A_\nu\left[D_{;\sigma},D_{;\mu}\right]=A_\sigma V_{[\mu\nu]}+A_\mu V_{[\nu\sigma]}+A_\nu V_{[\sigma\mu]} \\
&=A_{(\sigma}\left[D_{;\mu},D_{;\nu)}\right]=A_{(\sigma}V_{[\mu\nu])}
\end{aligned}. \tag{7.8}$$

Contrasting to (7.4) written as $\left[D_{;(\mu}, D_{;\nu}\right] A_{\sigma)} = V_{([\mu\nu]} A_{\sigma)}$, we see that *any* vector A_σ may be commuted with $V_{[\mu\nu]}$ when the spacetime indexes are cycles as in the above. This is an important commutativity relationship to have in mind when we regard A_σ as an NxN matrix of vectors in Yang Mills theory.

Speaking of which, let us do just that. If we set $A_\mu \to G_\mu = \lambda^i G^i_\mu$ again as we did for (7.6), then (7.8) becomes $G_{(\sigma}\left[D_{;\mu}, D_{;\nu)}\right] = G_{(\sigma} V_{[\mu\nu])}$, which is a cousin of the magnetic monopole equation (7.6) in which the gauge fields appear on the left rather than the right. But because the gauge fields are contained within $D_{;\mu} = \partial_{;\mu} - iG_\mu$, let us set the vector $A_\sigma \to D_{;\sigma}$ in both (7.4) written as $\left[D_{;(\mu}, D_{;\nu}\right] A_{\sigma)} = V_{([\mu\nu]} A_{\sigma)}$ and in (7.8), and then use the Jacobian (determinant-related) identity $\left[a,[b,c]\right] + \left[b,[c,a]\right] + \left[c,[a,b]\right] = 0$ to combine these into the single relationship:

$$\left[D_{;(\mu}, D_{;\nu}\right] D_{;\sigma)} = V_{([\mu\nu]} D_{;\sigma)} = D_{;(\sigma}\left[D_{;\mu}, D_{;\nu)}\right] = D_{;(\sigma} V_{[\mu\nu])} . \tag{7.9}$$

Because this commutes $D_{;(\sigma}$ to the left of the commutator $\left[D_{;\mu}, D_{;\nu)}\right]$, this sets up the ability to now incorporate the remaining Bianchi identity $\partial_{;(\alpha} R_{|\tau\sigma|\mu\nu)} \equiv \partial_{;\alpha} R_{\tau\sigma\mu\nu} + \partial_{;\mu} R_{\tau\sigma\nu\alpha} + \partial_{;\nu} R_{\tau\sigma\alpha\mu} = 0$ which underpins the expression $\partial_{;\nu}\left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R\right) = 0$ that is at the heart of gravitational theory. In the expression $\partial_{;(\alpha} R_{|\tau\sigma|\mu\nu)}$, we define the notation $|\tau\sigma|$ as a "wall" to seal off and fix the $\tau\sigma$ indexes (this is *not* an absolute value symbol as used here) from the $(\sigma\mu\nu)$ cycling of the remaining indexes. But before we do this, let us expand one of the gauge-covariant derivatives in (7.9) and also bring in (7.6) in the form $-iP_{\mu\nu\sigma} = -i\left[D_{;(\mu}, D_{;\nu}\right] G_{\sigma)} = -iV_{([\mu\nu]} G_{\sigma)}$, and also pull back in (4.6) which is our earlier monopole expression, to write the above as:

$$\begin{aligned} P_{\mu\nu\sigma} &= i\left(D_{;(\sigma}\left[D_{;\mu}, D_{;\nu)}\right] - \left[D_{;(\mu}, D_{;\nu}\right]\partial_{;\sigma)}\right) = i\left(D_{;(\sigma} V_{[\mu\nu])} - V_{([\mu\nu]}\partial_{;\sigma)}\right) \\ &= -i\left(\partial_{;(\sigma}\left[G_\mu, G_{\nu)}\right] + G_{(\sigma} D_{;[\mu} G_{\nu])}\right) \end{aligned} . \tag{7.10}$$

In this form, we have now turned the magnetic monopole density itself, into an operator!

Now, let's move on to the second Bianchi identity $\partial_{;(\alpha} R_{|\tau\sigma|\mu\nu)} = 0$. We start with (7.1) written in the form $\left[D_{;\mu}, D_{;\nu}\right] A_\sigma = R_{\tau\sigma\mu\nu} A^\tau + V_{[\mu\nu]} A_\sigma$. We operate on all three terms from the left using $D_{;\alpha}$. Thus, $D_{;\alpha}\left(\left[D_{;\mu}, D_{;\nu}\right] A_\sigma\right) = D_{;\alpha}\left(R_{\tau\sigma\mu\nu} A^\tau\right) + D_{;\alpha}\left(V_{[\mu\nu]} A_\sigma\right)$. Then we replicate this expression two more times via a simple renaming of indexes with a cycling of μ, ν, α. We then add all of these together, to fashion:

$$\begin{aligned}&D_{;\alpha}\left(\left[D_{;\mu},D_{;\nu}\right]A_\sigma\right)+D_{;\mu}\left(\left[D_{;\nu},D_{;\alpha}\right]A_\sigma\right)+D_{;\nu}\left(\left[D_{;\alpha},D_{;\mu}\right]A_\sigma\right)\\&=D_{;\alpha}\left(R_{\tau\sigma\mu\nu}A^\tau\right)+D_{;\mu}\left(R_{\tau\sigma\nu\alpha}A^\tau\right)+D_{;\nu}\left(R_{\tau\sigma\alpha\mu}A^\tau\right)+D_{;\alpha}\left(V_{[\mu\nu]}A_\sigma\right)+D_{;\mu}\left(V_{[\nu\alpha]}A_\sigma\right)+D_{;\nu}\left(V_{[\alpha\mu]}A_\sigma\right).\\&=D_{;(\alpha}\left(\left[D_{;\mu},D_{;\nu)}\right]A_\sigma\right)=D_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)}A^\tau\right)+D_{;(\alpha}\left(V_{[\mu\nu])}A_\sigma\right)\end{aligned} \tag{7.11}$$

It should be clear how the term $D_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)}A^\tau\right)$ sets up the ability to apply $\partial_{;(\alpha}R_{|\tau\sigma|\mu\nu)}=0$. So now let's proceed.

We can slightly expand the compacted form in the bottom line of (7.11) using $D_{;(\alpha}=\partial_{;(\alpha}-iG_{(\alpha}$, take the spacetime derivative $\partial_{;(\alpha}$ using the product rule, and make use of the Bianchi identity $\partial_{;(\alpha}R_{|\tau\sigma|\mu\nu)}=0$ to write $\partial_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)}A^\tau\right)=R_{\tau\sigma(\mu\nu}\partial_{;\alpha)}A^\tau$, thus obtaining:

$$\begin{aligned}D_{;(\alpha}\left(\left[D_{;\mu},D_{;\nu)}\right]A_\sigma\right)&=\partial_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)}A^\tau\right)-iG_{(\alpha}\left(R_{|\tau\sigma|\mu\nu)}A^\tau\right)+D_{;(\alpha}\left(V_{[\mu\nu])}A_\sigma\right)\\&=R_{\tau\sigma(\mu\nu}\partial_{;\alpha)}A^\tau-iG_{(\alpha}R_{|\tau\sigma|\mu\nu)}A^\tau+D_{;(\alpha}\left(V_{[\mu\nu])}A_\sigma\right)\end{aligned}. \tag{7.12}$$

That is it! We have now incorporated the Bianchi identity $\partial_{;(\alpha}R_{|\tau\sigma|\mu\nu)}=0$ which underlies the geometric heart of gravitational theory, $\partial_{;\nu}\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)=0$, directly into Yang-Mills. Now what remains is to rework (7.12) to make some of its meanings more transparent.

Continuing with (7.12), in the third line below we commute $G_{;(\alpha}R_{|\tau\sigma|\mu\nu)}=R_{\tau\sigma(\mu\nu}G_{;\alpha)}$, because while $R_{\tau\sigma\mu\nu}$ is a spacetime fourth rank tensor, it is simply a 1x1 matrix in Yang-Mills theory. In other words, while G_α and are $D_{;\alpha}$ and $V_{[\mu\nu]}$ are all NxN matrices which do not mutually commute with one another or even with themselves when the spacetime indexes are different, $R_{\tau\sigma\mu\nu}$ and (when it appears) $g_{\mu\nu}$ can be freely moved to any left-right position as desired. In the fourth line we consolidate the first and second term using $D_{;\alpha)}=\partial_{;\alpha)}-iG_{\alpha)}$. In the fifth line we use $D_{;\alpha)}=\partial_{;\alpha)}-iG_{\alpha)}$ to expand the $D_{;(\alpha}\left(V_{[\mu\nu])}A_\sigma\right)$ term. In the sixth line we apply the product rule for the ordinary derivative, and in the seventh line we reconsolidate the second and fourth terms using $D_{;(\alpha}=\partial_{;(\alpha}-iG_{(\alpha}$. The result is:

$$\begin{aligned}
D_{;(\alpha}\left(\left[D_{;\mu},D_{;\nu)}\right]A_\sigma\right) &= \partial_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)}A^\tau\right)-iG_{(\alpha}\left(R_{|\tau\sigma|\mu\nu)}A^\tau\right)+D_{;(\alpha}\left(V_{[\mu\nu])}A_\sigma\right)\\
&= R_{\tau\sigma(\mu\nu}\partial_{;\alpha)}A^\tau-iG_{(\alpha}R_{|\tau\sigma|\mu\nu)}A^\tau+D_{;(\alpha}\left(V_{[\mu\nu])}A_\sigma\right)\\
&= R_{\tau\sigma(\mu\nu}\partial_{;\alpha)}A^\tau-iR_{\tau\sigma(\mu\nu}G_{\alpha)}A^\tau+D_{;(\alpha}\left(V_{[\mu\nu])}A_\sigma\right)\\
&= R_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau+D_{;(\alpha}\left(V_{[\mu\nu])}A_\sigma\right) \qquad . \\
&= R_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau+\partial_{;(\alpha}\left(V_{[\mu\nu])}A_\sigma\right)-iG_{(\alpha}V_{[\mu\nu])}A_\sigma\\
&= R_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau+\partial_{;(\alpha}V_{[\mu\nu])}A_\sigma+V_{([\mu\nu]}\partial_{\alpha)}A_\sigma-iG_{(\alpha}V_{[\mu\nu])}A_\sigma\\
&= R_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau+\left(D_{;(\alpha}V_{[\mu\nu])}+V_{([\mu\nu]}\partial_{\alpha)}\right)A_\sigma
\end{aligned} \tag{7.13}$$

Now the "odd duck" is the $\left(D_{;(\alpha}V_{[\mu\nu])}+V_{([\mu\nu]}\partial_{\alpha)}\right)A_\sigma$ term which has as mix of gauge-covariant and ordinary-covariant derivatives commuted to opposite sides of $V_{[\mu\nu]}$. But from (7.10) with one index renamed, this is just $D_{;(\alpha}V_{[\mu\nu])}-V_{([\mu\nu]}\partial_{;\alpha)}=-iP_{\mu\nu\alpha}$, which is why we wanted to make the one final connection in (7.10) before turning to $\partial_{;(\alpha}R_{|\tau\sigma|\mu\nu)}=0$. So with the final substitution of (7.10) into (7.13), we obtain:

$$P_{\mu\nu\alpha}A_\sigma = iD_{;(\alpha}\left(\left[D_{;\mu},D_{;\nu)}\right]A_\sigma\right)-iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau = -i\partial_{;(\alpha}\left[G_\mu,G_{\nu)}\right]A_\sigma - iG_{(\alpha}D_{;[\mu}G_{\nu])}A_\sigma . \tag{7.14}$$

This is our final result for the magnetic source density written as an operator operating on any vector A^μ. Above, we have also included the monopole of (4.6). We can also manipulate the indexes to clearly display the spacetime symmetries:

$$g_{\sigma\tau}P_{\mu\nu\alpha}A^\tau = ig_{\sigma\tau}D_{;(\alpha}\left(\left[D_{;\mu},D_{;\nu)}\right]A^\tau\right)-iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau . \tag{7.15}$$

Of course, A^τ represents anything that transforms like a four-vector in spacetime. Among the specific vectors which may be of interest, are yet a fourth gauge covariant derivative $A^\mu \to D^{;\mu}$, and a gauge field $A^\mu \to G^\mu$ (which is implicit in $A^\mu \to D^{;\mu}$). Thus, it helps to use (7.14) to form:

$$D_{;(\alpha}\left(\left[D_{;\mu},D_{;\nu)}\right]D_{;\sigma}\right) = R_{\tau\sigma(\mu\nu}D_{;\alpha)}D^{;\tau}-iP_{\mu\nu\alpha}D_{;\sigma} . \tag{7.16}$$

In particular, this is now an operator identity which tells us what happens when we take four successive gauge covariant derivatives in the $D_{;(\alpha}\left(\left[D_{;\mu},D_{;\nu)}\right]D_{;\sigma}\right)$ cyclic combination.

Finally, in flat spacetime, where $R_{\tau\sigma\mu\nu}=0$ and $D_{;\mu}\to D_\mu$, (7.14) reduces simply to:

$$P_{\mu\nu\alpha}A_\sigma = iD_{(\alpha}\left(\left[D_\mu,D_{\nu)}\right]A_\sigma\right) . \tag{7.17}$$

So this monopole density operating on an arbitrary vector A_σ in the form $P_{\mu\nu\alpha}A_\sigma$ now becomes associated with the particular gauge-covariant triple derivative $iD_{(\alpha}\left(\left[D_\mu, D_{\nu)}\right]A_\sigma\right)$. For a succession of four gauge-covariant derivatives, this is:

$$P_{\mu\nu\alpha}D_\sigma = iD_{(\alpha}\left(\left[D_\mu, D_{\nu)}\right]D_\sigma\right). \tag{7.18}$$

So Hermann Weyl's curvature view of Yang-Mills theory teaches us quite a bit, in particular, about the nature of the Yang-Mills monopole densities. This ought not to be surprising, because the two Bianchi densities $R_{\tau\sigma\mu\nu} + R_{\tau\mu\nu\sigma} + R_{\tau\nu\sigma\mu} = 0$ and $\partial_{;\alpha}R_{\tau\sigma\mu\nu} + \partial_{;\mu}R_{\tau\sigma\nu\alpha} + \partial_{;\nu}R_{\tau\sigma\alpha\mu} = 0$ contain cyclic index structures just as do the monopoles. Above, we have illustrated the curvature *analogy* between gauge theory and gravitation, and embedded these two important identities of spacetime geometry in the Yang-Mills identities (7.14) and (7.15). Based on this embedding, however, we can go even further, to fully unify Yang-Mills gauge theory with classical gravitation via a single classical field equation which combines them both.

8. The Gravitational Field Equation for Yang-Mills Gauge Theory, Inclusive of Maxwell's Electrodynamics

Because the second Bianchi identity $\partial_{;\alpha}R_{\tau\sigma\mu\nu} + \partial_{;\mu}R_{\tau\sigma\nu\alpha} + \partial_{;\nu}R_{\tau\sigma\alpha\mu} = 0$ is embedded in (7.14) and (7.15), there should be some manipulation that will reveal a Yang-Mills analog to the equation $\partial_{;\nu}\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right) = 0$ which underlies gravitational theory. We now deduce that.

We start by manipulating (7.15) according to the following sequence of steps which apply $D_{;(\alpha} = \partial_{;(\alpha} - iG_{(\alpha}$ and the product rule for differentiation:

$$\begin{aligned} g_{\sigma\tau}P_{\mu\nu\alpha}A^\tau &= ig_{\sigma\tau}D_{;(\alpha}\left(\left[D_{;\mu}, D_{;\nu)}\right]A^\tau\right) - iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau \\ &= ig_{\sigma\tau}\partial_{;(\alpha}\left(\left[D_{;\mu}, D_{;\nu)}\right]A^\tau\right) + g_{\sigma\tau}G_{(\alpha}\left(\left[D_{;\mu}, D_{;\nu)}\right]A^\tau\right) - iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau \\ &= ig_{\sigma\tau}\partial_{;(\alpha}\left[D_{;\mu}, D_{;\nu)}\right]A^\tau + ig_{\sigma\tau}\left[D_{;(\mu}, D_{;\nu}\right]\partial_{;\alpha)}A^\tau + g_{\sigma\tau}G_{(\alpha}\left[D_{;\mu}, D_{;\nu)}\right]A^\tau - iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau \\ &= ig_{\sigma\tau}D_{;(\alpha}\left[D_{;\mu}, D_{;\nu)}\right]A^\tau + ig_{\sigma\tau}\left[D_{;(\mu}, D_{;\nu}\right]\partial_{;\alpha)}A^\tau - iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^\tau \end{aligned} . \tag{8.1}$$

Now, because A^τ is just a dummy operand which can be any four-vector, let us just lop it off of (8.1) entirely. The equations on each side of the equal sign will no longer have matching symmetries because $g_{\sigma\tau}$ is symmetric while $R_{\tau\sigma\mu\nu}$ is antisymmetric in these same two indexes. So we shall use a "=" sign, that is, an equal sign in quotes. Thus, we now write:

$$g_{\sigma\tau}P_{\mu\nu\alpha} \text{"="} ig_{\sigma\tau}D_{;(\alpha}\left[D_{;\mu}, D_{;\nu)}\right] + ig_{\sigma\tau}\left[D_{;(\mu}, D_{;\nu}\right]\partial_{;\alpha)} - iR_{\tau\sigma(\mu\nu}D_{;\alpha)}. \tag{8.2}$$

The two sides of this equation are only equal they operate on a vector as in (8.1), or if the symmetries can be restored in some other way. So we will need to now manipulate this in such a way that the symmetries on both sides once again become matching.

First, we fully expand the cyclators in (8.2) to obtain:

$$\begin{aligned} g_{\sigma\tau}P_{\mu\nu\alpha} \text{"}=\text{"} & ig_{\sigma\tau}D_{;\alpha}\left[D_{;\mu},D_{;\nu}\right]+ig_{\sigma\tau}D_{;\mu}\left[D_{;\nu},D_{;\alpha}\right]+ig_{\sigma\tau}D_{;\nu}\left[D_{;\alpha},D_{;\mu}\right] \\ & +ig_{\sigma\tau}\left[D_{;\mu},D_{;\nu}\right]\partial_{;\alpha}+ig_{\sigma\tau}\left[D_{;\nu},D_{;\alpha}\right]\partial_{;\mu}+ig_{\sigma\tau}\left[D_{;\alpha},D_{;\mu}\right]\partial_{;\nu} \\ & -iR_{\tau\sigma\mu\nu}D_{;\alpha}-iR_{\tau\sigma\nu\alpha}D_{;\mu}-iR_{\tau\sigma\alpha\mu}D_{;\nu} \end{aligned} . \tag{8.3}$$

Next, we use the terms $R_{\tau\sigma\mu\nu}D_{;\alpha}$ and the like as a guide and engage in the same manipulations normally used to derive $\partial_{;\nu}\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)=0$ from $\partial_{;\alpha}R_{\tau\sigma\mu\nu}+\partial_{;\mu}R_{\tau\sigma\nu\alpha}+\partial_{;\nu}R_{\tau\sigma\alpha\mu}=0$. We raise $\tau\sigma$ indexes everywhere to put the Riemann tensor into mixed form so we can extract the Ricci tensor. Then we contract one pair of indexes by setting $\nu=\tau$ and we start to reveal the Ricci tensor via $R^{\tau\sigma}{}_{\mu\tau}=R^{\sigma}{}_{\mu}$ including revealing one sign reversal. This yields the intermediate result:

$$\begin{aligned} g^{\tau\sigma}P_{\mu\tau\alpha} \text{"}=\text{"} & ig^{\tau\sigma}D_{;\alpha}\left[D_{;\mu},D_{;\tau}\right]+ig^{\tau\sigma}D_{;\mu}\left[D_{;\tau},D_{;\alpha}\right]+ig^{\tau\sigma}D_{;\tau}\left[D_{;\alpha},D_{;\mu}\right] \\ & +ig^{\tau\sigma}\left[D_{;\mu},D_{;\tau}\right]\partial_{;\alpha}+ig^{\tau\sigma}\left[D_{;\tau},D_{;\alpha}\right]\partial_{;\mu}+ig^{\tau\sigma}\left[D_{;\alpha},D_{;\mu}\right]\partial_{;\tau} \\ & -iR^{\sigma}{}_{\mu}D_{;\alpha}+iR^{\sigma}{}_{\alpha}D_{;\mu}-iR^{\tau\sigma}{}_{\alpha\mu}D_{;\tau} \end{aligned} . \tag{8.4}$$

Now we do a second index contraction by setting $\mu=\sigma$. This yields the Ricci scalar $R^{\sigma}{}_{\sigma}=R$ and allows another application of $R^{\tau\sigma}{}_{\alpha\sigma}=-R^{\tau}{}_{\alpha}$ with a second sign reversal. We then use the $g^{\tau\sigma}$ to raise indexes. Now we have:

$$\begin{aligned} P^{\tau}{}_{\tau\alpha}= & iD_{;\alpha}\left[D^{;\tau},D_{;\tau}\right]+iD^{;\tau}\left[D_{;\tau},D_{;\alpha}\right]+iD_{;\tau}\left[D_{;\alpha},D^{;\tau}\right] \\ & +i\left[D^{;\tau},D_{;\tau}\right]\partial_{;\alpha}+i\left[D_{;\tau},D_{;\alpha}\right]\partial^{;\tau}+i\left[D_{;\alpha},D^{;\tau}\right]\partial_{;\tau} \\ & -iRD_{;\alpha}+iR^{\sigma}{}_{\alpha}D_{;\sigma}+iR^{\tau}{}_{\alpha}D_{;\tau} \end{aligned} . \tag{8.5}$$

We have now removed the quotes from the equal sign, because now the only free index is α and there is no longer a mismatched symmetry. That is, the symmetry became mismatched when we looped off A^{τ} from (8.1) and it became restored when we contracted down to (8.4) which is a vector equation containing one free index α. But given the commutation properties in the above, $P^{\tau}{}_{\tau\alpha}=0$ because it is a third-rank totally antisymmetric tensor, and all of the other terms in the first and second lines also cancel out by inspection. So all that we have left in (8.5) after some very simple rearrangement, and applying the Einstein equation $-\kappa T^{\mu\nu}=R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R$, is:

$$-\kappa T^{\mu\nu}D_{;\nu}=\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)D_{;\nu}=0 . \tag{8.6}$$

<u>This is the gravitational field equation of Yang-Mills theory!</u> It resembles the usual $-\kappa\partial_{;\nu}T^{\mu\nu}=\partial_{;\nu}\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)=0$, but here, we have an operator equation, the derivative is moved to the right (it does not operate to differentiate $R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R$), *and it is a gauge-covariant derivative*. If we want to highlight the nexus to Yang-Mills theory in the clearest way possible, we may expand to above into the form:

$$-\kappa T^{\mu\nu}\left(\partial_{;\nu}-iG_{\nu}\right)=\left(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R\right)\left(\partial_{;\nu}-iG_{\nu}\right)=0\,. \tag{8.7}$$

And, if we then use this to operate on some arbitrary vector A_{σ}, we may further expand this to:

$$\begin{aligned}0&=-\kappa T^{\mu\nu}\left(\partial_{;\nu}A_{\sigma}-iG_{\nu}A_{\sigma}\right)=\left(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R\right)\left(\partial_{;\nu}A_{\sigma}-iG_{\nu}A_{\sigma}\right)\\&=-\kappa T^{\mu\nu}\left(\partial_{\nu}A_{\sigma}-\Gamma^{\tau}{}_{\nu\sigma}A_{\tau}-iG_{\nu}A_{\sigma}\right)=\left(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R\right)\left(\partial_{\nu}A_{\sigma}-\Gamma^{\tau}{}_{\nu\sigma}A_{\tau}-iG_{\nu}A_{\sigma}\right)\\&=-\kappa T^{\mu\nu}\left(\delta^{\tau}{}_{\sigma}\partial_{\nu}-\Gamma^{\tau}{}_{\nu\sigma}-i\delta^{\tau}{}_{\sigma}G_{\nu}\right)A_{\tau}=\left(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R\right)\left(\partial_{\nu}A_{\sigma}-\Gamma^{\tau}{}_{\nu\sigma}A_{\tau}-iG_{\nu}A_{\sigma}\right)\\&=-\kappa T^{\mu\nu}\left(g_{\tau\sigma}\partial_{\nu}-\Gamma_{\tau\nu\sigma}-ig_{\tau\sigma}G_{\nu}\right)A^{\tau}=\left(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R\right)\left(g_{\tau\sigma}\partial_{\nu}-\Gamma_{\tau\nu\sigma}-ig_{\tau\sigma}G_{\nu}\right)A^{\tau}\end{aligned}\,. \tag{8.8}$$

By the connection to $T^{\mu\nu}$, we further come to understand the coupling between gauge fields and source matter.

This brings Hermann Weyl full circle back to Albert Einstein, as there is no more concise way to express the role of geometry in spacetime and in gauge space then through the "Einstein-Weyl" unified field equation $\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)D_{;\nu}=0$. The term $R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R$ emerges from Einstein's understanding of parallel transport and curvature in spacetime, while $D_{\nu}=\partial_{;\nu}-iG_{\nu}$ emerges from Weyl's understanding of .parallel transport and curvature in gauge (phase) space. The contracted combination of $\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)D_{;\nu}=0$ marries the two together into one!

While we have developed the foregoing based on Yang-Mills gauge theory, and have generally regarded $D_{;\nu}=\partial_{;\nu}-iG_{\nu}=\partial_{;\nu}-i\lambda^{i}G^{i}_{\nu}$ to be an NxN matrix, this is not an absolute requirement. Weyl developed $D_{\nu}=\partial_{;\nu}-iG_{\nu}$ twenty five years before Yang and Mills came on the scene. So we can also take the gauge group to be $U(1)_{em}$ of electrodynamics, and we may regard the gauge field G_{ν} as Maxwell's electrodynamic vector potential A_{ν} (now we are *not* taking A_{ν} to be arbitrary but making a specific association with the electromagnetic potential). When we do so, *the geometric operator equation* $\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)\left(\partial_{;\nu}-iA_{\nu}\right)=0$ *now becomes the classical unified field equation for gravitation and electromagnetism*. All of classical field theory is geometry! Quantum field theory then emerges from path integration of the classical fields.

9. The Configuration Space Inverse of the Chromo-Electric Field Equation of Classical Yang-Mills Theory

Much of the focus in the last two sections was centered on the magnetic charge density $P_{\sigma\mu\nu}$, primarily because this has the same index-cyclic, antisymmetric tensor properties as the two Bianchi identities $R_{\tau\sigma\mu\nu}+R_{\tau\mu\nu\sigma}+R_{\tau\nu\sigma\mu}=0$ and $\partial_{;\alpha}R_{\tau\sigma\mu\nu}+\partial_{;\mu}R_{\tau\sigma\nu\alpha}+\partial_{;\nu}R_{\tau\sigma\alpha\mu}=0$ the Jacobian identity $\left[a,\left[b,c\right]\right]+\left[b,\left[c,a\right]\right]+\left[c,\left[a,b\right]\right]=0$ which were central to the development of the classical unified field equation in the various formulations of (8.6) to (8.8). Now it is time to return our focus largely to the field equation (4.3) of the chromo-electric charge density J^ν.

If we compare $J^\nu=\left(g^{\mu\nu}\left(D_{;\sigma}D^{;\sigma}+m^2\right)-D^{;\mu}D^{;\nu}\right)G_\mu$ which is the chromo-electric charge density field equation (4.3) side by side with $P^{\sigma\mu\nu}=-i\left(\partial^{;(\sigma}\left[G^\mu,G^{\nu)}\right]+G^{(\sigma}D^{;[\mu}G^{\nu])}\right)$ which is the chromo-magnetic charge density field equation (4.6) while keeping in mind that the gauge-covariant derivative $D^{;\mu}=\partial^{;\mu}-iG^\mu$, then we notice a remarkable thing: Mathematically, these two non-Abelian Maxwell's equations can be thought of as *a pair of parametric equations in which the gauge field G^μ is itself the parameter*. These means in turn that there is a definitive, albeit complicated relationship between the monopole density $P^{\sigma\mu\nu}$ and the charge density J^ν. As such, we should endeavor to find out more about this relationship. Keep in mind, this would never become a consideration in Abelian electrodynamics, because there, the magnetic sources $P^{\sigma\mu\nu}=0$. But this is *not* the case in Yang-Mills theory.

Additionally, the chromo-magnetic density $P^{\sigma\mu\nu}=-i\left(\partial^{;(\sigma}\left[G^\mu,G^{\nu)}\right]+G^{(\sigma}D^{;[\mu}G^{\nu])}\right)$ of (4.6) looks on the surface like a bundle of gluons G^μ. (Again, we avoid the term "glueball" to avert confusion with specific meanings that have already been given to this term.) But if we take a conservative view of field theory, wherein gauge field are always generated by some source, then the natural progression from (4.6) should be to inquire about the sources from which these gauge fields G^μ originate. Other than the monopole source $P^{\sigma\mu\nu}$, the only other logical source of G^μ is the chromo-electric source density J^ν.

Furthermore, in Dirac theory, an electric source density J^ν may in turn be expressed in terms of fermion wavefunctions ψ. Specifically, Dirac's equation says that $\left(i\gamma^\mu\partial_\mu-m\right)\psi=0$. For the adjoint spinor $\overline{\psi}=\psi^\dagger\gamma^0$ the field equation is $i\partial_\mu\overline{\psi}\gamma^\mu+m\overline{\psi}=0$. Adding yields $\partial_\mu\left(\overline{\psi}\gamma^\mu\psi\right)=0$ as is well known. And because the conserved current is expressed by $\partial_\mu J^\mu=0$, we identify the current density with $J^\mu=\overline{\psi}\gamma^\mu\psi$. In Yang-Mills theory, for a compact, simple gauge group SU(N), this generalizes to $J^\mu=\lambda^i_{AB}J^{i\mu}=\lambda^i_{AB}\overline{\Psi}_C\lambda^i_{CD}\gamma^\mu\Psi_D=\overline{\Psi}\gamma^\mu\Psi$, with Yang-Mills adjoint i and fundamental A,B,C,D indexes explicitly shown for illustration, where $\Psi=\Psi_A$ is an N-component column vector of 4-component elementary Dirac fermion wavefunctions ψ. Thus, $\overline{\Psi}\gamma^\nu\Psi=\left(g^{\mu\nu}\left(D_{;\sigma}D^{;\sigma}+m^2\right)-D^{;\mu}D^{;\nu}\right)G_\mu$ becomes another way to write (4.3). With this progression from $J^{\mu} \rightarrow \overline{\Psi}\gamma^{\mu}\Psi$, the gauge field G^{μ} now is the parameter which specifies a relationship between the magnetic sources $P^{\sigma\mu\nu}$ and the Dirac fermions Ψ. Because we already seen based on some of the symmetries outlined in section 5 that these $P^{\sigma\mu\nu}$ have attributes reminiscent of baryons, this parameterization may provide a way to "populate" these magnetic monopoles $P^{\sigma\mu\nu}$ with fermion eigenstates ψ. If, in turn, these fermions eigenstates exhibit the same symmetries as the quarks that we know reside inside baryons, this would provide support for regarding these ψ as quark wavefunctions, and the $P^{\sigma\mu\nu}$ themselves as baryon densities. So, we shall now proceed along these lines to populate the monopoles with fermions by developing the inverse field equations $G_{\mu} \equiv I_{\tau\mu}J^{\tau}$.

Specifically, we will want to *define* an inverse $I_{\tau\mu}$ such that $G_{\mu} \equiv I_{\tau\mu}J^{\tau}$. Then, we can insert $G_{\mu} = I_{\tau\mu}J^{\tau} = I_{\tau\mu}\overline{\Psi}\gamma^{\tau}\Psi$ into $P^{\sigma\mu\nu} = -i\left(\partial^{;(\sigma}\left[G^{\mu}, G^{\nu)}\right] + G^{(\sigma}D^{;[\mu}G^{\nu])}\right)$ for each occurrence of the gauge field G^{μ}, thereby populating $P^{\sigma\mu\nu}$ with fermions. It helps to briefly review how this inversion is done in electrodynamics, to prepare for the more complicated calculation required for Yang-Mills theory.

In electrodynamics, we use the classical field equation mentioned just after (4.3) to specify this inverse $G_{\mu} \equiv I_{L\tau\mu}J^{\tau}$, namely:

$$J^{\nu} = \left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\mu}\partial^{;\nu}\right)G_{\mu} = \delta^{\nu}{}_{\tau}J^{\tau} \equiv \left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\mu}\partial^{;\nu}\right)I_{L\tau\mu}J^{\tau} \quad . \tag{9.1}$$

We have specifically denoted this inverse $I_{L\tau\mu}$ with a "L" subscript to keep note of the fact that this is the *linear* inverse of Abelian gauge theory. We will shortly derive the more complicated inverse $I_{YM\,\tau\mu}$ which includes all the effects of Yang-Mills theory both linear and non-linear, and then from this will form a $I_{P\tau\mu} \equiv I_{YM\,\tau\mu} - I_{L\tau\mu}$ which tells us the precise portion of the complete Yang-Mills inverse $I_{YM\,\tau\mu}$ arises from the *perturbative* effects which account for the difference between $I_{YM\,\tau\mu}$ and $I_{L\tau\mu}$. This follows the approach introduced prior to (6.10) where we found that the perturbative-only contribution to the current density is $J_P^{\nu} = \left(g^{\mu\nu}V - V^{\mu\nu}\right)G_{\mu}$.

Dropping J^{τ} from the last two terms with index renaming then allows us to sift out:

$$\delta^{\mu}{}_{\nu} \equiv \left(g^{\mu\tau}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\tau}\partial^{;\mu}\right)I_{L\nu\tau} \quad . \tag{9.2}$$

Looking at the momentum space operator $g^{\mu\tau}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\tau}\partial^{;\mu}$, we see that in flat spacetime this will be symmetric in its μ,τ indexes, but in curved spacetime it will not. In curved spacetime, the Riemann tensor $\left[\partial_{;\mu}, \partial_{;\nu}\right]G_{\alpha} \equiv R^{\sigma}{}_{\alpha\mu\nu}G_{\sigma}$ is non-zero as noted just prior to (4.4), and so left-right ordering matters. Especially since the non-Abelian $g^{\mu\tau}\left(D_{;\sigma}D^{;\sigma} + m^2\right) - D^{;\mu}D^{;\tau}$

with $D^{;\mu} = \partial^{;\mu} - iG^{\mu}$ where $G^{\mu} = \lambda^{i}_{AB} G^{i\mu}$ is an NxN matrix for SU(N) is manifestly not μ, τ symmetric *even in flat spacetime*, it will be important to pay attention right away to commutativity issues. One will also discern from this, that except in flat spacetime for Abelian gauge theory, inverse $I_{\nu\tau}$ will be non-symmetric between its ν, τ indexes. Thus, the definitional choice $G_{\nu} \equiv I_{\tau\nu} J^{\tau}$ where the left index in the inverse is summed with the current density is different than the reversed-index definition $G_{\nu} \equiv I_{\nu\tau} J^{\tau}$ in which the right index is so-summed.

Based on the terms in (9.2), we may surmise that $I_{L\nu\tau} \equiv g_{\nu\tau} A + \partial_{;\nu} \partial_{;\tau} B$ will be the general form of the inverse, with $I_{L\nu\tau}$ defined to have the same index ordering as $\partial_{;\nu} \partial_{;\tau}$, and with *A* and *B* being unknowns we shall now deduce. We define *A* and *B* to the right, so that when we insert $I_{L\nu\tau}$ into (9.2) to specify:

$$\delta^{\mu}{}_{\nu} \equiv \left(g^{\mu\tau} \left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right) - \partial^{;\tau} \partial^{;\mu} \right) \left(g_{\nu\tau} A + \partial_{;\nu} \partial_{;\tau} B \right) \ , \tag{9.3}$$

the *A* and *B* will not come between the known terms. Again, this is part of our desire to pay very close attention to commutativity order, which will be especially important when we progress to Yang-Mills theory.

Now we expand (9.3) to obtain:

$$\delta^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} \left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right) A - \partial_{;\nu} \partial^{;\mu} A + \left(\left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right) \partial_{;\nu} \partial^{;\mu} - \partial^{;\tau} \partial^{;\mu} \partial_{;\nu} \partial_{;\tau} \right) B \ , \tag{9.4}$$

where we may freely commute $g^{\mu\nu}$, and where we then make use of $\delta^{\mu}{}_{\nu} = g^{\mu\tau} g_{\nu\tau}$ and also use the remaining metric tensors to raise or lower indexes as appropriate. The first step is to eliminate the $\delta^{\mu}{}_{\nu} \left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right) A$ term by setting $\left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right) A = 1$, and more specifically, by left-multiplying with $\left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right)^{-1}$ to write:

$$A = \left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right)^{-1} \left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right) A = \left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right)^{-1} \ , \tag{9.5}$$

Because $\partial_{;\sigma} \partial^{;\sigma} + m^2$ is not a matrix (shortly, its Yang-Mills counterpart will be), the use of inverses is not required and we can employ the more-common $A = 1 / \left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right)$. But this "overkill" will be important for Yang-Mills theory. Inserting (9.5) back into (9.4) while maintaining all the "overkill" of ordering and taking inverses yields, with some rearrangement:

$$\partial^{;\nu} \partial^{;\mu} \left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right)^{-1} = \left(\left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right) \partial^{;\nu} \partial^{;\mu} - \partial_{;\tau} \partial^{;\mu} \partial^{;\nu} \partial^{;\tau} \right) B \ . \tag{9.6}$$

Multiplying from the left by $\left(\left(\partial_{;\sigma} \partial^{;\sigma} + m^2 \right) \partial^{;\nu} \partial^{;\mu} - \partial_{;\tau} \partial^{;\mu} \partial^{;\nu} \partial^{;\tau} \right)^{-1}$ then yields:

$$B=\left(\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)\partial^{;\nu}\partial^{;\mu}-\partial_{;\tau}\partial^{;\mu}\partial^{;\nu}\partial^{;\tau}\right)^{-1}\partial^{;\nu}\partial^{;\mu}\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)^{-1} . \tag{9.7}$$

Now using (9.5) and (9.7) in $I_{L\nu\tau}\equiv g_{\nu\tau}A+\partial_{;\nu}\partial_{;\tau}B$ we obtain:

$$I_{\nu\tau}=\left[g_{\nu\tau}+\partial_{;\nu}\partial_{;\tau}\left(\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)\partial^{;\alpha}\partial^{;\beta}-\partial_{;\sigma}\partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma}\right)^{-1}\partial^{;\alpha}\partial^{;\beta}\right]\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)^{-1} . \tag{9.8}$$

Since these inverses have a Yang-Mills dimension of NxN=1x1, they are not Yang-Mills matrices and may be placed into denominators in customary manner. Thus (9.8) becomes:

$$I_{L\nu\tau}=\frac{g_{\nu\tau}+\dfrac{\partial_{;\nu}\partial_{;\tau}\partial^{;\alpha}\partial^{;\beta}}{\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)\partial^{;\alpha}\partial^{;\beta}-\partial_{;\sigma}\partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma}}}{\partial_{;\sigma}\partial^{;\sigma}+m^2} . \tag{9.9}$$

In flat spacetime where the derivatives may be freely commuted, we can factor out the $\partial^{;\alpha}\partial^{;\beta}$ terms and which leaves a $\partial_{;\sigma}\partial^{;\sigma}-\partial_{;\sigma}\partial^{;\sigma}=0$ which also zeros out. Thus, we convert to momentum space via $\partial^{\mu}\rightarrow ik^{\mu}$ and add the $+i\varepsilon$ prescription yields the inverse for a massive vector boson, thus obtaining:

$$I_{L\nu\tau}=\frac{g_{\nu\tau}+\dfrac{\partial_{\nu}\partial_{\tau}}{m^2}}{\partial_{\sigma}\partial^{\sigma}+m^2}=\frac{-g_{\nu\tau}+\dfrac{k_{\nu}k_{\tau}}{m^2}}{k_{\sigma}k^{\sigma}-m^2}\overset{+i\varepsilon}{\Rightarrow}\frac{-g_{\nu\tau}+\dfrac{k_{\nu}k_{\tau}}{m^2}}{k_{\sigma}k^{\sigma}-m^2+i\varepsilon} . \tag{9.10}$$

We make note of the fact that up to a factor of *i*, this inverse is identical to the QED propagator $\pi_{\nu\tau}$, i.e., that $\pi_{\nu\tau}=iI_{L\nu\tau}$. Finally, we return to use the above in $G_{\nu}\equiv I_{L\tau\nu}J^{\tau}$ (note reversed index ordering versus (9.10) traceable to (9.2)), which yields:

$$G_{\nu}=\frac{-g_{\tau\nu}+\dfrac{k_{\tau}k_{\nu}}{m^2}}{k_{\sigma}k^{\sigma}-m^2+i\varepsilon}J^{\tau}=-\frac{1}{k_{\sigma}k^{\sigma}-m^2+i\varepsilon}J_{\nu}\overset{m=0}{\Rightarrow}-\frac{1}{k_{\sigma}k^{\sigma}+i\varepsilon}J_{\nu} . \tag{9.11}$$

After a final flat spacetime commutation $\left[\partial_{\nu},\partial_{\tau}\right]=-\left[k_{\nu},k_{\tau}\right]=0$, the final reduction occurs via conservation of charge density $\partial_{\tau}J^{\tau}=0$, which in momentum space, is $k_{\tau}J^{\tau}=0$ (e.g., [2] after I.5(4)).

Now, it is easy to see from (9.10), via $k_{\nu}k_{\tau}/m^2\rightarrow\infty$ hence $I_{L\nu\tau}\rightarrow\infty$ as $m\rightarrow 0$, which is why the configuration space operator $g^{\mu\nu}\partial_{;\sigma}\partial^{;\sigma}-\partial^{;\mu}\partial^{;\nu}$ for a *massless* vector particle in *flat spacetime* has no inverse (e.g., [2] section 3.4). But what happens in curved spacetime, use $+i\varepsilon$, and set $m\rightarrow 0$? This will be instructive for our monetary consideration of Yang-Mills. In this circumstance, using (9.9) in $G_{\nu}=I_{L\tau\nu}J^{\tau}$, the inverse equation corresponding to (9.11) becomes:

$$G_\nu \equiv \frac{g_{\tau\nu} + \dfrac{\partial_{;\tau}\partial_{;\nu}\partial^{;\alpha}\partial^{;\beta}}{\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)\partial^{;\alpha}\partial^{;\beta} - \partial_{;\sigma}\partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma}}}{\partial_{;\sigma}\partial^{;\sigma}+m^2} J^\tau \underset{+i\varepsilon}{\overset{m=0}{\Rightarrow}} \frac{g_{\tau\nu} + \dfrac{\partial_{;\tau}\partial_{;\nu}\partial^{;\alpha}\partial^{;\beta}}{\partial_{;\sigma}\left(\partial^{;\sigma}\partial^{;\alpha}\partial^{;\beta} - \partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma}\right)}}{\partial_{;\sigma}\partial^{;\sigma}+i\varepsilon} . \tag{9.12}$$

None of the reductions of (9.10) or (9.11) occur. To obtain $\partial_{;\nu}\partial^{;\alpha}\partial^{;\beta}\partial_{;\tau}J^\tau = 0$ from $\partial_{;\tau}\partial_{;\nu}\partial^{;\alpha}\partial^{;\beta}J^\tau$ one would need to commute $\partial_{;\tau}$ to the right past all of $\partial_{;\nu}\partial^{;\alpha}\partial^{;\beta}$, generating several new non-vanishing terms containing the Riemann and Ricci tensors. But of particular interest is what happens if we set $m=0$ (and also added $+i\varepsilon$), as we have done on the rightmost expression above. This, of course, describes the photon. Even here, with $m=0$ (so long as we use $+i\varepsilon$), the inverse is only singular in the circumstance where $\partial^{;\sigma}\partial^{;\alpha}\partial^{;\beta} - \partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma} = 0$, i.e., in flat spacetime. In curved spacetime, the commutator $\partial^{;[\sigma}\partial^{;\alpha}\partial^{;\beta]} \neq 0$, and so while the inverse of $g^{\mu\nu}\partial_{;\sigma}\partial^{;\sigma} - \partial^{;\mu}\partial^{;\nu}$ will still become very large in relatively flat regions of spacetime, *so long as there is a modicum of gravitational curvature, formally speaking, the inverse will never become infinite*. In the real physical world – as opposed to the mathematical idealization that is flat spacetime – anywhere there is matter there is gravitation. So in the real physical world where one cannot escape at least some modicum of matter which inherently gravitates, the inverse in (9.12) will always be finite. Of course, we still need to add $+i\varepsilon$ in the bottom denominator, because for a massless photon on-shell, $\partial_{;\sigma}\partial^{;\sigma} \Rightarrow -k_\sigma k^\sigma = 0$, this inverse will still become singular even in curved spacetime. We point this out because these types of behaviors due to non-commuting derivatives will be manifest very pervasively in Yang-Mills theory, and will actually fill the mass gap.

Now we turn back to the Yang-Mills inverses. Here, we start with the classical chromo-electric field strength (4.3) which we cast in a form analogous to (9.1), namely:

$$J^\nu = \left(g^{\mu\nu}\left(D_{;\sigma}D^{;\sigma}+m^2\right) - D^{;\mu}D^{;\nu}\right)G_\mu = \delta^\nu{}_\tau J^\tau \equiv \left(g^{\mu\nu}\left(D_{;\sigma}D^{;\sigma}+m^2\right) - D^{;\mu}D^{;\nu}\right)I_{YM\,\tau\mu}J^\tau , \tag{9.13}$$

where $I_{YM\,\tau\mu}$ is now the Yang-Mills inverse and $G_\mu \equiv I_{YM\,\tau\mu}J^\tau$ to include *all* the effects of Yang-Mills, both linear and perturbative, $I_{YM\,\tau\mu} \equiv I_{L\,\tau\mu} + I_{P\,\tau\mu}$. The calculation then proceeds exactly in the manner of (9.2) to (9.8), but now the "overkill" of being very careful about inverses and left-right ordering is essential. Completely analogously to (9.8), but with the Yang-Mills "minimal coupling" discussed in relation to the "gauge theory on steroids" view of (3.6), with the simple replacement of $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$, we obtain:

$$I_{YM\,\nu\tau} = \left[g_{\nu\tau} + D_{;\nu}D_{;\tau}\left(m^2 D^{;\alpha}D^{;\beta} + D_{;\sigma}D^{;\sigma}D^{;\alpha}D^{;\beta} - D_{;\sigma}D^{;\beta}D^{;\alpha}D^{;\sigma}\right)^{-1} D^{;\alpha}D^{;\beta}\right]\left(D_{;\sigma}D^{;\sigma}+m^2\right)^{-1} . \tag{9.14}$$

Here, not only is the left-right ordering essential because the $G^\mu = \lambda^i_{AB}G^{i\mu}$ are all Yang-Mills matrices, but so is the specification of matrix inverses which are *not ordinary denominators*. To express (9.14) in a way that facilitates visual comparison to (9.9) for Abelian gauge theory, we shall now adopt a "quoted denominator" notion whereby we represent the inverse of any matrix M according to $1/"M"\equiv M^{-1}$, and to keep track of the proper placement of an inverse in the overall series of matrix multiplications, we use a "$\vee$" down-arrow as a placement marker. In this notation, (9.14) now is written as:

$$I_{YM\,\nu\tau}=\frac{g_{\nu\tau}+\dfrac{D_{;\nu}D_{;\tau}\ {}_{\vee}\ D^{;\alpha}D^{;\beta}}{"m^2D^{;\alpha}D^{;\beta}+D_{;\sigma}D^{;\sigma}D^{;\alpha}D^{;\beta}-D_{;\sigma}D^{;\beta}D^{;\alpha}D^{;\sigma}"}\ {}_{\vee}}{"D_{;\sigma}D^{;\sigma}+m^2"}\quad. \tag{9.15}$$

By comparison to (9.9), we see in stark relief the manner in which Yang-Mills gauge theory – at least at the classical level – is simply Gauge theory on steroids with the minimal coupling principle $\partial^\mu\rightarrow D^\mu=\partial^\mu-iG^\mu$. On should note two factorizations which are available in the upper denominator of (9.15). The first two terms may be written as $\left(m^2+D_{;\sigma}D^{;\sigma}\right)D^{;\alpha}D^{;\beta}$ which matches up with the $D^{;\alpha}D^{;\beta}$ in the top numerator. But these do not simply factor out as they did going from (9.9) to (9.10) because of the Yang-Mills matrices and the inverses involved. And the latter two terms in the upper denominator may be written as $D_{;\sigma}\left(D^{;\sigma}D^{;\alpha}D^{;\beta}-D^{;\beta}D^{;\alpha}D^{;\sigma}\right)$. As discussed after (9.12), this helps avert a singular numerator even if we set $m=0$, because this will remain finite to the degree that $D^{;\sigma}D^{;\alpha}D^{;\beta}-D^{;\beta}D^{;\alpha}D^{;\sigma}=D^{;[\sigma}D^{;\alpha}D^{;\beta]}\neq 0$.

We note finally, referring back to sections 7 and 8, that the symmetries of sequences of covariant derivatives is integrally connected to the "curvature view" of Yang-Mills theory and helped us to derive the Einstein-Weyl equation (8.6). Along the way, we obtained several useful identities involving the commutativity properties of taking three of four successive covariant derivatives, specifically (7.9), and (7.13) to (7.17). Clearly, based on these identities, as a general rule $D^{;[\sigma}D^{;\alpha}D^{;\beta]}\neq 0$. Thus, *(9.15) will not become infinite even if we set $m=0$ and even if we do not include $+i\varepsilon$ and even if the gauge particles for which (9.15) is the inverse are placed on shell without $+i\varepsilon$.* This property of (9.15) will become essential for filling the mass gap.

10. Populating Yang-Mills Monopoles with Fermions, and the Recursive Nature of the Yang-Mills: A Sixth View of Yang-Mills

We will examine (9.14) and (9.15) much more closely in the next section. But for the moment, let us return to the complete the goal established at the start of the last section, which is to "populate" these magnetic monopoles $P^{\sigma\mu\nu}$ with fermion eigenstates ψ. Via $G_\mu\equiv I_{YM\,\tau\mu}J^\tau$, we now use the final line of (4.6) to populate the magnetic monopole density (4.6) with inverses $I_{YM\,\tau\mu}$ and current densities J^τ, and we further make use of the Dirac relationship between fermion wavefunctions and chromo-electric current source densities as discussed at the outset of the last section, namely $J^\mu=\overline{\Psi}\gamma^\mu\Psi=\lambda^i_{AB}J^{i\mu}=\lambda^i_{AB}\overline{\Psi}_C\lambda^i_{CD}\gamma^\mu\Psi_D$, to write:

$$\begin{aligned}P^{\sigma\mu\nu} &= -i\left(\partial^{;(\sigma}\left[I_{YM}^{\alpha\mu}J_\alpha, I_{YM}^{\beta\nu)}J_\beta\right] + I_{YM}^{\tau(\sigma}J_\tau D^{;[\mu}I_{YM}^{\beta\nu])}J_\beta\right)\\
&= -i\left(\partial^{;(\sigma}\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_{YM}^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{;[\mu}I_{YM}^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right)\\
&= -i\begin{pmatrix}\partial^{;\sigma}\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu}\overline{\Psi}\gamma_\beta\Psi\right] + \partial^{;\mu}\left[I_{YM}^{\alpha\nu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\sigma}\overline{\Psi}\gamma_\beta\Psi\right] + \partial^{;\nu}\left[I_{YM}^{\alpha\sigma}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\mu}\overline{\Psi}\gamma_\beta\Psi\right]\\
+I_{YM}^{\tau\sigma}\overline{\Psi}\gamma_\tau\Psi D^{;[\mu}I_{YM}^{\beta\nu]}\overline{\Psi}\gamma_\beta\Psi + I_{YM}^{\tau\mu}\overline{\Psi}\gamma_\tau\Psi D^{;[\nu}I_{YM}^{\beta\sigma]}\overline{\Psi}\gamma_\beta\Psi + I_{YM}^{\tau\nu}\overline{\Psi}\gamma_\tau\Psi D^{;[\sigma}I_{YM}^{\beta\mu]}\overline{\Psi}\gamma_\beta\Psi\end{pmatrix}\end{aligned} \quad .(10.1)$$

The Yang-Mills monopole is now fully populated with fermion wavefunctions. We now explicitly can see the fermion sources from which the gauge fields originate. *All* of the non-linear plus non-linear/perturbative (L+P) aspects of Yang-Mills gauge theory are fully included in the above.

In fact, it is critically-important to observe that if we wish to do so, we may explicitly substitute the $I_{YM\nu\tau}$ of (9.14) with a renaming and raising of some indexes into (10.1) to obtain an even more detailed expression. And then, we can employ the gauge-covariant derivative $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$ throughout the inverses to *reintroduce additional gauge fields*. And then, we can use $G_\mu \equiv I_{YM\,\tau\mu}J^\tau$ to replace these new gauge fields with current densities and then use $J^\mu = \overline{\Psi}\gamma^\mu\Psi$ to add more fermion wavefunctions and then use $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$ to again replace gauge fields and repeat this cycle iteratively, recursively, *ad infinitum!* So while (10.1) represents this Yang-Mills monopole in its most compact form, this is a recursive expression because of the fact that if we use $G_\mu \equiv I_{YM\,\tau\mu}J^\tau$ to write gauge field G_τ in terms of the current density J^ν via:

$$G_\tau = \left[g_{\nu\tau} + D_{;\nu}D_{;\tau}\left(m^2 D^{;\alpha}D^{;\beta} + D_{;\sigma}D^{;\sigma}D^{;\alpha}D^{;\beta} - D_{;\sigma}D^{;\beta}D^{;\alpha}D^{;\sigma}\right)^{-1}D^{;\alpha}D^{;\beta}\right]\left(D_{;\sigma}D^{;\sigma} + m^2\right)^{-1}J^\nu \quad .(10.2)$$

we obtain a host of terms with $D^\mu = \partial^\mu - iG^\mu$ *which specify the gauge field* G^μ *recursively in terms of itself.*

In other words, it is very important to observe that (10.2) is <u>*not*</u> a closed expression, because G_μ is self-defined *recursively* in terms of itself. To obtain a closed expression, one would have to repeatedly insert G_μ into itself, *ad infinitum*. It may well be possible to discern the patterns and develop a closed form of (10.2), but for the moment, we simply note this recursion as yet a *sixth view* of Yang-Mills gauge theory. Yang Mills field theory is *1) non-commuting, 2) non-linear, 3) steroidal, 4) perturbative, 5) geometrically-curved and now 6), based on (10.2),* <u>*recursive*</u>. And as noted in section 3, all of these views are alternative, equivalent, and complementary. The $P^{\sigma\mu\nu} = -i\left(\partial^{;(\sigma}\left[I'^{\alpha\mu}J_\alpha, I_{YM}^{\beta\nu)}J_\beta\right] + I'^{\tau(\sigma}J_\tau D^{;[\mu}I_{YM}^{\beta\nu])}J_\beta\right)$ of (10.1), is the compact, irreducible kernel of the recursive specification of the Yang-Mills monopole, with *all* non-linear aspects of Yang-Mills inherently included to infinite recursive order. This is the same monopole (7.10), (7.14) used in section 8, starting with (8.1), to derive the classical unified Einstein-Weyl field equation $-\kappa T^{\mu\nu}D_{;\nu} = \left(R^{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}R\right)D_{;\nu} = 0$ of (8.6).

On page 7 of [1], Jaffe and Witten state:

> "Since the inception of quantum field theory, two central methods have emerged to show the existence of quantum fields on non-compact configuration space (such as Minkowski space). These known methods are (i) Find an exact solution in closed form; (ii) Solve a sequence of approximate problems, and establish convergence of these solutions to the desired limit."

The foregoing suggests a third method which is really a hybrid of (i) and (ii): find an exact *recursive kernel* in closed form, and then expand that kernel in successive iterations to see how the recursion behaves in the limit of infinite recursive nesting.

It will of course be of great interest to examine the behavior of (9.14) a.k.a. (9.15) to see if it is exhibits suitable convergence under infinite recursive nesting, and how this relates to expression obtained during efforts to quantize Yang-Mills. If we look at the numerator *N* in (9.15) and raise one free index to turn $g_{\nu\tau}$ into $\delta_\nu{}^\tau$ which is a unit matrix *I*, we see that this has the skeletal mathematical form $N = 1 + A/B$. Noting that one definition of e^x includes the similar form $e^x = \lim_{n\to\infty}(1+x/n)^n$, and thinking for example how Pe^{RT} expresses the continuous growth of a "principal" *P* at a rate *R* for a time *T* which principal is, in essence, recursively fed into itself for compounding, we may think of e^x as the quintessential, self-feeding, recursive mathematical function. So we ask if there is also an explicitly-*recursive* definition for e^x which might give some insight into how to tame expressions such as (9.14), (9.15). If we define a dummy variable $x \equiv 1 + Bx/n$ and feed this into itself, each time setting *n* to the number of the nesting level, it turns out that as the nesting approaches infinity, we obtain e^B:

$$x = 1 + \frac{Bx}{1} \to 1 + \frac{B(1+\frac{Bx}{2})}{1} \to 1 + \frac{B(1+\frac{B(1+\frac{Bx}{3})}{2})}{1} \to \frac{B(1+\frac{B(1+\frac{B(1+\frac{Bx}{4})}{3})}{2})}{1}\ldots \qquad (10.3)$$
$$= 1 + B + \frac{1}{2!}B^2 + \frac{1}{3!}B^3 + \frac{1}{4!}B^4 x \overset{n\to\infty}{\to} e^B$$

In other words, the infinite recursive nesting of $x \equiv 1 + Ax/n$ with n set to the nesting level is *another way to define* e^B. This is not to say that (9.15) will necessarily turn out to have an exponential form, but rather to point out how a Maclaurin series for e^A may be recursively defined from the recursive kernel $x \equiv 1 + Bx/n$ where *n* is the nesting level. *It would seem a fruitful mathematical exercise to develop similar recursive definitions for other mathematical functions via their series*, and then, armed with those definitions, to take a fresh look at (9.15) and see if that provides further insight into understanding this recursive series and the circumstances under which this series diverges or tractably-converges, and what it looks like in truly closed form.

The other very important insight to carry away from the recursive expression (10.2), in light of (10.3), which is a mathematical insight with possible physical implications is this: In (10.3) x is a "dummy" variable that gets stripped away in the infinite application of recursion. This means that in (10.2) the gauge field G_τ is the dummy variable that will get stripped away by the recursion as the nesting reaches infinity, that that what will remain behind is the single G_τ on the left (10.2) expressed as an infinite series in powers of the source current J^ν. Possibly analogously, when we take a path integral, such as in QED:

$$\begin{aligned} Z &= \int DG_\mu \exp i \int d^4x \left(\tfrac{1}{2} G_\mu \left(g^{\mu\nu} \left(\partial_\sigma \partial^\sigma + m^2 \right) - \partial^\mu \partial^\nu \right) G_\nu - J^\mu G_\mu \right) \\ &\equiv \mathcal{C} \exp\left(iW(J) \right) = \mathcal{C} \exp \left(-\frac{1}{2} i \int \frac{d^4k}{(2\pi)^4} J^\mu \frac{-g_{\mu\nu} + \frac{k_\mu k_\nu}{m^2}}{k_\sigma k^\sigma - m^2} J^\nu \right) \end{aligned} , \tag{10.4}$$

the gauge field G_τ is the variable of integration, it also gets stripped away as the integration takes place, and what is left behind is an infinite series in powers of the source current J^ν.

With this in mind, using what Zee [2] in Appendix A refers to as the "central identity of quantum field theory" (we have reversed the sign for J because we are using the electrodynamic convention in which the units of charge (electrons) are negative whereas Zee uses a positive charge sign convention):

$$\int D\phi \exp\left(-\tfrac{1}{2} \phi \cdot K \cdot \phi - V(\phi) - J \cdot \phi \right) = \mathcal{C} \exp\left(V\left(\delta / \delta J \right) \right) \exp\left(\tfrac{1}{2} J \cdot K^{-1} \cdot J \right), \tag{10.5}$$

it would be a very interesting mathematical exercise to see whether the core Gaussian integral:

$$\int dx \exp\left(-\tfrac{1}{2} A x^2 - Jx \right) = \left(-2\pi / A \right)^{.5} \exp\left(J^2 / 2A \right) \tag{10.6}$$

can be fully reformulated in terms of a recursive functions. As s start toward this, it helps to develop what may be a new mathematical notation to represent this sort of recursive nesting. Analogously to how series are summarized using the symbol $\Sigma_{n=1}^{\infty}$, we shall now create an infinite nest symbol represented by a pair of nested parenthesis $(())_{n=1}^{\infty}$. In the function to be nested, we shall enclose the dummy variable (which was x in (10.3)) in the form $((x))$. This, in this (possibly new) notation, we may write (10.3) in compact form as:

$$e^B \equiv \left(\left(\ \right) \right)_{n=1}^{\infty} \left(1 + B\left((x) \right) / n \right) \tag{10.7}$$

This means that the Gaussian integral (10.6) may be *recursively* written as:

$$\int dx \exp\left(-\tfrac{1}{2}Ax^2 - Jx\right) = \left(-2\pi/A\right)^{.5} \exp\left(J^2/2A\right) e^{J^2/2A} = \left(-2\pi/A\right)^{.5} \left(\left(\ \right)\right)_{n=1}^{\infty} \left(1 + \frac{J^2\left(\left(x\right)\right)}{2An}\right) \quad (10.8)$$

where x (an abstracted gauge field), which is a *dummy variable of integration*, is what gets stripped away during the infinite recursion as a *dummy variable of recursion*. It is not at this point clear whether this sort of recursive analysis can be helpful in breaking through to enable an exact, analytical path integral quantization of Yang-Mills theory in closed form, but it is worthwhile to see what contributions can be made by a recursive analysis in which the physical field to be subjected to path integration is instead regarded as a dummy variable in a recursive expansion. What is absolutely clear, however, is that Yang-Mills theory, in the form of (10.1) and (10.2), forces upon us the need to analyze, understand, and better develop its recursive features, which are yet a sixth view of Yang Mills in which all of the non-linearities are expressed and developed through recursive mathematics.

It is also worth observing that the magnetic monopole (10.1), now populated with fermions (which we will later show are quarks) is really, at bottom, a non-Abelian combination of *both* of Maxwell's classical equations (4.1) and (4.2) into a *single equation*. Specifically, the chromo-electric charge equation combined with Dirac wavefunction theory via $J^\nu = D_{;\mu}F^{\mu\nu} = D_{;\mu}D^{[\mu}G^{\nu]} = \overline{\Psi}\gamma^\nu\Psi$ is represented in inverse form via (10.2) and then inserted into the monopole density (4.6) to arrive at (10.1). Einstein, in his final paper [18] at page 159 points out the "surprising" finding that Maxwell's two equations, taken together, possess a field strength $z_1 = 12$ which is the exact same strength as the equation $R_{\mu\nu} = 0$ for pure geometry. This would suggest (10.1), which is a field equation relating all three of $J^\nu = \overline{\Psi}\gamma^\nu\Psi$, $P^{\sigma\mu\nu}$ and G^μ (two sources and one gauge field) to one another, and which merges both of Maxwell's equations together, will also have a strength of $z_1 = 12$ interrelating each of its $i = 1,2,3...N^2 - 1$ Abelian sources $J^{i\nu}$, $P^{i\sigma\mu\nu}$ and fields $G^{i\mu}$.

The final, very important point to note is that because of its origin in (4.2) and (4.6) as a Yang-Mills monopole, (10.1) contains *three additive terms* in index-cyclic configuration of the form $\partial^{;(\sigma}\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right]$, and similarly $I_{YM}^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{;[\mu}I_{YM}^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi$. Further, $\Psi = \Psi_A$ is an *N*-component column vector of 4-component Dirac spinor wavefunctions ψ for whatever gauge group SU(N) we choose to employ. To this moment, we have been exploring Yang-Mills gauge theory *in general*, but have made no selection of any specific gauge group. Now that is about to change. Because $P^{\sigma\mu\nu}$ is the density of a *single* magnetic monopole, $P^{\sigma\mu\nu}$ must be regarded as a system which contains these $\Psi = \Psi_A$. But by virtue of the three additive terms, it would appear to contain three such fermions. This was the source of the "three-ness" discussed at some length toward the end of section 5. Dirac-Fermi-Pauli exclusion tells us to make certain that that the fermions in each of these terms are in different eigenstates, so that this monopole system does not contain any two fermions in the same state. Because there are three additive terms, the *smallest* group we are permitted to choose is SU(3). By Occam's razor, we make this smallest permitted selection, *and so do choose SU(3).*

Once we choose SU(3), we place each of the now-three ψ of $\Psi = \Psi_A$, A=1,2,3 into a distinct eigenstate. In order to discuss this, we need to name these states. So we will name them Red, Green and Blue, and denote them ψ_R, ψ_G and ψ_B. And with that, we move from Yang-Mills gauge theory generally, to Chromodynamics specifically. And while we start with three fermions ψ_R, ψ_G and ψ_B which we shall soon establish may be interpreted as quarks, the recursive nature of (10.1) via (10.2) and $D^\mu = \partial^\mu - iG^\mu$ and $G_\mu = I'_{\tau\mu} J^\tau = I'_{\tau\mu} \overline{\Psi} \gamma^\tau \Psi$ ensures us the monopole system of (10.1) will be teeming with non-linear physics and many additional quarks and antiquarks that arise at the first, second, thousandth, and millionth recursive order, which may thought of as various "excited" baryonic states. This will all be developed in detail in section 12.

In this light, and as we shall detail in the forthcoming development, QCD is not a theory of first principle, it is a theory of second principle. The theory of first principle is Maxwell's electrodynamics as extended into non-Abelian domains by Yang-Mills gauge theory. QCD is then *derived by deduction* as a consequence of enforcing exclusion for the fermions contained in the non-vanishing magnetic monopoles of Yang-Mills gauge theory, and choosing a gauge group no larger than is necessary to enforce this exclusion. In the process, we fully explain why nature chooses three quarks per baryon (in the "ground" state of zero-recursive order) rather than some other number.

Now we turn to two specific showings: First, we shall show how the relation (10.2), which of course is contained to infinite recursive order in (10.1), fills the mass gap. To preview: if we set $m = 0$ in (10.1), due the non-commuting nature of Yang-Mills theory, we still retain terms which create mass-like effects and which, because of the specific matrix inversion $\left(D_{;\sigma} D^{;\sigma} + m^2\right)^{-1}$ in (10.2), yield a mass eigenvalue spectrum, which one expects will come to be associated with the masses of the observed mesons. Second, as has already been developed to some degree in section 5, we shall show from a more formal standpoint how and why (10.1) contains all of the expected color symmetries of a baryon, and at the same time confines its quarks and its gauge fields, while permitting the flux of colorless quark combinations that we observe in the form of mesons.

11. The Mass Gap Solution

Let us now show how the solution to the mass gap that is embedded in the infinitely-recursive equation (9.14), which we shall discuss using the more "user-friendly" representation (9.15).

The configuration space inverse (9.15) represents *all* of the non-linear, recursive features of Yang-Mills theory. As we have done previously, let us now identify how much of this inverse arises strictly from the perturbations P. As we did earlier with (6.10), let us used the framework YM=L+P (total Yang-Mills effects are the sum of linear effects plus perturbative effects) to calculate $I_{P\nu\tau} = I_{YM\,\nu\tau} - I_{L\nu\tau}$, which is simply the difference between the entire, holistic (see [2] at page 356) inverse (9.15) and the linear inverse (9.9). So what we shall now be studying is what Yang-Mills theory brings to the table (perturbations in the perturbative view), *above and beyond* what Abelian gauge theories such as electrodynamics already bring to the table. So we can study *only* the impact of Yang-Mills theory separated from any impact due to spacetime curvature, we represent both of (9.9) and (9.15) in flat spacetime, and so turn the gravitationally-covariant derivatives $\partial^{;\sigma}$ into ordinary ones ∂^{σ}. Thus, we form:

$$I_{P\nu\tau} = I_{YM\,\nu\tau} - I_{L\nu\tau}$$
$$= \frac{g_{\nu\tau} + \dfrac{D_\nu D_\tau \vee D^\alpha D^\beta}{"m^2 D^\alpha D^\beta + D_\sigma D^\sigma D^\alpha D^\beta - D_\sigma D^\beta D^\alpha D^\sigma" \vee}}{"D_\sigma D^\sigma + m^2"} - \frac{g_{\nu\tau} + \dfrac{\partial_\nu \partial_\tau \partial^\alpha \partial^\beta}{m^2 \partial^\alpha \partial^\beta + \partial_\sigma \partial^\sigma \partial^\alpha \partial^\beta - \partial_\sigma \partial^\beta \partial^\alpha \partial^\sigma}}{\partial_\sigma \partial^\sigma + m^2} \quad .(11.1)$$

The ordinary derivatives in the right hand term commute and the denominators are real denominators, not matrix inverses. So the above readily reduces to (see (9.9) to (9.10) where we did the same reduction earlier):

$$I_{P\nu\tau} = I_{YM\,\nu\tau} - I_{L\nu\tau}$$
$$= \frac{g_{\nu\tau} + \dfrac{D_\nu D_\tau \vee D^\alpha D^\beta}{"m^2 D^\alpha D^\beta + D_\sigma D^\sigma D^\alpha D^\beta - D_\sigma D^\beta D^\alpha D^\sigma" \vee}}{"D_\sigma D^\sigma + m^2"} - \frac{g_{\nu\tau} + \dfrac{\partial_\nu \partial_\tau}{m^2}}{\partial_\sigma \partial^\sigma + m^2} \,. \quad (11.2)$$

The term on the right, of course, is the inverse for a massive spin-1 vector field (vector boson), it is identical to what we found in (9.10), and when we convert over to momentum space, it is the same thing as the vector boson propagator up to a factor of *i*, $\pi_{\nu\tau} = iI_{L\nu\tau}$. The QED path integration which establishes that $\pi_{\nu\tau} = iI_{L\nu\tau}$, is displayed in (10.4). The term $D_\sigma D^{[\sigma} D^\alpha D^{\beta]} = D_\sigma D^\sigma D^\alpha D^\beta - D_\sigma D^\beta D^\alpha D^\sigma$, which will be at the heart of the discussion to follow, contains a succession of four covariant derivatives, and as we can see from the identities developed in section 7 and especially (7.18), this term $D_\sigma D^{[\sigma} D^\alpha D^{\beta]}$ is non-vanishing everywhere there are non-zero perturbations.

Now let us return to (6.6) for two successive gauge-covariant derivatives, and write this in momentum space in flat spacetime via $\partial^\mu \to ik^\mu$, as

$$D^\mu D^\nu = -k^\mu k^\nu + V^{\mu\nu} = -k^\mu k^\nu + k^\mu G^\nu + G^\mu k^\nu - G^\mu G^\nu \,, \quad (11.3)$$

which also means that:

$$V^{\mu\nu} = k^\mu G^\nu + G^\mu k^\nu - G^\mu G^\nu \,. \quad (11.4)$$

So we expand the various $D^\mu D^\nu = -k^\mu k^\nu + V^{\mu\nu}$ in (11.2) and convert into momentum space, to obtain:

$$\begin{aligned}
I_{P\nu\tau} &= I_{YM\,\nu\tau} - I_{L\nu\tau} \\
&= \frac{-g_{\nu\tau} + \dfrac{\left(k_\nu k_\tau - V_{\nu\tau}\right)_\vee \left(-k^\alpha k^\beta + V^{\alpha\beta}\right)}{\text{"}m^2\left(-k^\alpha k^\beta + V^{\alpha\beta}\right) + \left(k_\sigma k^\sigma - V_\sigma{}^\sigma\right)\left(k^\alpha k^\beta - V^{\alpha\beta}\right) - \left(k_\sigma k^\beta - V_\sigma{}^\beta\right)\left(k^\alpha k^\sigma - V^{\alpha\sigma}\right)\text{"}_\vee}}{\text{"}k_\sigma k^\sigma - V_\sigma{}^\sigma - m^2\text{"}} \\
&\quad - \frac{-g_{\nu\tau} + \dfrac{k_\nu k_\tau}{m^2}}{k_\sigma k^\sigma - m^2}
\end{aligned} \quad . \tag{11.5}$$

We of course see the perturbative-only inverse $I_{P\nu\tau} \to 0$ if all the perturbations are turned off, $V^{\alpha\beta} \to 0$, as is to be expected. Again, we are now largely working in the perturbative view of Yang-Mills.

What we now wish to consider is this: In the full Yang Mills inverse $I_{YM\,\nu\tau}$ in (11.5), the m^2 is from the Proca mass of the Yang-Mills gauge bosons, introduced by hand back in (4.3). That mass has followed us all the way through the development since, but as originally pointed out, it is a red flag mass that we want to eventually be able to zero out and – if there are massive particles to be found the in the physics we are describing – to be able to reintroduce those masses in some other way without ruining the gauge invariance and the renormalizability of the theory. So now, that time has come to set the Proca mass in $I_{YM\,\nu\tau}$ to zero. But we shall leave the Proca mass as is in $I_{L\nu\tau}$ for reasons to be momentarily discussed. With setting $m^2 = 0$ in $I_{YM\,\nu\tau}$, the above now reduces to:

$$\begin{aligned}
I_{P\nu\tau} &= I_{YM\,\nu\tau} - I_{L\nu\tau} \\
&= \frac{-g_{\nu\tau} + \dfrac{\left(k_\nu k_\tau - V_{\nu\tau}\right)_\vee \left(-k^\alpha k^\beta + V^{\alpha\beta}\right)}{\text{"}\left(k_\sigma k^\sigma - V_\sigma{}^\sigma\right)\left(k^\alpha k^\beta - V^{\alpha\beta}\right) - \left(k_\sigma k^\beta - V_\sigma{}^\beta\right)\left(k^\alpha k^\sigma - V^{\alpha\sigma}\right)\text{"}_\vee}}{\text{"}k_\sigma k^\sigma - V_\sigma{}^\sigma\text{"}} - \frac{-g_{\nu\tau} + \dfrac{k_\nu k_\tau}{m^2}}{k_\sigma k^\sigma - m^2}
\end{aligned} \quad . \tag{11.6}$$

While we are at it, let us even go a step further, by setting the now-massless gauge bosons in $I_{YM\,\nu\tau}$ to be on mass shell, with $k_\sigma k^\sigma = 0$ (which means that the term $k_\sigma k^\beta k^\alpha k^\sigma \to 0$ because the k^σ can commute since we have assumed flat spacetime to isolate the effects of Yang-Mills all by themselves), while at the same time adding $+i\varepsilon$ to the linear inverse $I_{L\nu\tau}$ and also introducing the gauge number ξ, which for $\xi = 1$ is the Feynman gauge and for $\xi = 0$ is the Landau gauge. This gauge number is associated with in the Faddeev-Popov method and was originally developed by Feynman, see, e.g., [2] section III.4. The latter $\xi = 0$ is the gauge of (11.6). Let us also raise the free ν index everywhere. Thus, (11.16) now becomes:

$$I_P{}^\nu{}_\tau = I_{YM}{}^\nu{}_\tau - I_L{}^\nu{}_\tau$$
$$= \frac{-\delta^\nu{}_\tau + \dfrac{\left(k^\nu k_\tau - V_{\nu\tau}\right)_\vee \left(-k^\alpha k^\beta + V^{\alpha\beta}\right)}{"V_\sigma{}^\sigma\left(V^{\alpha\beta} - k^\alpha k^\beta\right) + k_\sigma k^\beta V^{\alpha\sigma} - V_\sigma{}^\beta V^{\alpha\sigma} + V_\sigma{}^\beta k^\alpha k^\sigma"_\vee}}{"-V_\sigma{}^\sigma"} - \frac{-\delta^\nu{}_\tau + (1-\xi)\dfrac{k^\nu k_\tau}{m^2}}{k_\sigma k^\sigma - m^2 + i\varepsilon} . \quad (11.7)$$

And, to simplify our consideration of $I_L{}^\nu{}_\tau$ a bit, let us choose the Feynman gauge $\xi = 1$ which is what transpires anyway the moment one contracts the inverse $I_L{}^\nu{}_\tau$ with a current density via $k_\tau J^\tau = 0$, see (9.11). Thus, (11.7) now becomes:

$$I_P{}^\nu{}_\tau = I_{YM}{}^\nu{}_\tau - I_L{}^\nu{}_\tau = \frac{-\delta^\nu{}_\tau + \dfrac{\left(k^\nu k_\tau - V^\nu{}_\tau\right)_\vee \left(-k^\alpha k^\beta + V^{\alpha\beta}\right)}{"V_\sigma{}^\sigma\left(V^{\alpha\beta} - k^\alpha k^\beta\right) + k_\sigma k^\beta V^{\alpha\sigma} - V_\sigma{}^\beta V^{\alpha\sigma} + V_\sigma{}^\beta k^\alpha k^\sigma"_\vee}}{"-V_\sigma{}^\sigma"} - \frac{-\delta^\nu{}_\tau}{k_\sigma k^\sigma - m^2 + i\varepsilon} . \quad (11.8)$$

Now we arrive at the point: even after we set the Proca mass to zero to keep the Yang-Mills gauge bosons massless and preserve renormalizability, and even after we further set those zero-mass gauge bosons on-shell, so long as the perturbations $V^{\alpha\beta}$ and $V_\sigma{}^\sigma$ are not zero – which means that so long as Yang-Mills theory is doing something more than Abelian gauge theory – the inverse $I_{YM\,\nu\tau}$ remains entirely finite and well-behaved. We do not need the Proca mass at all, and we do not even need $+i\varepsilon$ to avoid the pole that occurs in $I_{L\nu\tau}$ when $k_\sigma k^\sigma - m^2 = 0$ (or when $k_\sigma k^\sigma = 0$ with $m^2 = 0$). The $1/"V_\sigma{}^\sigma" = \left(V_\sigma{}^\sigma\right)^{-1} = \left(k_\sigma G^\sigma + G_\sigma k^\sigma - G_\sigma G^\sigma\right)^{-1}$ term keeps $I_{YM\,\nu\tau}$ well-behaved in exactly the same way that $k_\sigma k^\sigma - m^2 + i\varepsilon$ keeps the linear $I_{L\nu\tau}$ well-behaved. But – at the heart of the matter – $1/"V_\sigma{}^\sigma" = \left(V_\sigma{}^\sigma\right)^{-1} = \left(k_\sigma G^\sigma + G_\sigma k^\sigma - G_\sigma G^\sigma\right)^{-1}$ is an NxN *matrix inverse* that arises with *no artifice* from the essential non-linear core of Yang-Mills theory. In contrast, in $k_\sigma k^\sigma - m^2 + i\varepsilon$, the m^2 is a renormalization-destroying Proca mass which has us asking why, for example, the strong interaction can be a short range interaction even though its gauge boson masses are zero which means we cannot introduce a Proca mass even though we need a Proca mass to make the strong interaction short range and give the inverse / propagator $\pi_{\nu\tau} = iI_{L\nu\tau}$ a non-exploding inverse. And in further contrast, $+i\varepsilon$ is another artifice introduced by hand, to avoid the pole of an on-shell boson. Similarly, as we even saw following (9.12), the moment we set $m^2 = 0$, the numerator term $k_\nu k_\tau / m^2 \to \infty$ in $I_{L\nu\tau}$, unless the spacetime is curved. Here, where we are considering Yang-Mills alone and have removed any effects of gravitational curvature, the corresponding "denominator" in (11.7), $\left(V_\sigma{}^\sigma\left(V^{\alpha\beta} - k^\alpha k^\beta\right) + k_\sigma k^\beta V^{\alpha\sigma} - V_\sigma{}^\beta V^{\alpha\sigma} + V_\sigma{}^\beta k^\alpha k^\sigma\right)^{-1}$, plays the analogous role to the spacetime curvature , and is perfectly well-behaved so long as the perturbations $V^{\alpha\beta}$ and $V_\sigma{}^\sigma$ are not zero, which is exactly what Yang-Mills theory is all about.

So, now, to the mass gap: The Klein Gordon equation (6.1) for a massless scalar field ϕ with gauge symmetry, plus a hand-added Proca mass term for a vector boson with mass, has an associated Lagrangian density (every Lagrangian density is multiplied by 2 in Yang Mills because of the generator normalization $Tr\left(\lambda^i\lambda^j\right)=\frac{1}{2}\delta^{ij}$, see the start of section 3):

$$\begin{aligned}\mathfrak{L}&=\left(D_\mu\phi\right)^\dagger\left(D^\mu\phi\right)-m^2G_\mu G^\mu=\phi\left(\overleftarrow{\partial}_\mu-iG_\mu\right)\left(\partial^\mu-iG^\mu\right)\phi-m^2G_\mu G^\mu\\&=\phi\overleftarrow{\partial}_\mu\partial^\mu\phi-i\phi G^\mu\partial_\mu\phi-i\phi\overleftarrow{\partial}_\mu G^\mu\phi-\phi G_\mu G^\mu\phi-m^2G_\mu G^\mu\end{aligned}\tag{11.9}$$

Above, we represent $\left(\left(\partial_\mu-iG_\mu\right)\phi\right)^\dagger=\phi\left(\overleftarrow{\partial}_\mu-iG_\mu\right)$ due to the hermicity of the gauge fields $G_\mu=\lambda^iG^i_\mu$ which is in turn due to $\lambda^i=\lambda^{i\dagger}$ for the Yang-Mills generators. (While we are here, contrast $\left(D_\mu\phi\right)^\dagger D^\mu\phi$ above to one possible use of the Einstein-Weyl equation (8.6) so as to operate on a scalar field, namely, $\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)D_{;\nu}\varphi=0$.) Although the only ingredients we started with in (11.9) were a scalar ϕ for which we took the gauge-covariant derivative $D^\mu\phi$, we ended up with a term $\phi G_\mu G^\mu\phi$. When we then expand the scalar around the vacuum using a Higgs fields in the form $\phi=v+h(x)+...$ and rescale $G_\mu\to gG_\mu$ to explicitly show the gauge coupling, this gauge-created term:

$$-g^2\phi G_\mu G^\mu\phi=-g^2\left(v+h+...\right)G_\mu G^\mu\left(v+h+...\right)=-\left(vg\right)^2G_\mu G^\mu-g^2\left(2vh+h^2+...\right)G_\mu G^\mu\tag{11.10}$$

reveals the term $-\left(vg\right)^2G_\mu G^\mu$. So now (11.9) contains $-\left(vg\right)^2G_\mu G^\mu-m^2G_\mu G^\mu$. But the term $m^2G_\mu G^\mu$ was introduced by hand with a Proca mass and it ruins the gauge symmetry. The term $-\left(vg\right)^2G_\mu G^\mu$, on the other hand, is a direct result of the gauge symmetry. In fact, the gauge symmetry would be ruined if we did *not* have this term. So we remove the Proca mass (set it to zero) and in its place we regard the term $-\left(vg\right)^2G_\mu G^\mu$ to represent the massive boson and vg to represent the mass of the boson. The experimental confirmation of electroweak theory, of course, validates this result, and at the same time, by using $-\left(vg\right)^2G_\mu G^\mu$ rather than $m^2G_\mu G^\mu$ as the boson mass term, we keep the gauge theory remains renormalizable.

The exact same sort of thing is happening in (11.8). Based on what we know from Abelian gauge theory, we have come to expect that massive vector bosons will have a propagator $\pi_{\nu\tau}=iI_{L\nu\tau}$. The term $I_{L\nu\tau}=-i\pi_{\nu\tau}$ in (11.8) is completely analogous to the term $m^2G_\mu G^\mu$ in (11.9). Each contains a hand-added, renormalization-destroying Proca mass. And (11.8) does (11.9) one better, because it also has a hand-added $+i\varepsilon$ to ensure that the world does not come to an end when a boson is on-shell. But in strong interaction theory, we expect the gauge bosons to be massless. Were we to set m=0 in the $I_{L\nu\tau}$ of (11.7) before we gauged out this term in

(11.8), everything would blow up. Were we to set the boson on-shell in $I_{L\nu\tau}$ and not use $+i\varepsilon$, everything would blow up. But the compete inverse in Yang-Mills theory is $I_{YM\,\nu\tau}$, not $I_{L\nu\tau} = -i\pi_{\nu\tau}$. So $I_{YM\,\nu\tau}$, not $I_{L\nu\tau}$, is the inverse in which we should set *m*=0. And while we are at it, if we want the gauge bosons to be on-shell, $I_{YM\,\nu\tau}$ is also the inverse in which we should set $k_\sigma k^\sigma = 0$. In (11.7) we have already done all of this. The mass is zero, the bosons are on-shell, and we have done nothing by hand that is artificial. And what great catastrophe has befallen $I_{YM\,\nu\tau}$ in (11.7)? Absolutely none! This remains a completely finite, well-behaved *matrix* expression, so long as $V^{\alpha\sigma} \neq 0$ and $V_\sigma{}^\sigma \neq 0$. But where and how, exactly, mathematically, do we fill the mass gap?

This is where the matrix expressions and the inverses come in. Written out expressly in terms of matrices and inverses, (11.8) really says:

$$\begin{aligned}&I_P{}^\nu{}_{\tau AB} = I_{YM}{}^\nu{}_{\tau AB} - I_L{}^\nu{}_\tau \delta_{AB} = \\ &\left(-\delta^\nu{}_\tau + \left(k^\nu k_\tau - V^\nu{}_\tau\right)\left(V_\sigma{}^\sigma\left(V^{\alpha\beta} - k^\alpha k^\beta\right) + k_\sigma k^\beta V^{\alpha\sigma} - V_\sigma{}^\beta V^{\alpha\sigma} + V_\sigma{}^\beta k^\alpha k^\sigma\right)^{-1}\left(-k^\alpha k^\beta + V^{\alpha\beta}\right)\right)\left(-V_\sigma{}^\sigma\right)^{-1}{}_{AB} \\ &-\left(-\delta^\nu{}_\tau\right)/\left(k_\sigma k^\sigma - m^2 + i\varepsilon\right)\delta_{AB}\end{aligned} \quad .(11.11)$$

We have taken special pains to make explicit, the NxN matrix structure, noting that $I_{YM}{}^\nu{}_{\tau AB}$ is a complete, non-commuting, rather complicated NxN Yang-Mills matrix for SU(N), and that $I_L{}^\nu{}_\tau = -\delta^\nu{}_\tau / \left(k_\sigma k^\sigma - m^2 + i\varepsilon\right)$ is not a Yang-Mills matrix . Rather, when we subtract $I_L{}^\nu{}_\tau$ from $I_P{}^\nu{}_\tau$, we must put $I_L{}^\nu{}_\tau$ into the diagonal positions of the Yang-Mills unit matrix δ_{AB}, thus forming $I_L{}^\nu{}_\tau \delta_{AB}$.

But (11.7) is in the form of an eigenvalue equation for the matrix $I_{YM}{}^\nu{}_{\tau AB}$. So if we use this to operate on any Yang-Mills column vector V_B, then $I_L{}^\nu{}_\tau$ will represent the eigenvalues, i.e., the *observables*, of the matrix $I_{YM}{}^\nu{}_{\tau AB}$. But we don't even need to posit a vector V_B because we may obtain these eigenvalues directly from (11.11) via the eigenvalue equation $\left|M - I\lambda\right| = \det\left(M - I\lambda\right) = 0$ which uses the determinant of a matrix *M* to compute its eigenvalues *λ*, and which equation for (11.11) takes the form:

$$\left|I_P{}^\nu{}_{\tau AB}\right| = \left|I_{YM}{}^\nu{}_{\tau AB} - I_L{}^\nu{}_\tau \delta_{AB}\right| = 0 \quad . \tag{11.12}$$

That is it! Once we deduce a non-zero eigenvalue $I_{L\nu\tau}$ via the above from some perturbations $V^{\alpha\sigma} \neq 0$ and $V_\sigma{}^\sigma \neq 0$ in $I_{YM}{}^\nu{}_{\tau AB}$, we then know that the mass *m* will be related to this by:

$$\frac{-\delta^\nu{}_\tau}{k_\sigma k^\sigma - m^2 + i\varepsilon} = I_L{}^\nu{}_\tau \quad . \tag{11.13}$$

In this way, we may deduce both the mass *m* and, if an eigenvalue $I_L{}^\nu{}_\tau$ is a *complex number* with any imaginary component (which it may be because the λ^i generators systemically generate complex numbers once one takes a matrix *inverse* such as $\left(V_\sigma{}^\sigma\right)^{-1}$), the imaginary magnitude $+i\varepsilon$ which corresponds not to the mass – but to a *half-life*. (See, e.g., [**Error! Bookmark not defined.**], page 150.)

So, we turn to the mass gap problem in [1], which states at page 3:

> ". . . for QCD to describe the strong force successfully . . . It must have a "mass gap;" namely there must be some constant Δ > 0 such that every excitation of the vacuum has energy at least Δ,"

and which at page 6 then sets forth the problem:

> "Prove that for any compact simple gauge group G, a non-trivial quantum Yang–Mills theory exists on R4 and has a mass gap > 0 . . . namely there must be some constant Δ > 0 such that every excitation of the vacuum has energy at least Δ . . ."

The solution to the mass gap is as follows: For a compact simple gauge group G which may be "*any*" gauge group SU(N) with $N \ge 2$ and generators λ^i and gauge bosons $G^\mu = \lambda^i G^{i\mu}$, the complete, holistic, non-Abelian, non-linear classical inverse $I_{YM\,\nu\tau}$ associated with these gauge fields G^μ and defined by $G_\mu \equiv I_{YM\,\tau\mu} J^\tau$, with a hand-added Proca mass *m*, will be the $I_{YM\,\nu\tau}$ included in (11.1) generally, and included in (11.2) in flat spacetime. As pointed out already, the term $D_\sigma D^{[\sigma} D^\alpha D^{\beta]}$ in (11.2) is non-vanishing. To maintain the renormalizability of gauge group G, we *must* set this Proca mass to zero, as we do in (11.6). This means that the gauge bosons are now massless. If one takes the gauge group to be $SU(3)_C$ then the gauge bosons are gluons and these gluons are now massless. But we are in no way restricted to $SU(3)_C$ or to any other specific gauge group G. For good measure, though not essential, we even place the gauge bosons on-shell as in (11.7).

Now that the gauge bosons are massless, the question becomes how, for every excitation of the vacuum, "there *must* be constant Δ > 0 such that every excitation of the vacuum has energy at least Δ." The "excitations of the vacuum," in Yang-Mills, are the perturbations $V^{\mu\nu} = k^\mu G^\nu + G^\mu k^\nu - G^\mu G^\nu$ of (11.4). For every such perturbation / excitation, $V^{\mu\nu} \ne 0$ and $V_\sigma{}^\sigma \ne 0$, by definition. Wherever $0 < V^{\mu\nu} < \infty$ and $0 < V_\sigma{}^\sigma < \infty$, the matrix $I_{YM\,\nu\tau}$ will be finite and well behaved, and the eigenvalues of $I_{YM}{}^\nu{}_\tau$ obtained through the eigenvalue equation (11.12) will be finite and non-zero and given by $I_L{}^\nu{}_\tau$. These eigenvalues, which are physical observables, may, in the process, also be complex. These eigenvalues in turn, are related to boson masses and lifetimes via (11.13). This means that the mass *m* in (11.13) will also be non-zero, that is, will have a value > Δ where Δ is some non-zero value, *notwithstanding the fact that* *we have set m=0 in (11.6).* And because this mass is contained within an inverse $I_L{}^{\nu}{}_{\tau}$ which is an eigenvalue of $I_{YM}{}^{\nu}{}_{\tau}$, this mass is deducible (as are possible non-infinite lifetimes) via (11.13). This works for *any* compact simple gauge group G, which is to say, at no point in this *completely general* development have we assumed or needed to assume one particular group over any other. (Though as we have pointed out toward the end of the last section based on (11.1), Yang-Mills monopoles give us reason for regarding SU(3) as a particularly important group, which will be developed further in the next section.)

The mass *m* in $I_L{}^{\nu}{}_{\tau}$ in (11.13) is similar to $m^2 G_\mu G^\mu$ in (11.9). It is a hand-added version of a mass that we observe in the physical world but may not put into the theory by hand without ruining the renormalizability of the theory. So we look for ways for this "anticipated" mass to be revealed by the theory in some other way. In (11.12), this mass associated with this mass gap is revealed in the theory because the excitations in (11.11) give this mass non-zero eigenvalues via (11.13) and the non-zero eigenvalues $I_L{}^{\nu}{}_{\tau}$ which are the reciprocals of $k_\sigma k^\sigma - m^2 + i\varepsilon$, even though the gauge boson masses have been set to zero. If we stay on-shell, then these eigenvalues $I_L{}^{\nu}{}_{\tau}$ are simply the reciprocals of $-m^2 + i\varepsilon$, which is a pure mass number with infinite lifetime for real eigenvalues, a pure mass number and finite lifetime number for complex eigenvalues. The mass gap is filled, and we then have the basis for explaining why Yang-Mills interactions – most notably the strong interaction – are able to have a short range which requires massive gauge bosons and that the same time have gauge bosons which are massless. The mass gap is filled because (11.12) "reveals" a non-zero mass in the inverse (11.13) without ever introducing that mass by hand, in exactly the same way that (11.10) reveals a non-zero mass in the Lagrangian density (11.9) without ever introducing that mass by hand.

Having now filled the mass gap, we return to show why it is that the Yang-Mills monopoles (10.1) have all the chromodyanmic color symmetries required of a baryon, and at the same time confines its quarks and its gauge fields, while permitting the flux of colorless quark combinations that we observe in the form of mesons. Given the mass gap now filled here, this in turn would mean that the nuclear forces associated with these baryon/monopoles have short range.

12. Populating Yang-Mills Monopoles with Fermions to Reveal the Chromodynamic Symmetries of Baryons and Mesons

Let us return to the monopole (10.1) which we have populated with the fermion sources Ψ from which its gauge fields arise. As we did in the last section, we write the inverses in the form $I_{YM\,\nu\tau} = I_{L\nu\tau} + I_{P\nu\tau}$ to show the sum of the linear plus perturbative contributions to the complete Yang-Mills inverse $I_{YM\,\nu\tau}$. And, let us stay in flat spacetime and thereby set all spacetime-covariant derivatives to ordinary derivatives, $\partial_{;\mu} \to \partial_\mu$. So, substituting $I_{YM\,\nu\tau} = I_{L\nu\tau} + I_{P\nu\tau}$ into (10.1) yields:

$$\begin{aligned}
P^{\sigma\mu\nu} &= -i\left(\partial^{(\sigma}\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_{YM}^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_{YM}^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&= -i\left(\partial^{(\sigma}\left[\left(I_L^{\alpha\mu}+I_P^{\alpha\mu}\right)\overline{\Psi}\gamma_\alpha\Psi, \left(I_L^{\beta\nu)}+I_P^{\beta\nu)}\right)\overline{\Psi}\gamma_\beta\Psi\right] + \left(I_L^{\tau(\sigma}+I_P^{\tau(\sigma}\right)\overline{\Psi}\gamma_\tau\Psi D^{[\mu}\left(I_L^{\beta\nu])}+I_P^{\beta\nu])}\right)\overline{\Psi}\gamma_\beta\Psi\right) \\
&= -i\left(\partial^{(\sigma}\left[I_L^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_L^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_L^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_L^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&\quad -i\left(\partial^{(\sigma}\left[I_L^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_P^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_L^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_P^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&\quad -i\left(\partial^{(\sigma}\left[I_P^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_L^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_P^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_L^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&\quad -i\left(\partial^{(\sigma}\left[I_P^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_P^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_P^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_P^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&\equiv P_{LL}^{\sigma\mu\nu} + P_{LP}^{\sigma\mu\nu} + P_{PL}^{\sigma\mu\nu} + P_{PP}^{\sigma\mu\nu}
\end{aligned} \quad .(12.1)$$

At the end, we have respectively denoted each of the four main terms as $P_{LL}^{\sigma\mu\nu}$, $P_{LP}^{\sigma\mu\nu}$, $P_{PL}^{\sigma\mu\nu}$ and $P_{PP}^{\sigma\mu\nu}$ to specify the four combinations of linear (L) and perturbative (P) inverses they contain. Because our goal is to understand the symmetry properties of $P^{\sigma\mu\nu}$, let us zero in on the $P_{LL}^{\sigma\mu\nu}$ terms, which we segregate out as:

$$P_{LL}^{\sigma\mu\nu} = -i\left(\partial^{(\sigma}\left[I_L^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_L^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_L^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_L^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \equiv P_{LL1}^{\sigma\mu\nu} + P_{LL2}^{\sigma\mu\nu} . \quad (12.2)$$

We have further used $P_{LL1}^{\sigma\mu\nu}$ and $P_{LL2}^{\sigma\mu\nu}$ to separately denote each of the terms in the above. Zeroing in even more, let's look at:

$$P_{LL1}^{\sigma\mu\nu} = -i\partial^{(\sigma}\left[I_L^{\alpha\mu}J_\alpha, I_L^{\beta\nu)}J_\beta\right] = -i\partial^{(\sigma}\left[I_L^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_L^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right], \quad (12.3)$$

where we have also used $J_\alpha = \overline{\Psi}\gamma_\alpha\Psi$ to consolidate back to a source density. Now, let us substitute the linear inverse derived in (9.10) sans $+i\varepsilon$ into the above, to obtain:

$$\begin{aligned}
P_{LL1}^{\sigma\mu\nu} &= -i\partial^{(\sigma}\left[I_L^{\alpha\mu}J_\alpha, I_L^{\beta\nu)}J_\beta\right] = -i\partial^{(\sigma}\left[\frac{-g^{\alpha\mu}+k^\alpha k^\mu / m^2}{k_\sigma k^\sigma - m^2}J_\alpha, \frac{-g^{\beta\nu)}+k^\beta k^{\nu)} / m^2}{k_\sigma k^\sigma - m^2}J_\beta\right] \\
&= i\partial^{(\sigma}\left[\frac{J^\mu}{k_\sigma k^\sigma - m^2}, \frac{J^{\nu)}}{k_\sigma k^\sigma - m^2}\right] = i\partial^{(\sigma}\left[\frac{\overline{\Psi}\gamma^\mu\Psi}{k_\sigma k^\sigma - m^2}, \frac{\overline{\Psi}\gamma^{\nu)}\Psi}{k_\sigma k^\sigma - m^2}\right]
\end{aligned}, \quad (12.4)$$

The terms $k^\alpha k^\mu / m^2$ etc. are eliminated via the conserved current $k^\alpha J_\alpha = 0$, see (9.11), and then we raise the index on the current and the $-g^{\alpha\mu}$ absorbed into the current flips the overall sign. Finally, let us expand the cylcator in the final expression in (12.4) as such:

$$P_{LL1}^{\sigma\mu\nu}=i\left(\partial^\sigma\left[\frac{\overline{\Psi}\gamma^\mu\Psi}{k_\sigma k^\sigma-m^2},\frac{\overline{\Psi}\gamma^\nu\Psi}{k_\sigma k^\sigma-m^2}\right]+\partial^\mu\left[\frac{\overline{\Psi}\gamma^\nu\Psi}{k_\sigma k^\sigma-m^2},\frac{\overline{\Psi}\gamma^\sigma\Psi}{k_\sigma k^\sigma-m^2}\right]+\partial^\nu\left[\frac{\overline{\Psi}\gamma^\sigma\Psi}{k_\sigma k^\sigma-m^2},\frac{\overline{\Psi}\gamma^\mu\Psi}{k_\sigma k^\sigma-m^2}\right]\right)$$
$$=i\frac{1}{k_\sigma k^\sigma-m^2}\left(\partial^\sigma\frac{\overline{\Psi}\gamma^{[\mu}\Psi\overline{\Psi}\gamma^{\nu]}\Psi}{k_\sigma k^\sigma-m^2}+\partial^\mu\frac{\overline{\Psi}\gamma^{[\nu}\Psi\overline{\Psi}\gamma^{\sigma]}\Psi}{k_\sigma k^\sigma-m^2}+\partial^\nu\frac{\overline{\Psi}\gamma^{[\sigma}\Psi\overline{\Psi}\gamma^{\mu]}\Psi}{k_\sigma k^\sigma-m^2}\right) \tag{12.5}$$

Now, let's develop the above in some depth. The development to follow parallels sections 2, 3 and 5 of [8], but streamlines and simplifies that development considerably and, perhaps more importantly, put that development in the overall context being developed in this paper.

To start, we note the spin sum relationship which is often normalized (but not here) such that $N^2=\mathrm{E}+m$. This spin sum *prior to normalization* is:

$$\sum\nolimits_{\text{spins}}U\overline{U}=\frac{N^2}{E+m}\left(\not{p}+m\right). \tag{12.6}$$

Also seeing the emergent $\Psi\overline{\Psi}=U\overline{U}$ in each of the three terms in (12.5), we take the $\Psi\overline{\Psi}=U\overline{U}$ in all three of these terms in (12.5), and use (12.6) to write:

$$P_{LL1}^{\sigma\mu\nu}=i\frac{1}{k_\sigma k^\sigma-m^2}\frac{N^2}{E+m}\left(\partial^\sigma\frac{\overline{\Psi}\gamma^{[\mu}\left(\not{p}+m\right)\gamma^{\nu]}\Psi}{k_\sigma k^\sigma-m^2}+\partial^\mu\frac{\overline{\Psi}\gamma^{[\nu}\left(\not{p}+m\right)\gamma^{\sigma]}\Psi}{k_\sigma k^\sigma-m^2}+\partial^\nu\frac{\overline{\Psi}\gamma^{[\sigma}\left(\not{p}+m\right)\gamma^{\mu]}\Psi}{k_\sigma k^\sigma-m^2}\right) \tag{12.7}$$

Next, we keep in mind that the fermion propagator

$$\frac{\not{p}+m}{p^\tau p_\tau-m^2}=\frac{\not{p}+m}{\left(\not{p}+m\right)\left(\not{p}-m\right)}=\left(\not{p}-m\right)^{-1}, \tag{12.8}$$

while also noting the appearance of $\left(\not{p}+m\right)/\left(k_\tau k^\tau-m^2\right)$ throughout (12.7) which is very similar in form to (12.8). So, if we can find some rationale (see section 3 of [8]) to associate the k^τ with p^τ which is the four-momentum of the fermion, then we will have established that there are propagating fermion wavefunctions populating the monopole $P^{\sigma\mu\nu}$. Observing that the $1/\left(k_\tau k^\tau-m^2\right)$ represents propagation for a Proca-massive vector boson with *three* degrees of freedom and that fermions have *four* degrees of freedom, we shift one degree of freedom from the leading $1/\left(k_\tau k^\tau-m^2\right)$ over to the fermions by setting m=0 to turn that leading term into a *massless* boson propagator. That is, for each term in (12.7), we shift:

$$\frac{1}{k_\tau k^\tau-m^2}\partial^\sigma\frac{\overline{\psi}\gamma^{[\mu}\left(\not{p}+m\right)\gamma^{\nu]}\psi}{k_\tau k^\tau-m^2}\Rightarrow\frac{1}{k_\tau k^\tau}\partial^\sigma\frac{\overline{\psi}\gamma^{[\mu}\left(\not{p}+m\right)\gamma^{\nu]}\psi}{p_\tau p^\tau-m^2}. \tag{12.9}$$

and now take p^τ to represent the fermion four-momentum. It should be clear that both parts of (3.17) contain a total of six degrees of freedom; they have just been shifted from a 3+3 to a 2+4 configuration not dissimilarly to how a degree of freedom is shifted from a Higgs scalar to a massless gauge boson to create massive vector bosons using the Goldstone mechanism. Thus, following this shifting of degrees of freedom, (12.7) becomes:

$$P_{LL1}^{\sigma\mu\nu} = i\frac{1}{k_\sigma k^\sigma}\frac{N^2}{E+m}\left(\partial^\sigma\frac{\overline{\Psi}\gamma^{[\mu}\left(\not{p}+m\right)\gamma^{\nu]}\Psi}{p_\sigma p^\sigma - m^2} + \partial^\mu\frac{\overline{\Psi}\gamma^{[\nu}\left(\not{p}+m\right)\gamma^{\sigma]}\Psi}{p_\sigma p^\sigma - m^2} + \partial^\nu\frac{\overline{\Psi}\gamma^{[\sigma}\left(\not{p}+m\right)\gamma^{\mu]}\Psi}{p_\sigma p^\sigma - m^2}\right) \tag{12.10}$$

If we now normalize such that $N^2 = \left(E+m\right)k_\tau k^\tau$, then via (12.8) we can reduce (12.10) to:

$$P_{LL1}^{\sigma\mu\nu} = i\left(\partial^\sigma\left(\overline{\Psi}\gamma^{[\mu}\left(\not{p}-m\right)^{-1}\gamma^{\nu]}\Psi\right) + \partial^\mu\left(\overline{\Psi}\gamma^{[\nu}\left(\not{p}-m\right)^{-1}\gamma^{\sigma]}\Psi\right) + \partial^\nu\left(\overline{\Psi}\gamma^{[\sigma}\left(\not{p}-m\right)^{-1}\gamma^{\mu]}\Psi\right)\right) \tag{12.11}$$

which contains three additive terms each containing a propagating fermion wavefunction.

Now, we resume the discussion toward the end of section 10 where we noted that because $P^{\sigma\mu\nu}$ is the density of a *single* magnetic monopole, $P^{\sigma\mu\nu}$ must be regarded as a system which contains these $\Psi = \Psi_A$, and that Dirac-Fermi-Pauli exclusion tells us to make certain that that the fermions in each of these terms are in different eigenstates. Thus, as already stated, because there are three additive terms, the *smallest* group we are permitted to choose is SU(3), and by Occam's razor, we make this smallest permitted selection, *and so do choose SU(3)*. So let us now implement this.

As already stated at the end of section 10, once we choose SU(3), we place each of the now-three ψ of $\Psi = \Psi_A$, A=1,2,3 into a distinct eigenstate. In order to discuss this, we need to name these states. So we will name them Red, Green and Blue, and denote them ψ_R, ψ_G and ψ_B. The generators are $\lambda^i; i = 1,2,3...8$, the eight gauge bosons are $G^\mu = \lambda^i G^{i\mu}$, and the three fermion eigenstates are ψ_R, ψ_G and ψ_B. Specifically, we define these eigenstates as:

$$\Psi_1 \equiv \left|\lambda^8 = \tfrac{1}{\sqrt{3}}; \lambda^3 = 0\right\rangle = \begin{pmatrix}\psi_R\\0\\0\end{pmatrix}; \Psi_2 \equiv \left|\lambda^8 = -\tfrac{1}{2\sqrt{3}}; \lambda^3 = \tfrac{1}{2}\right\rangle = \begin{pmatrix}0\\\psi_G\\0\end{pmatrix}; \Psi_3 \equiv \left|\lambda^8 = -\tfrac{1}{2\sqrt{3}}; \lambda^3 = -\tfrac{1}{2}\right\rangle = \begin{pmatrix}0\\0\\\psi_B\end{pmatrix}. \tag{12.12}$$

This means that:

$$\Psi_1\overline{\Psi}_1 = \begin{pmatrix}\psi_R\overline{\psi}_R & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{pmatrix}; \quad \Psi_2\overline{\Psi}_2 = \begin{pmatrix}0 & 0 & 0\\0 & \psi_G\overline{\psi}_G & 0\\0 & 0 & 0\end{pmatrix}; \quad \Psi_3\overline{\Psi}_3 = \begin{pmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & \psi_B\overline{\psi}_B\end{pmatrix}. \tag{12.13}$$

Then we go back and use (12.13) to display the explicit 3x3 matrix character of $P_{LL1}^{\sigma\mu\nu} = P_{LL1\,AB}^{\sigma\mu\nu}$ in (12.5):

$$P^{\sigma\mu\nu}_{LL1\ AB} = i\frac{1}{k_\tau k^\tau - m^2}\begin{pmatrix} \partial^\sigma \frac{\overline{\psi}_R \gamma^{[\mu}\psi_R \overline{\psi}_R \gamma^{\nu]}\psi_R}{k_\sigma k^\sigma - m^2} & 0 & 0 \\ 0 & \partial^\mu \frac{\overline{\psi}_G \gamma^{[\nu}\psi_G \overline{\psi}_G \gamma^{\sigma]}\psi_G}{k_\sigma k^\sigma - m^2} & 0 \\ 0 & 0 & \partial^\nu \frac{\overline{\psi}_B \gamma^{[\sigma}\psi_B \overline{\psi}_B \gamma^{\mu]}\psi_B}{k_\sigma k^\sigma - m^2} \end{pmatrix}. \quad (12.14)$$

Then, repeating the same steps that brought us from (12.5) to (12.11), we may turn this into:

$$P^{\sigma\mu\nu}_{LL1\ AB} = i\begin{pmatrix} \partial^\sigma\left(\overline{\psi}_R\gamma^{[\mu}\left(\not{p}_R - m_R\right)^{-1}\gamma^{\nu]}\psi_R\right) & 0 & 0 \\ 0 & \partial^\mu\left(\overline{\psi}_G\gamma^{[\nu}\left(\not{p}_G - m_G\right)^{-1}\gamma^{\sigma]}\psi_G\right) & 0 \\ 0 & 0 & \partial^\nu\left(\overline{\psi}_B\gamma^{[\sigma}\left(\not{p}_B - m_B\right)^{-1}\gamma^{\mu]}\psi_B\right) \end{pmatrix}. \quad (12.15)$$

where $\not{p}_C, m_C; C = R, G, B$ now represent the daggered momentum $\not{p} = \gamma^\tau p_\tau$ and mass m of each of each of the three fermion eigenstates. The trace equation $\mathrm{Tr}P^{\sigma\mu\nu}_{LL1} = P^{\sigma\mu\nu}_{LL1\ AA}$ is then easily deduced to be:

$$\mathrm{Tr}P^{\sigma\mu\nu}_{LL1} = i\left(\partial^\sigma\left(\overline{\psi}_R\gamma^{[\mu}\left(\not{p}_R - m_R\right)^{-1}\gamma^{\nu]}\psi_R\right) + \partial^\mu\left(\overline{\psi}_G\gamma^{[\nu}\left(\not{p}_G - m_G\right)^{-1}\gamma^{\sigma]}\psi_G\right) + \partial^\nu\left(\overline{\psi}_B\gamma^{[\sigma}\left(\not{p}_B - m_B\right)^{-1}\gamma^{\mu]}\psi_B\right)\right). \quad (12.16)$$

This is now the fully-developed Yang-Mills magnetic monopole term $\mathrm{Tr}P^{\sigma\mu\nu}_{LL1\ AB}$, populated with three colored quarks, and it is formally equivalent to [5.5] of [8]. There are of course other terms that we see in (12.1) and (12.2), but we are working with this specific term because it most clearly displays the chromodynamic symmetries of the monopole $P^{\sigma\mu\nu}$. Although we are working with the one term $\mathrm{Tr}P^{\sigma\mu\nu}_{LL1\ AB}$, the assignment (12.12) is systemic: with (12.12), every single Ψ in the complete monopole $P^{\sigma\mu\nu}$ of (12.1) has been turned into an SU(3) column vector with three color eigenstates.

If we now associate each color wavefunction with the spacetime index in the related ∂^σ operator in (12.16), i.e., $\sigma \sim R$, $\mu \sim G$ and $\nu \sim B$, and keeping in mind that $\mathrm{Tr}P^{\sigma\mu\nu}_{LL1}$ is antisymmetric in all spacetime indexes, we express this antisymmetry with wedge products as $\sigma \wedge \mu \wedge \nu \sim R \wedge G \wedge B = R[G,B] + G[B,R] + B[R,G]$. *This is the exact colorless wavefunction that is expected of a baryon.* Indeed, the antisymmetric character of the spacetime indexes in a magnetic monopole should have been a good tipoff that magnetic monopoles would naturally make good baryons. We now may assert that this Yang-Mills monopole has the exact colorless QCD symmetry required of a baryon.

Furthermore, if we apply Gauss' / Stokes' theorem to (12.16) and also include from (5.3) in trace form the finding that $\oiint \mathrm{Tr}G^2 = 3\oiint \mathrm{Tr}\left[G^\mu, G^\nu\right] dx_\mu dx_\nu$, we find that:

$$\begin{aligned}&\iiint \mathrm{Tr}P_{LL1} = \oiint \mathrm{Tr}F_{LL1} = -i\oiint \mathrm{Tr}G^2{}_{LL1} = -3i\oiint \mathrm{Tr}\left[G^\mu, G^\nu\right]_{LL1} dx_\mu dx_\nu \\ &= i\oiint \left(\overline{\psi}_R \gamma^{[\mu}\left(\not{p}_R - m_R\right)^{-1}\gamma^{\nu]}\psi_R + \overline{\psi}_G \gamma^{[\mu}\left(\not{p}_G - m_G\right)^{-1}\gamma^{\nu]}\psi_G + \overline{\psi}_B \gamma^{[\mu}\left(\not{p}_B - m_B\right)^{-1}\gamma^{\nu]}\psi_B\right) dx_\mu dx_\nu\end{aligned} \quad .(12.17)$$

What is the color wavefunction for the $-3i\left[G^\mu, G^\nu\right]$ entities? By inspection, $\overline{R}R + \overline{G}G + \overline{B}B$. So *quarks do <u>not</u> net flow in and out of closed two-dimensional surfaces surrounding* P_{LL1}*, except in the colorless combination of a meson!* So (3.25) validates the suspicion expressed at the end of section III.3 that the appearance of a "3" in front of $\left[G^\mu, G^\nu\right]$ has something to do with there being three colors of quark inside the magnetic monopole.

Of course, (12.17) does beg the question of what flows in and out of the complete monopole (12.1), because (12.17) only considers the term P_{LL1}. So if we go back to (12.1) to apply Gauss'/Stokes' theorem, we obtain:

$$i\iiint P = \oiint I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi dx_\mu dx_\nu + \iiint I_{YM}^{\tau\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu} I_{YM}^{\beta\nu]}\overline{\Psi}\gamma_\beta\Psi dx_\sigma dx_\mu dx_\nu . \quad (12.18)$$

The first term in (12.1), because of the leading $\partial^{(\sigma}$ in (12.1) is fully integrable via Gauss'/Stokes theorem. The second term in (12.1) is *not* integrable, and so it tells us about all of the physics that is confined inside the overall volume of the monopole. But the point made by (12.17), is that whatever does flow across a closed surface pursuant to $\oiint I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi dx_\mu dx_\nu$ in the above, will have the color wavefunction $\overline{R}R + \overline{G}G + \overline{B}B$ of a meson!

So returning to the MIT bag model as discussed in section 5, we now see that for the magnetic monopole (12.1) with surface flux shown in the first term in (12.18), 1) the color wavefunction is that of a baryon, namely $R[G,B]+G[B,R]+B[R,G]$; 2) from (5.4) and (5.5), $\oiint \mathrm{Gluons} = 0$; 3) from (12.17), $\oiint \mathrm{Mesons} \neq 0$ and 4) $\oiint \mathrm{Quarks}=0$ except in the colorless combination $\overline{R}R + \overline{G}G + \overline{B}B$ of a meson. Thus, on a formal basis, with the MIT Bag Model telling us to look at what flows across the surface of any theoretical entity proposed to be a baryon, and we see that the Yang-Mills magnetic monopole has *precisely the required formal symmetries and boundary flows required for a baryon.*

On page 3 of [1], Jaffe and Witten note that QCD:

> ". . . must have "quark confinement," that is, even though the theory is described in terms of elementary fields, such as the quark fields, that transform non-trivially under SU(3), the physical particle states—such as the proton, neutron, and pion—are SU(3)-invariant."

Equation (12.16) shows how the magnetic monopoles of Yang-Mills, with an antisymmetric color wavefunction $R[G,B]+G[B,R]+B[R,G]$, are indeed SU(3) invariant, notwithstanding that the individual fermion eigenstates transform non-trivially under SU(3). This makes the monopoles them well-suited to represent the physical particle states such as protons and neutrons, and makes the fermion eigenstates well-suited to represent quark fields. We further see from (12.17) that all the flux across a closed surface of the monopole has the symmetric color wavefunction $\overline{R}R + \overline{G}G + \overline{B}B$ which is also SU(3) invariant and so make the physical particle states which the spacetime geometry does permit to net flow across closed surfaces well-suited to represent mesons including the pion. And in the process, as discussed at the end of section 10, QCD itself is fully reproduced, but it not a theory of first principle, but rather a secondary theory derived by deduction from Maxwell's electrodynamics as extended into non-Abelian domains by Yang-Mills gauge theory.

Of course, if we wish to associate these magnetic monopole with physical baryons, we still need to make these monopoles topologically stable and see how to use them to represent protons and neutrons which are the most important baryons, see section 6 through 8 of [8], and we need to calculate their energies to see if they make sense in relation to empirical data, see sections 11 and 12 of [8]. Insofar as topological stability, we simply note that the trace equation (12.16) is non-vanishing, but that $\mathrm{Tr}P^{\sigma\mu\nu} = \mathrm{Tr}\left(\lambda^i_{AB} P^{i\sigma\mu\nu}\right) = 0$ *if we regard the gauge group as SU(3)*, because all of λ^i are traceless. In other words, the left and right sides of (12.16) do not match up because one side is traceless and the other is not, *if we assume the simple group SU(3)*. It is on this basis that we introduce the product group $SU(3)_C \times U(1)_{B-L}$, and then obtain the monopole (12.16) (and generally, (12.1)) from the spontaneous symmetry breaking of larger SU(4) gauge groups with a $B-L$ (baryon minus lepton number) generator which yields the quantum numbers required to turn these baryons into proton and neutrons and ensure that these magnetic monopoles are topologically stable. These details are in sections 6 through 8 of [8], they fully apply to the development here, and so they need little if any elaboration or modification here.

More to be added.

References

[1] A. Jaffe and E. Witten, *Quantum Yang–Mills Theory*, Clay Mathematics Institute (2000)
[2] Zee, A., *Quantum Field Theory in a Nutshell*, Princeton (2003)
[3] K. G. Wilson, *Quarks and strings on a lattice*, in New Phenomena In Subnuclear Physics, Proceedings of the 1975 Erice School, A. Zichichi, ed., Plenum Press, New York, 1977.
[4] Niels Bohr Library & Archives, *Oral History Transcript — Dr. Donald F. Geesaman*, http://www.aip.org/history/ohilist/35671.html (2012)
[5] R. Feynman, *The quantitative behavior of Yang–Mills theory in 2+1 dimensions*, Nucl. Phys. B188 (1981), 479-512.
[6] I. M. Singer, *The geometry of the orbit space for nonabelian gauge theories*, Phys. Scripta 24 (1981), 817-820.
[7] A. Einstein, *On the Method of Theoretical Physics*, The Herbert Spencer Lecture, delivered at Oxford (10 June 1933); also published in *Philosophy of Science*, Vol. 1, No. 2 (April 1934), pp. 163-169., p. 165
[8] J. Yablon, *Why Baryons Are Yang-Mills Magnetic Monopoles*, Hadronic Journal, Volume 35, Number 4, 401-468 (2012) Link: http://www.hadronicpress.com/issues/HJ/VOL35/HJ-35-4.pdf
[9] Yang, C. N.; Mills, R. (1954). *Conservation of Isotopic Spin and Isotopic Gauge Invariance*. *Physical Review* 96 (1): 191–195. Bibcode:1954PhRv...96..191Y. doi:10.1103/PhysRev.96.191
[10] H.C. Ohanian, *Gravitation and Spacetime*, Norton (1976)

[11] Misner, C. W., Thorne, K. S., and Wheeler, J. A., *Gravitation*, W. H. Freeman & Co. (1973)
[12] Close, F. E., *An Introduction to Quarks and Partons*, Academic Press (1979)
[13] H. Weyl, *Gravitation and Electricity*, Sitzungsber. Preuss. Akad.Wiss., 465-480. (1918).
[14] H. Weyl, *Space-Time-Matter* (1918)
[15] H. Weyl, Electron und Gravitation, Zeit. f. Physik, 56 (1929), 330
[16] Minkowski, Hermann (1908/9), *Space and Time*, Physikalische Zeitschrift 10 (1909) 75–88
[17] A. Einstein, *On the Electrodynamics of Moving Bodies*, *Annalen der Physik* 17 (1905) 891
[18] A. Einstein, *Relativistic Theory of the Non-Symmetric Field*, in *The Meaning of Relativity*, Princeton University Press, December 1954