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

Jay R. Yablon
Schenectady, New York 12309
jyablon@nycap.rr.com

November 7, 2013

Abstract: We show how $SU(3)_C$ chromodynamics, which is the theory of strong interactions, is a corollary theory emerging naturally from the combination of nothing other than Maxwell / Weyl gauge theory with Yang-Mills theory. In the process, we show not only the emergence from the Maxwell / Yang-Mills combination of all that is to be expected from $SU(3)_C$ chromodynamics, but additionally, we show how the observed baryons containing three colored quarks in the ground state are the magnetic charges of Yang-Mills gauge theory and how these magnetic charges naturally confine their quarks and gluons but do pass mesons in order to interact. That is, we explain quark and gluon confinement and how it is that strong interactions are mediated by mesons but not gauge fields. Additionally, we demonstrate how the inherent non-linearity of Yang-Mills theory may be used to solve the "mass gap" problem and yield a nuclear interaction that is short range notwithstanding its being based on massless gluon gauge fields. We further demonstrate the origin of "chiral symmetry breaking" in strong interactions. We find that the non-linear nature of Yang-Mills theory contains a recursive aspect which provides a useful tool for solving the Yang-Mills path integral in order to exactly, analytically arrive at quantum Yang-Mills theory. As a result of further developing Weyl's original geometric view of gauge theory, we uncover a classical field equation unifying gravitational theory with Weyl's gauge theory including both its Maxwell / Abelian and Yang-Mills variants, at the level of the Einstein equation for gravitation. Finally, we use the recursive aspects of Yang-Mills theory to develop and solve an exact, closed recursive path integral for Quantum Yang-Mills Theory and thereby prove the existence of a non-trivial quantum Yang–Mills theory on R^4 for any simple gauge group G.

PACS: 12.38.Aw; 12.40.-y; 14.20.-c; 14.40.-n

Contents

1. Introduction

In this paper we study the strong "chromodynamic" interactions for which the Yang-Mills gauge group is $SU(3)_C$. But contrary to how chromodynamic interactions are commonly approached, we make no *a priori* supposition about Yang-Mills $SU(3)_C$ being the theory of strong interactions. We simply postulate that Maxwell's $U(1)_{em}$ electrodynamics is a correct theory of nature and that any other non-gravitational interactions have the exact same form as electrodynamics with the sole exception that they employ gauge groups SU(N) with all spacetime derivatives ∂^μ in the Maxwell Lagrangian and the classical field equations including those operating on gauge fields and on the field strength replaced by $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$, and so are non-Abelian versions of Maxwell's electrodynamics.

Starting from this view, we show how chromodynamics in the form of an $SU(3)_C$ gauge theory need not be posited at all, but emerges entirely as a *corollary theory* based on positing Maxwell gauge theory with Yang-Mills extension as the underlying, *fundamental theory*. But in the process, extending beyond the pedagogical utility of this viewpoint, we not only uncover $SU(3)_C$ chromodynamics in its usual expected form, but we also come upon baryons and show them to be the magnetic monopoles of these Yang-Mills extensions of Maxwell. We further find out how and why interactions between observed strong particle states such as protons and neutrons are mediated by mesons, we develop certain important connections to gravitational Riemannian geometry, and we solve the Yang Mills mass gap and confinement problems.

In laying out the "Yang-Mills and Mass Gap" problem which the present paper solves, Jaffe and Witten point out at page 3 of [1] that:

> ". . . for QCD to describe the strong force successfully, it must have at the quantum level the following three properties, each of which is dramatically different from the behavior of the classical theory: 1) It must have a "mass gap;" namely there must be some constant $\Delta > 0$ such that every excitation of the vacuum has energy at least Δ. (2) It 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. (3) It must have "chiral symmetry breaking," which means that the vacuum is potentially invariant (in the limit, that the quark-bare masses vanish) only under a certain subgroup of the full symmetry group that acts on the quark fields."

They further proceed to state that:

> "The first point is necessary to explain why the nuclear force is strong but short-ranged; the second is needed to explain why we never see individual quarks; and the third is needed to account for the 'current algebra' theory of soft pions that was developed in the 1960s."

They then continue (emphasis added, original references renumbered):

> "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, [2]). *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.*"

Moving past a statement of the problem to *how* the mass gap might be solved, Jaffe and Witten 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 finally observe that:

> "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 [3], and may be tied to curvature in the space of connections, see [4]."

This is the view of the Yang-Mills mass gap that will be developed here and used to solve this problem. It is in accord Occam's Razor as restated by Einstein [5], 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." 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. *But more importantly than theoretical economy, this view actually does lead to confinement and a solution to the mass gap and chiral symmetry breaking*.

In other words, we show how confinement and the mass gap and chiral symmetry breaking can be *fully* explained using no more than a Yang-Mills field strength $F = dA + gA \wedge A$ via the quartic action terms $(A \wedge A)^2$. This places the mass gap and confinement and chiral solutions *entirely on the shoulders of Yang-Mills theory without any supplement*. Because the classical Yang-Mills equations are simply those of Maxwell extended into the non-Abelian domain, this would entirely explain nuclear short range and quark and gauge field confinement and chiral symmetry breaking on the basis of "Maxwell's equations . . . replaced by the Yang–Mills equations, $0 = d_A F = d_A {*F}$" ([1] pages 1-2), and so *reveals Maxwell's theory, with the simple replacement of all ordinary derivatives in the Lagrangian and classical field equations by gauge-covariant derivatives and nothing more, to be the governing theory of nuclear physics.*

In sum, by taking a view that the fundamental theory of Yang-Mills electrodynamics naturally gives birth to $SU(3)_C$ as a corollary, secondary theory of strong interactions, we see how $SU(3)_C$ naturally emerges such that there is a built in, non-trivial $SU(3)_C$ transformation for elementary quark and gluon fields concurrent with $SU(3)_C$ invariance for the physical particle states which leads to a naturally-emergent, built-in form of quark and gluon confinement, meson interaction, chiral symmetry breaking, and a mass gap. *These features are not easily seen if one starts out by assuming $SU(3)_C$ to be the theory of strong interactions. But they are discovered if one starts out only with Maxwell and Yang-Mills and then derives QCD as a corollary theory.* The purpose of this paper is to convincingly demonstrate this.

What is novel about his paper is the following: 1) In section 7, we are able to obtain a classical unification of gravitational theory with gauge theory at the level of the Einstein field equation, see (7.6). 2) In section 9 we uncover an infinite recursion which does not appear to have previously been found, and which could provide a tool for carrying out Yang-Mills path integration in an exact, analytical fashion, and thereby quantizing Yang-Mills theory, exactly. 3) In section 10 we solve the mass gap, see (10.12) and (10.13), which explains how nuclear interactions can have short range yet at the same time be based on massless gluons. 4) In section 11 we solve confinement and show how QCD naturally emerges as a corollary theory from Yang-Mills gauge theory, and specifically how the Yang-Mills monopoles are synonymous with baryons consisting of three colored quarks in the ground state and interacting solely via meson exchange with individual quarks and gluons remaining strictly confined, see (11.1) and (11.18) and section 11 generally. 5) In section 12, we uncover the origins of chiral symmetry breaking in strong interactions, and particularly, of the vector (V) and axial (A) character of the phenomenologically-observed mesons. 6) In section 13, we use the recursive aspects of Yang-Mills theory earlier uncovered in section 9 to develop and solve an exact, closed recursive path integral for Quantum Yang-Mills Theory, which proves the existence of a non-trivial quantum Yang–Mills theory on $\mathbb{R}^4$ for any simple gauge group G.

Now, we provide a brief overview of this paper: 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. So in section 2, we begin by reviewing Yang-Mills theory from three equivalent viewpoints: that of a gauge theory for non-commuting gauge fields; that of a gauge theory with non-linear interactions between gauge fields, and that of an Abelian gauge theory "on steroids" by virtue of a "minimal coupling" principle through which all ordinary spacetime derivatives in the Lagrangian and classical field equations are replaced by gauge-covariant derivatives and the theory is consequently turned into a non-Abelian gauge theory.

In section 3, we examine the classical Maxwell equations for the electric and magnetic charge densities, and demonstrate how the non-commuting nature of Yang-Mills theory naturally gives rise to non-zero magnetic charge densities. Section 4 begins to show how the Yang-Mills magnetic charge densities have a number of symmetry characteristics which are reminiscent of baryons, most notably, that there is no net flux of a Yang-Mills gauge field across any closed surface surrounding a Yang-Mills monopole for the exact same formal reasons that there are no monopoles *at all* in an Abelian gauge theory such as that of Maxwell. We return to this discussion in section 11 following further development at which point we are able to formally identify these Yang-Mills monopoles with baryons containing three colored quarks in the ground state and showing that these monopoles have all of the required features of quark and gluon confinement was well as interactions which transpire via mesons.

In section 5 we develop a fourth, perturbative view of Yang-Mills theory, and in section 6 we develop a fifth view of gauge theory – which is the original view of Hermann Weyl, the founder of gauge theory – based on geometric curvature in a gauge / phase space. In section 7 we make use of this view to uncover in (7.6) a "twin" of the Einstein equation which is the gravitational field equation of Yang-Mill gauge theory. Because this field equation remains valid even for Abelian gauge theory, this unifies gravitation with the non-gravitational interactions including electrodynamics, at the classical level.

While sections 4 through 7 focus largely on the magnetic charge densities, section 8 returns to the electric charge densities. Observing that the magnetic and electric charge densities are essentially a set of linked equations parameterized by the gauge fields, in section 8 we invert the electric charge density so that the gauge fields appearing in the magnetic charge density may be replaced by the source currents form which they originate, which in turn enables us to replace the source currents with the fermion wavefunctions from which they arise and thus "populate" the monopole densities with fermion wavefunctions. In section 9 we make use of this inverse to in fact "populate" the monopole densities with fermion wavefunctions. In so doing, we come to see that the inverse $I_{\tau\mu}$ defined such that $G_\mu \equiv I_{\tau\mu} J^\tau$ which is used to replace the gauge fields with the current densities and then with the fermion wavefunctions is actually a *recursive* expression which embeds an infinite recursive nesting of gauge fields and thus an infinite succession of current densities and fermion wavefunctions. This finding of an infinite recursion represents yet a sixth view of the non-linear character of Yang-Mills theory which may be of help in developing an exact, analytical solution to the Yang-Mills path integral and thus yielding quantum Yang-Mills theory on an exact footing.

Sections 10, 11 and 12 then present the solutions to the three main aspects of the mass gap problem, namely, the mass gap itself, quark confinement, and chiral symmetry breaking. Section 10, in equations (10.12) and (10.13) contains the mass gap solution. Section 11 completes the development first started in section 4 and shows how and why we are able to formally identify the Yang-Mills monopoles with baryons containing three colored quarks in the ground state and show that these monopoles have all of the required features of quark and gluon confinement was well as interactions which transpire via mesons. Section 12 shows the origin of chiral symmetry breaking in the quaternion nature of the Dirac gamma matrices, and in the infinite recursion of gauge fields and current densities developed in section 9.

Finally, in section 13, we use the recursive aspects of Yang-Mills theory earlier uncovered in section 9 to develop and solve an exact, closed recursive path integral for Quantum Yang-Mills Theory, which proves the existence of a non-trivial quantum Yang–Mills theory on $\mathbb{R}^4$ for any simple gauge group G. Section 14 concludes.

2. Classical Yang-Mills Theory: Three Equivalent Viewpoints

Yang-Mills gauge theories, first developed in 1954 [6] by C. N. Yang and R. Mills, rest mathematically upon the generalization of the 2x2 Pauli matrices of SU(2) into SU(N) matrices of any NxN dimensionality. These Pauli matrices for which $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = -i\sigma_1\sigma_2\sigma_3 = I$ and which have the commutation relationship $\left[\sigma_i, \sigma_j\right] = 2i\varepsilon_{ijk}\sigma_k$, are in turn the direct descendants of the quaternions $i^2 = j^2 = k^2 = ijk = -1$ which Hamilton first carved with his penknife into the Brougham Bridge in Dublin, Ireland in 1843, presaging what has since become the use of non-commuting numbers throughout modern physics. 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 becomes $\left[\sigma_i, \sigma_j\right] = i\varepsilon_{ijk}\sigma_k$ for the normalization $Tr\left(\sigma^i \sigma^j\right) = \frac{1}{2}\delta^{ij}$. Each generator λ^i is an NxN matrix and so can be written $\lambda^i{}_{AB}; A, B = 1,2,3...N$, but in general it is simpler and more compact 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 G^μ is an NxN matrix containing the N^2-1 spacetime 4-vector gauge potentials. Similarly, one forms a set of N^2-1 field strength tensors $F^{i\mu\nu}$, each of which is a bivector containing a "chromo-electric" field $\mathbf{E_i}$ and a chromo-magnetic field $\mathbf{B_i}$ in the usual manner, aside from the N^2-1-fold replication of these fields. 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., [7], equation IV.5(16)):

$$F^{\mu\nu} = \partial^\mu G^\nu - \partial^\nu G^\mu - i\left[G^\mu, G^\nu\right] = \partial^{[\mu} G^{\nu]} - i\left[G^\mu, G^\nu\right]. \tag{2.1}$$

Because the gauge fields G^μ are NxN Yang-Mills matrices $G^{\mu}{}_{AB} \equiv \lambda^i{}_{AB} G^{i\mu}$, this commutator $\left[G^\mu, G^\nu\right] = G^\mu G^\nu - G^\nu G^\mu$ is non-vanishing, $\left[G^\mu, G^\nu\right] \neq 0$. *Much of what differentiates Yang-Mills gauge theory from an Abelian gauge theory such as QED, originates 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 field strength (2.1), there are several different, fully equivalent ways in which one can think about Yang-Mills gauge theories. 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 (2.1), 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, and will eventually become associated with the observed baryons including protons and neutrons.

For a second way to think about Yang-Mills, it is worth being reminded how to expand (2.1) using $F^{\mu\nu}=\lambda^i F^{i\mu\nu}$, $G^\mu=\lambda^i G^{i\mu}$ and $\left[\lambda_i,\lambda_j\right]=if_{ijk}\lambda_k$. Renaming summed indexes as needed, this expansion yields:

$$\begin{aligned}\lambda^i F^{i\mu\nu}&=\partial^\mu\lambda^i G^{i\nu}-\partial^\nu\lambda^i G^{i\mu}-i\left[\lambda^i G^{i\mu},\lambda^j G^{j\nu}\right]=\lambda^i\partial^\mu G^{i\nu}-\lambda^i\partial^\nu G^{i\mu}-i\left[\lambda^i,\lambda^j\right]G^{i\mu}G^{j\nu}\\&=\lambda^i\partial^\mu G^{i\nu}-\lambda^i\partial^\nu G^{i\mu}+f^{kji}\lambda^i G^{k\mu}G^{j\nu}\end{aligned}. \qquad (2.2)$$

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

$$F^{i\mu\nu}=\partial^\mu G^{i\nu}-\partial^\nu G^{i\mu}+f^{ijk}G^{j\mu}G^{k\nu}=\partial^{[\mu}G^{i\nu]}+f^{ijk}G^{j\mu}G^{k\nu}. \qquad (2.3)$$

If we now use (2.3) to form a Lagrangian density akin to the QED $\mathfrak{L}=-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ for a pure gauge field, we obtain the also familiar (see, e.g., [7], equations (VII.1.(1)-(2)):

$$\begin{aligned}\mathfrak{L}=-\tfrac{1}{4}F^{i\mu\nu}F_{i\mu\nu}&=-\tfrac{1}{4}\left(\partial^{[\mu}G^{i\nu]}+f^{ijk}G^{j\mu}G^{k\nu}\right)\left(\partial_{[\mu}G_{i\nu]}+f_{ilm}G_{l\mu}G_{m\nu}\right)\\&=-\tfrac{1}{4}\partial^{[\mu}G^{i\nu]}\partial_{[\mu}G_{i\nu]}-\tfrac{1}{2}f_{ijk}\partial^{[\mu}G^{i\nu]}G_{j\mu}G_{k\nu}-\tfrac{1}{4}f^{ijk}f_{ilm}G^{j\mu}G^{k\nu}G_{l\mu}G_{m\nu}\end{aligned}. \qquad (2.4)$$

The first term, $-\frac{1}{4}\partial^{[\mu}G^{i\nu]}\partial_{[\mu}G_{i\nu]}$, a "harmonic oscillator" term, is quadratic in the gauge fields, and is fully analogous and indeed identical in form to the term $-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}=-\frac{1}{4}\partial^{[\mu}G^{\nu]}\partial_{[\mu}G_{\nu]}$ in the Lagrangian density of electrodynamics. But the remaining terms $-\frac{1}{2}f_{ijk}\partial^{[\mu}G^{i\nu]}G_{j\mu}G_{k\nu}$ and $-\frac{1}{4}f^{ijk}f_{ilm}G^{j\mu}G^{k\nu}G_{l\mu}G_{m\nu}$, the "perturbation" terms, represent vertices with three and four interacting gauge fields. This is not seen in electrodynamics, and makes Yang-Mills a *non-linear* theory. *So the second way to think about Yang-Mills theory is that of (2.4), in which the gauge fields do* <u>*not*</u> *act like photons by foregoing interactions with 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.*

As Zee points out in section VII.1 of [7], present methods used to calculate in Yang-Mills theory, such as perturbation theory or lattice gauge theory, are severely truncated methods 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, which is highlighted by the separation of terms in (2.4), in Zee's description, is "an unnatural act as it involves brutally splitting [the Lagrangian density] L into two parts: a part quadratic in the fields and the rest." Lattice gauge theory [2], in contrast, "does violence to Lorentz invariance rather than to gauge invariance." Further, as a fundamentally computational rather than analytical method based on small but finite lattice spacing, Lattice gauge theory is akin to doing calculus in Yang-Mills gauge theory using the finite limits that were used before Newton taught us how to do calculus with infinitesimal limits. 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 do violence to neither gauge symmetry nor Poincare symmetry, and which fully employ all the tools of modern calculus. Because doing exact calculations with (2.4) is difficult, in general we will find it unhelpful to split (2.4) into harmonic and perturbative parts as is done in perturbative gauge theory, or to spoil the Lorentz invariance or be restricted by finite limits as in lattice gauge theory, and will look to other approaches.

A third way to think about Yang-Mills gauge theory is to expand the commutator in (2.1) and then reconsolidate using gauge covariant derivatives $D^{\mu} \equiv \partial^{\mu} - iG^{\mu}$, as such: (In general, for compactness, we scale the interaction charge strength g into the gauge field via $gG^{\mu} \rightarrow G^{\mu}$. This g can always be extracted back out when explicitly needed.):

$$F^{\mu\nu} = \partial^{\mu}G^{\nu} - \partial^{\nu}G^{\mu} - iG^{\mu}G^{\nu} + iG^{\nu}G^{\mu} = \left(\partial^{\mu} - iG^{\mu}\right)G^{\nu} - \left(\partial^{\nu} - iG^{\nu}\right)G^{\mu} = D^{\mu}G^{\nu} - D^{\nu}G^{\mu} = D^{[\mu}G^{\nu]} . \tag{2.5}$$

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 $\partial^{\mu} \rightarrow D^{\mu} = \partial^{\mu} - iG^{\mu}$. *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 $\phi \rightarrow e^{i\theta(x)}\phi$ or $\psi \rightarrow e^{i\theta(x)}\psi$ and insists that the field equation or Lagrangian remain invariant under this transformation. What does one do to ensure such invariance? Make the replacement $\partial^{\mu} \rightarrow D^{\mu} = \partial^{\mu} - iG^{\mu}$. So, one then changes $\partial_{\mu}\phi \rightarrow D_{\mu}\phi$ and $\partial_{\mu}\psi \rightarrow D_{\mu}\psi$ with the consequence that ϕ or ψ *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 $\partial^{\mu} \rightarrow D^{\mu} = \partial^{\mu} - iG^{\mu}$ so that $F^{\mu\nu} = D^{[\mu}G^{\nu]}$, which is (2.5). 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 (2.4), with the exception that in the form of (2.5), 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):

$$\mathfrak{L} = -\tfrac{1}{2}\mathrm{Tr}F^{\mu\nu}F_{\mu\nu} = -\tfrac{1}{2}\mathrm{Tr}D^{[\mu}G^{\nu]}D_{[\mu}G_{\nu]} . \tag{2.6}$$

Given that (2.4) and (2.6) state *exactly the same physics*, it should be clear that (2.6) is a much easier expression to work with than (2.4) and does not "brutally split" anything. *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., [8] 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, in which the $-iG^{\mu}$ represents the gauge (really, phase) curvature based on a relative relationship between non-observable phases. This curvature view will be developed at length in sections 6 and 7.

These first and third views of Yang-Mills are the ones laid out by Jaffe and Witten in [1] at pages 1-2 when they point out that for Yang-Mills 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, $0 = d_AF = d_A{*F}$, where d_A is the gauge-covariant extension of the exterior derivative."

This view of Yang-Mills theory as simply being Maxwell's theory on steroids with a $\partial^{;\mu} \to D^{;\mu} = \partial^{;\mu} - iG^{\mu}$ replacement throughout ($d \to d_A$ in the above passage) is actually very attractive and mathematically simplifying. Physically, it says that the weak and strong interactions which are based respectively on SU(2) and SU(3), are just steroidal versions of Maxwell's electrodynamics in which all spacetime derivatives ∂^{μ} including those which act on gauge fields G^{ν} or field strengths $F^{\mu\nu} = D^{[\mu}G^{\nu]}$ are replaced with D^{μ}. It tells us that Maxwell already discovered the governing classical equations for the other non-gravitational (weak and strong) interactions but for the fact that he used commuting gauge fields $\left[G^{\mu}, G^{\nu}\right] = 0$ rather than non-commuting ones $\left[G^{\mu}, G^{\nu}\right] \neq 0$. And, as (2.5) teaches, *non-commuting a.k.a. non-Abelian gauge fields* <u>*inherently flow*</u> *from using gauge-covariant derivatives to define the field strength as* $F^{\mu\nu} = D^{[\mu}G^{\nu]}$, i.e., from putting Maxwell on steroids. So from this view, weak and strong interactions are simply governed by Maxwell's electrodynamics on steroids. The questions then become not about the nature of the governing theory for these interactions, but about 1) why SU(2) and SU(3) and not some other groups are used for these interactions; 2) what group G serves to unify these interactions and 3) what is the nature of the symmetry breaking that yields the phenomenological $G \to SU(3)_C \times SU(2)_W \times U(1)_Y \to SU(3)_C \times U(1)_{EM}$. The focus here will be on the first question, and specifically, how it is that everything needed to deduce $SU(3)_C$ and explain confinement and chiral symmetry breaking and solve the mass gap is embodied in this view of Yang-Mills gauge theory as Maxwell's electrodynamics on steroids.

3. 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 $0 = d_AF = d_A{*F}$ 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:

$$J^{\nu} = D_{;\mu} F^{\mu\nu} , \tag{3.1}$$

$$P^{\sigma\mu\nu} = D^{;\sigma} F^{\mu\nu} + D^{;\mu} F^{\nu\sigma} + D^{;\nu} F^{\sigma\mu} \equiv D^{;(\sigma} F^{\mu\nu)} = \partial^{;(\sigma} F^{\mu\nu)} - iG^{(\sigma} F^{\mu\nu)} . \tag{3.2}$$

In (3.2), we have also defined a "cyclator" notation $(\sigma\mu\nu)$ to represent the cycling of three free indexes over three terms, as shown, which will be useful for compacting the somewhat lengthy expressions we shall soon be deriving for $P^{\sigma\mu\nu}$. We have also regarded the spacetime to be curved and so have included the gravitationally-covariant derivatives $\partial_{;\mu} G^{\nu} \equiv \partial_{\mu} G^{\nu} + \Gamma^{\nu}_{\mu\sigma} G^{\sigma}$ (which become exterior derivatives when used in differential forms). Here in (3.1) and (3.2) 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.

Referring to the "three views" of Yang-Mills just reviewed, we shall find that for Yang-Mills magnetic sources $P^{\sigma\mu\nu}$ of (3.2), it is most helpful to view Yang-Mills theory in the form of (2.1), 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 the Yang-Mills electric sources of (3.1), the more convenient view is that of (2.6), in which we view Yang-Mills as gauge theory on steroids. So, as a first step, taking the "gauge theory on steroids" view of Yang-Mills, and employing spacetime-covariant derivatives, we substitute the field strength represented as $F^{\mu\nu} = D^{;[\mu} G^{\nu]}$ from (2.5) into (3.1), while taking the "non-commuting gauge fields" view of Yang-Mills, we substitute $F^{\mu\nu} = \partial^{;[\mu} G^{\nu]} - i\left[G^{\mu}, G^{\nu} \right]$ of (2.1), which is entirely equivalent to (2.5), into (3.2).

So for the Yang-Mills electric source density (3.1), using $D^{;\mu} \equiv \partial^{;\mu} - iG^{\mu}$ and (2.5) and some well-known index gymnastics, we obtain:

$$\begin{aligned} J^{\nu} &= D_{;\mu} F^{\mu\nu} = D_{;\mu} D^{;[\mu} G^{\nu]} = D_{;\mu} D^{;\mu} G^{\nu} - D_{;\mu} D^{;\nu} G^{\mu} = \left(g^{\mu\nu} D_{;\sigma} D^{;\sigma} - D^{;\mu} D^{;\nu} \right) G_{\mu} \\ &\overset{+m^2}{\Rightarrow} \left(g^{\mu\nu} \left(D_{;\sigma} D^{;\sigma} + m^2 \right) - D^{;\mu} D^{;\nu} \right) G_{\mu} \end{aligned} . \tag{3.3}$$

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 (3.3), 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 "reveal" masses from a Lagrangian with gauge symmetry via spontaneous symmetry breaking or some analogous method to reveal masses, 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 filling mass gap in section 10, where we shall eventually set this mass to zero and show how even with this mass going to zero there will be non-zero vector 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., [7], 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 review in section 8.

The above (3.3) 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. We see the gauge theory "minimal coupling principle" at work here: in (3.3) 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 (3.3) is $g^{\mu\nu}\left(D_{;\sigma}D^{;\sigma} + m^2\right) - D^{;\mu}D^{;\nu}$, in contrast to the analogous operator $g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\mu}\partial^{;\nu}$ in electrodynamics. These operators will play an important role in the development here, and in section 8 we shall be obtaining their inverses.

For the Yang-Mills magnetic source density (3.2), it will help to first review how the monopole density (3.2) 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 $R^\sigma{}_{\alpha\mu\nu}$ may be *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 $R_\tau{}^{(\nu\sigma\mu)} = R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma} = 0$, with a cycling of indexes identical to that which obtains in the magnetic monopole field equation (3.2). Writing (3.2) 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 electrodynamic calculation is as follows:

$$\begin{aligned}
P^{\sigma\mu\nu} &= \partial^{;\sigma}F^{\mu\nu} + \partial^{;\mu}F^{\nu\sigma} + \partial^{;\nu}F^{\sigma\mu} \\
&= \partial^{;\sigma}\left(\partial^{;\mu}G^\nu - \partial^{;\nu}G^\mu\right) + \partial^{;\mu}\left(\partial^{;\nu}G^\sigma - \partial^{;\sigma}G^{;\nu}\right) + \partial^{;\nu}\left(\partial^{;\sigma}G^\mu - \partial^{;\mu}G^\sigma\right) \\
&= \left[\partial^{;\sigma}, \partial^{;\mu}\right]G^\nu + \left[\partial^{;\mu}, \partial^{;\nu}\right]G^\sigma + \left[\partial^{;\nu}, \partial^{;\sigma}\right]G^\mu \\
&= \left(R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma}\right)G^\tau = \mathbf{0}
\end{aligned} \tag{3.4}$$

This is a very important result, because it tells us that vanishing magnetic monopoles in Maxwell's theory (and to be discussed later, the confinement of color in QCD), are brought about not only via the trivial relationship $\left[\partial^\mu, \partial^\nu\right] = 0$ for the commuting of derivatives in flat spacetime, but also via the Bianchi identity $R_\tau{}^{(\nu\sigma\mu)}=0$ 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, wherein $P^{\sigma\mu\nu}=\partial^{;(\sigma}F^{\mu\nu)}=0$ is geometrically-rooted in $R_\tau{}^{(\nu\sigma\mu)}=0$. In the language of "differential forms," (3.4) 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., [9] §4.6.

It will also be of interest here to consider the monopole equation (3.4) 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 (3.4), one writes $P=dF=ddG=0$ as (antisymmetric 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):

$$\iiint P=\iiint dF=\iiint ddG=\oiint F=\oiint F^{\mu\nu}dx_\mu dx_\nu=\oiint dG=0. \tag{3.5}$$

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.

With this review of the vanishing of magnetic charges in Maxwell's Abelian theory, we now turn back to the non-Abelian $F^{\mu\nu}=\partial^{;[\mu}G^{\nu]}-i\left[G^\mu,G^\nu\right]$ of (2.1). Using this in the non-Abelian (3.2), also making use of $D^{;\mu}=\partial^{;\mu}-iG^\mu$, noting as just reviewed in (3.4) that $\left(R_\tau{}^{\nu\sigma\mu}+R_\tau{}^{\sigma\mu\nu}+R_\tau{}^{\mu\nu\sigma}\right)G^\tau=0$, and at the end condensing with the cyclator $(\sigma\mu\nu)$, we obtain:

$$\begin{aligned}
P^{\sigma\mu\nu} &= D^{;\sigma}F^{\mu\nu} + D^{;\mu}F^{\nu\sigma} + D^{;\nu}F^{\sigma\mu} \\
&= D^{;\sigma}\left(\partial^{;[\mu}G^{\nu]} - i\left[G^{\mu},G^{\nu}\right]\right) + D^{;\mu}\left(\partial^{;[\nu}G^{\sigma]} - i\left[G^{\nu},G^{\sigma}\right]\right) + D^{;\nu}\left(\partial^{;[\sigma}G^{\mu]} - i\left[G^{\sigma},G^{\mu}\right]\right) \\
&= \left(R_{\tau}{}^{\nu\sigma\mu} + R_{\tau}{}^{\sigma\mu\nu} + R_{\tau}{}^{\mu\nu\sigma}\right)G^{\tau} - i\left(\partial^{;\sigma}\left[G^{\mu},G^{\nu}\right] + \partial^{;\mu}\left[G^{\nu},G^{\sigma}\right] + \partial^{;\nu}\left[G^{\sigma},G^{\mu}\right]\right) \\
&\quad - i\left(G^{\sigma}\partial^{;[\mu}G^{\nu]} + G^{\mu}\partial^{;[\nu}G^{\sigma]} + G^{\nu}\partial^{;[\sigma}G^{\mu]}\right) - \left(G^{\sigma}\left[G^{\mu},G^{\nu}\right] + G^{\mu}\left[G^{\nu},G^{\sigma}\right] + G^{\nu}\left[G^{\sigma},G^{\mu}\right]\right) \\
&= \mathbf{0} - i\left(\partial^{;\sigma}\left[G^{\mu},G^{\nu}\right] + \partial^{;\mu}\left[G^{\nu},G^{\sigma}\right] + \partial^{;\nu}\left[G^{\sigma},G^{\mu}\right] + G^{\sigma}\partial^{;[\mu}G^{\nu]} + G^{\mu}\partial^{;[\nu}G^{\sigma]} + G^{\nu}\partial^{;[\sigma}G^{\mu]}\right) \\
&\qquad - \left(G^{\sigma}\left[G^{\mu},G^{\nu}\right] + G^{\mu}\left[G^{\nu},G^{\sigma}\right] + G^{\nu}\left[G^{\sigma},G^{\mu}\right]\right) \\
&= \mathbf{0} - i\left(\partial^{;(\sigma}\left[G^{\mu},G^{\nu)}\right] + G^{(\sigma}\partial^{;[\mu}G^{\nu])}\right) - G^{(\sigma}\left[G^{\mu},G^{\nu)}\right] \\
&= \mathbf{0} - i\left(\partial^{;(\sigma}\left[G^{\mu},G^{\nu)}\right] + G^{(\sigma}D^{;[\mu}G^{\nu])}\right)
\end{aligned} \quad . \tag{3.6}$$

So, in sum, (3.3) is the classical electric source field equation of Yang-Mills gauge theory corresponding to Maxwell's equation $J^{\nu} = \partial_{;\mu}F^{\mu\nu}$ for electric charges, and (3.6) is the classical magnetic source field equation of Yang-Mills gauge theory corresponding to Maxwell's equation $0 = \partial^{;\sigma}F^{\mu\nu} + \partial^{;\mu}F^{\nu\sigma} + \partial^{;\nu}F^{\sigma\mu}$ for (vanishing in $U(1)_{em}$) magnetic charges.

4. The 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 (3.6) 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 relation to (3.4) and (3.5) above, this is why there are no magnetic monopoles in QED. But because $\left[G^{\mu},G^{\nu}\right] \neq 0$, we have some non-zero remaining terms $-i\left(\partial^{;(\sigma}\left[G^{\mu},G^{\nu)}\right] + G^{(\sigma}\partial^{;[\mu}G^{\nu])}\right) - G^{(\sigma}\left[G^{\mu},G^{\nu)}\right]$, and consequently *these magnetic monopoles are non-vanishing.* So if one believes in Maxwell's electrodynamics and one believes in Yang-Mills gauge theory, then one must also believe that the magnetic monopoles (3.6) exist somewhere, in some form, in the physical universe. Indeed, t'Hooft [10] and Polyakov [11] were among the first to recognize this. What form they exist in, however, remains an open question to this day. Whether these monopoles 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 and are the "colour magnetic charges" referenced by Cheng and Li [12] at 472-473 (which the author contends in [13] 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 (3.6) 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 (2.1) now takes on the compacted form $F = dG - iG^2 = DG$. As a result, (3.6) is written compactly with $D = d - iG$ as:

$$\begin{aligned} P = DF &= \left(d - iG\right)F = D\left(dG - iG^2\right) = \left(d - iG\right)\left(dG - iG^2\right) = ddG - idG^2 - iGdG - G^3 \\ &= \mathbf{0} - i\left(dG^2 + GdG\right) - G^3 = \mathbf{0} - i\left(dG^2 + GDG\right) \end{aligned}, \tag{4.1}$$

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 (3.6) and (4.1) are $dG^2 \Leftrightarrow \partial^{;(\sigma}\left[G^\mu, G^{\nu)}\right]$, $GdG \Leftrightarrow G^{(\sigma}\partial^{;[\mu}G^{\nu])}$, $G^3 \Leftrightarrow G^{(\sigma}\left[G^\mu, G^{\nu)}\right]$ and $GDG \Leftrightarrow G^{(\sigma}D^{;[\mu}G^{\nu])}$. So now, via (4.1) and the use of Gauss'/Stokes' theorem $\iint dX = \oint X$ in differential forms, the Yang-Mills magnetic monopole equation in integral form is:

$$\begin{aligned} \iiint P = \iiint DF &= \oiint F - \iiint iGF = \iiint\left(ddG - i\left(dG^2 + GdG\right) - G^3\right) = \iiint\left(-i\left(dG^2 + GdG\right) - G^3\right) \\ &= \oiint dG - i\oiint G^2 - \iiint\left(iGdG + G^3\right) = \mathbf{0} - i\oiint G^2 - \iiint\left(iGdG + G^3\right) \\ &= \oiint dG - i\oiint G^2 - i\iiint GDG = \mathbf{0} - i\oiint G^2 - i\iiint GDG \end{aligned} .\tag{4.2}$$

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 $G^3 \Leftrightarrow G^{(\sigma}\left[G^\mu, G^{\nu)}\right]$ or $GDG \Leftrightarrow G^{(\sigma}D^{;[\mu}G^{\nu])}$. Note also that (4.2) embeds $\oiint dG = 0$, which in (3.5) for electrodynamics tells us that there is no *net* magnetic field flux across any *closed* two-dimensional surface. Above, the magnetic charge equation (3.5) of Maxwell's theory, $\iiint P = \oiint F = 0$, now becomes $\iiint P = \oiint F - i\iiint GF = -i\oiint G^2 - i\iiint GDG$.

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

$$\begin{aligned} &-i\iiint dG^2 = -i\iiint \partial^{;(\sigma}\left[G^\mu, G^{\nu)}\right]dx_\sigma dx_\mu dx_\nu \\ &= -i\iiint\left(\partial^{;\sigma}\left[G^\mu, G^\nu\right] + \partial^{;\mu}\left[G^\nu, G^\sigma\right] + \partial^{;\nu}\left[G^\sigma, G^\mu\right]\right)dx_\sigma dx_\mu dx_\nu \; . \\ &= -3i\oiint\left[G^\mu, G^\nu\right]dx_\mu dx_\nu = -i\oiint G^2 \end{aligned} \tag{4.3}$$

Then let us use this with (3.6) to expand some key terms in (4.2), and thereafter consolidate using $D^{;\mu} = \partial^{;\mu} - iG^\mu$ thus $-iGdG - G^3 = -iGDG$ and some summed index renaming as follows:

$$\begin{aligned}
\iiint P &= \iiint P^{\sigma\mu\nu} dx_\sigma dx_\mu dx_\nu \\
&= \iiint \left(R_\tau{}^{\nu\sigma\mu} + R_\tau{}^{\sigma\mu\nu} + R_\tau{}^{\mu\nu\sigma} \right) G^\tau dx_\sigma dx_\mu dx_\nu \\
&-i \iiint \left(\partial^{;\sigma} \left[G^\mu, G^\nu \right] + \partial^{;\mu} \left[G^\nu, G^\sigma \right] + \partial^{;\nu} \left[G^\sigma, G^\mu \right] \right) dx_\sigma dx_\mu dx_\nu \\
&-i \iiint \left(G^\sigma \partial^{;[\mu} G^{\nu]} + G^\mu \partial^{;[\nu} G^{\sigma]} + G^\nu \partial^{;[\sigma} G^{\mu]} \right) dx_\sigma dx_\mu dx_\nu \\
&- \iiint \left(G^\sigma \left[G^\mu, G^\nu \right] + G^\mu \left[G^\nu, G^\sigma \right] + G^\nu \left[G^\sigma, G^\mu \right] \right) dx_\sigma dx_\mu dx_\nu \\
&= \mathbf{0} - 3i \oiint \left[G^\mu, G^\nu \right] dx_\mu dx_\nu - 3i \iiint G^\sigma D^{;[\mu} G^{\nu]} dx_\sigma dx_\mu dx_\nu \\
&= \oiint dG - i \oiint G^2 - i \iiint G \partial G - \iiint G^3 = \oiint dG - i \oiint G^2 - i \iiint GDG \\
&= \mathbf{0} - i \oiint G^2 - i \iiint G \partial G - \iiint G^3 = \mathbf{0} - i \oiint G^2 - i \iiint GDG
\end{aligned} \quad . \tag{4.4}$$

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 demonstrate in section 11 that this term represents a net flow of mesons through the closed surfaces.

Third, making (3.6) even more interesting, as detailed in section 1 of the author's [13], if we perform a local transformation $F \to F' = F - dG$ on the field strength F, which in expanded form is written as $F^{\mu\nu} \to F^{\mu\nu}{}' = F^{\mu\nu} - \partial^{[\nu} G^{\mu]}(x)$, then we find from (4.2) as a direct result of $\oiint dG = 0$ which in electrodynamics includes the Maxwell equation $\oiint \vec{B} \cdot d\vec{A} = 0$ and Faraday's law $\oint \vec{E} \cdot d\vec{l} = -\iint (\partial \vec{B} / \partial t) \cdot d\vec{A}$, see after (3.5), that:

$$\iiint P = \oiint F \to \oiint F' = \oiint \left(F - dG \right) = \oiint F \quad . \tag{4.5}$$

This means that the *net* flow of the field strength $\oiint F = \oiint dG - i \oiint G^2 = -i \oiint G^2$ across a closed *two* dimensional surface is *invariant* under the local gauge-like transformation $F^{\mu\nu} \to F^{\mu\nu}{}' = F^{\mu\nu} - \partial^{[\nu} G^{\mu]}$, and that this invariance is caused by the equation $\oiint dG = 0$ which in

Maxwell theory is responsible for Faraday's law and the absence of magnetic monopoles. So in Yang-Mills theory, $\oiint dG = 0$ is responsible for the *symmetry principle* expressed in (4.5).

Fourth, we see from (4.4) that $\iiint G^3 = \iiint G^{(\sigma}\left[G^\mu, G^{\nu)}\right] dx_\sigma dx_\mu dx_\nu$ 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., [14] 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 strong $SU(3)_C$ are represented by the eight gluons of $G^\tau = \lambda^i G^{i\tau}$ with $i = 1, 2, 3...8$) is symbolically specified by $\oiint \text{Gluons} = 0$. Similarly, the confinement of individual quarks (which are represented by the $SU(3)_C$ Dirac wavefunction $\psi_A;\, A = 1, 2, 3$ with three color eigenstates *R*, *G*, *B*) is specified symbolically by $\oiint \text{Quarks} = 0$. 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 (4.2) through (4.5) contain a lot of information about what does and does not flow across the closed $\oiint$ 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 $\oiint dG = 0$ in (4.4) and is the exact same expression which yields the absence of magnetic monopoles entirely, in Abelian electrodynamics, review (3.4). 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 from any net surface flow as a direct result of its coupling through the Riemannian geometry in the configuration of the first Bianchi identity, which upon Gauss' / Stokes' integration yields $\oiint dG = 0$. 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. Recall that in electrodynamics, magnetic fields 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 interpretation of (4.4) as saying that there is no net flow of gauge fields across a closed Yang-

Mills monopole surface is strengthened by the fact displayed in (4.5) that $\oiint F \to \oiint F' = \oiint F$ is invariant under the local 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 G^{μ} (gluons in QCD) *not observable* with respect to net flux through the closed surface. *This may mean as argued in section 1 of [13] that gauge fields are confined within the non-vanishing magnetic monopoles of Yang-Mills gauge 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 term $\iiint dG^2 = \oiint G^2 = 3\oiint \left[G^{\mu}, G^{\nu}\right] dx_{\mu} dx_{\nu}$ detailed in (4.3). This is only non-vanishing integrable term in (4.4), as so tells 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, see (2.1). If these $3\left[G^{\mu}, G^{\nu}\right]$ do *not* turn out to represent individual quarks, then because there are no other non-vanishing integrable terms in (4.4), what (4.4) 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 surfaces of a Yang-Mills magnetic monopole, that is, that $\oiint \text{Gluons} = 0$ 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 $\bar{R}R + \bar{G}G + \bar{B}B$. 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 do in section 11, then these magnetic monopoles, in the setting of spacetime geometry, 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.

A third point is made by 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 (4.3) and (4.4). Although these arise from the three additive terms in the various expressions in (4.4), "3" 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 $\bar{R}R + \bar{G}G + \bar{B}B$. 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* that there are three quarks per baryon as is presently done in QCD, and could instead *require* us to have three quarks per baryon upon 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 baryon including confinement and meson interaction, 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 as a corollary theory to Yang-Mills gauge theory!* 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 (4.4). So we need to now see if there is some way to "populate" these magnetic monopoles with quarks. This brings us back to (3.3), which is the field equation relating Yang-Mills *electric* charge densities J^ν to the gauge fields G_μ, and which we shall be inverting in section 8. This is because when (3.3) is inverted to express G_μ as a function of J^ν, see (9.2) infra, it becomes possible to replace all of the gauge fields in the monopole (4.4) by the source currents from which they originate, and then to replace these source currents with fermion wavefunctions via Dirac's $J^\nu = \overline{\psi}\gamma^\nu\psi$, and finally to identify these fermions with quarks. But at this point, to lay the foundation for this, it we must first explore two more views of Yang-Mills theory, namely the "perturbative" view to be developed in section 5, and the "curvature" view to be developed in section 6. Not only are these two views helpful as to how we conceptualize Yang-Mills theory, but they will greatly simplify the mathematical development of Yang-Mills theory in order to readily perform the inversion in section 8.

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

In section 2, we described three equivalent "views" of Yang-Mills gauge theory: as a field theory of non-commuting gauge fields (2.1); as a theory of non-linear interactions among the gauge fields (2.4); and as a minimally-coupled gauge theory on steroids (2.6), (3.1), (3.2) 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 (3.1), (3.2) when the field strength is expressed as $F^{\mu\nu} = D^{;[\mu}G^{\nu]}$ in the steroidal view of (2.5). This "perturbative" view is rooted in the Klein-Gordon equation

$$\begin{aligned} 0 &= \left(D_\sigma D^\sigma + m^2\right)\phi = \left(\left(\partial_\sigma - iG_\sigma\right)\left(\partial^\sigma - iG^\sigma\right) + m^2\right)\phi = \left(\partial_\sigma\partial^\sigma + m^2 - i\partial_\sigma G^\sigma - iG_\sigma\partial^\sigma - G_\sigma G^\sigma\right)\phi \\ &= \left(\partial_\sigma\partial^\sigma + m^2 + V\right)\phi \end{aligned} \tag{5.1}$$

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

$$V \equiv -i\partial_\sigma G^\sigma - iG_\sigma\partial^\sigma - G_\sigma G^\sigma . \tag{5.2}$$

In virtually identical fashion, we may use (2.5) and $D^\mu \equiv \partial^\mu - iG^\mu$ to rewrite the Yang-Mills chromo-electric field equation (3.3) as:

$$\begin{aligned}J^{\nu} &= \left(g^{\mu\nu}\left(\left(\partial_{;\sigma}\partial^{;\sigma} - i\left(\partial_{;\sigma}G^{\sigma} + G_{\sigma}\partial^{;\sigma}\right) - G_{\sigma}G^{\sigma}\right) + m^{2}\right) - \left(\partial^{;\mu}\partial^{;\nu} - i\left(\partial^{;\mu}G^{\nu} + G^{\mu}\partial^{;\nu}\right) - G^{\mu}G^{\nu}\right)\right)G_{\mu} \\ &= \left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + V + m^{2}\right) - \left(\partial^{;\mu}\partial^{;\nu} + V^{\mu\nu}\right)\right)G_{\mu}\end{aligned} ,(5.3)$$

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

$$V^{\mu\nu} \equiv -i\left(\partial^{;\mu}G^{\nu} + G^{\mu}\partial^{;\nu}\right) - G^{\mu}G^{\nu}, \tag{5.4}$$

$$V_{AB} = V_{\sigma}{}^{\sigma} = -i\partial_{\sigma}G^{\sigma} - iG_{\sigma}\partial^{\sigma} - G_{\sigma}G^{\sigma} = -i\partial_{\sigma}G^{\sigma}_{AB} - iG_{AB\sigma}\partial^{\sigma} - G_{AC\sigma}G^{\sigma}_{CB}. \tag{5.5}$$

The perturbation scalar is *identical* in form to (5.2), but in Yang-Mills theory, it is an NxN Yang-Mills matrix of spacetime scalars, as we are reminded about by the explicit showing of Yang-Mills indexes in (5.5).

Noting that for any two successive gauge-covariant derivatives:

$$D^{;\mu}D^{\nu} = \left(\partial^{;\mu} - iG^{\mu}\right)\left(\partial^{;\nu} - iG^{\nu}\right) = \partial^{;\mu}\partial^{;\nu} - i\partial^{;\mu}G^{\nu} - iG^{\mu}\partial^{;\nu} - G^{\mu}G^{\nu} = \partial^{;\mu}\partial^{;\nu} + V^{\mu\nu}, \tag{5.6}$$

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

$$V^{[\mu\nu]} = V^{\mu\nu} - V^{\nu\mu} = \left[D^{;\mu}, D^{;\nu}\right] = \left[D^{\mu}, D^{\nu}\right]. \tag{5.7}$$

So the anti-symmetrized $V^{[\mu\nu]}$ is synonymous with the commutator of the Yang-Mills covariant derivatives. But in *curved* spacetime, using (5.7) 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:

$$\left[D^{;\mu}, D^{;\nu}\right]A^{\sigma} = \left[\partial^{;\mu}, \partial^{;\nu}\right]A^{\sigma} + V^{[\mu\nu]}A^{\sigma} = \left(R_{\tau}{}^{\sigma\mu\nu} + \delta_{\tau}{}^{\sigma}V^{[\mu\nu]}\right)A^{\tau}. \tag{5.8}$$

Applying (5.8) and $F^{\mu\nu} = D^{[\mu}G^{\nu]}$ to the magnetic monopole (3.6), the curvature terms vanish as in (3.4) via $R_{\tau}{}^{\nu\sigma\mu} + R_{\tau}{}^{\sigma\mu\nu} + R_{\tau}{}^{\mu\nu\sigma} = 0$, and *in both curved and flat spacetime*, we obtain:

$$\begin{aligned}P^{\sigma\mu\nu} &= D^{;\sigma}D^{;[\mu}G^{\nu]} + D^{;\mu}D^{;[\nu}G^{\sigma]} + D^{;\nu}D^{;[\sigma}G^{\mu]} \\ &= \left[D^{;\sigma}, D^{;\mu}\right]G^{\nu} + \left[D^{;\mu}, D^{;\nu}\right]G^{\sigma} + \left[D^{;\nu}, D^{;\sigma}\right]G^{\mu}. \\ &= V^{[\sigma\mu]}G^{\nu} + V^{[\mu\nu]}G^{\sigma} + V^{[\nu\sigma]}G^{\mu} = V^{([\sigma\mu]}G^{\nu)}\end{aligned} \tag{5.9}$$

The Yang-Mills electric and magnetic field equations (3.1), (3.2) expressed in the respective wholly equivalent forms of (5.3) and (5.9), 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 $V^{\mu\nu}$ of (5.4) and functions of this 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 (see [7] at page 356) 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 aspect P and a linear aspect L. Thus, from (5.3), 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 (5.3) 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, is given by:

$$\begin{aligned} J_P^\nu \equiv J_{YM}^\nu - J_L^\nu &= \left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + V + m^2\right) - \left(\partial^{;\mu}\partial^{;\nu} + V^{\mu\nu}\right)\right)G_\mu - \left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\mu}\partial^{;\nu}\right)G_\mu \\ &= \left(g^{\mu\nu}V - V^{\mu\nu}\right)G_\mu \end{aligned} . \quad (5.10)$$

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 the current density J_L^ν of Abelian 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 as we are reminded by (3.4) the monopole densities of Abelian gauge theory are zero, $P_L^{\sigma\mu\nu} = 0$. We know this of course from (3.4), but we also see this by inspection from (5.9) in which the non-vanishing magnetic monopole arises completely from the index-cyclical application of the antisymmetrized perturbation operator $V^{[\mu\nu]}$ to Yang-Mills gauge fields G^σ, i.e., from $P^{\sigma\mu\nu} = V^{([\sigma\mu]}G^{\nu)}$. If $V^{\mu\nu} \to 0$, clearly the monopole densities $P^{\sigma\mu\nu} \to 0$. Yang-Mills monopoles are thus 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. Those of course, are the four views of Yang-Mills theory that we have articulated so far. Now we turn to a fifth view, which is the *geometric curvature view* first articulated by Herrmann Weyl in the wake of Einstein's 1915 General Theory of Relativity [15] based on the curvature of spacetime.

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

Hermann Weyl in 1918 [16], [17] first conceived the idea that electrodynamics might be unified with gravitation by analyzing a "twisting" of vectors under parallel transport to measure the geometric curvature of a gauge space. While Weyl first conceived of this as a local "gauge" symmetry, in 1929 [18] he corrected his original misconception into the modern view of a local "phase" symmetry. Notwithstanding, the original misnomer "gauge" is still used to name Weyl's theory, perhaps as a reminder to posterity that even 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 $\Re^4$ first pioneered by Minkowski [19] based on Einstein's 1905 development of Lorentz invariance into Special Relativity [20], Weyl's theory operates along the circle of an abstract phase space using a non-observable 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 (5.8) already 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 the degree to which the spacetime derivatives do or do not commute. In fact, lowering all of the indexes on in (5.8), 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. That is:

$$g_{\tau\sigma}\left[D_{;\mu},D_{;\nu}\right]A^\tau=\left(R_{\tau\sigma\mu\nu}+g_{\tau\sigma}V_{[\mu\nu]}\right)A^\tau . \tag{6.1}$$

(Note that the ability to apply $\partial_{;\beta}A_\sigma=g_{\sigma\tau}\partial_{;\beta}A^\tau$ for raising and lowering indexes on a vector $A_\sigma=g_{\sigma\tau}A^\tau$ operated on by $\partial_{;\beta}$ relies on the metricity $\partial_{;\mu}g_{\nu\sigma}=0$ of the metric tensor $g_{\nu\sigma}$, and specifically, on the calculation $\partial_{;\beta}A_\sigma=\partial_{;\beta}\left(g_{\sigma\tau}A^\tau\right)=\partial_{;\beta}g_{\sigma\tau}A^\tau+g_{\sigma\tau}\partial_{;\beta}A^\tau=g_{\sigma\tau}\partial_{;\beta}A^\tau$. This will be implicitly used in a number of the upcoming index manipulations.) So just as $R_{\tau\sigma\mu\nu}$ represents curvature in spacetime, $g_{\tau\sigma}V_{[\mu\nu]}$ represents curvature in Weyl's gauge / phase space. We note the leading role of the anti-symmetrized perturbation $V_{[\mu\nu]}$ in this curvature connection space. It is also worth noting 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, so that both sides of (6.1) have balanced spacetime symmetries.

In fact, we can and should apply the same curvature analysis to the gauge-covariant derivative in curved spacetime, $D_{;\mu}=\partial_{;\mu}-iG_\mu$, which we now write operating on A_ν as:

$$D_{;\mu}A_\nu=\partial_{;\mu}A_\nu-iG_\mu A_\nu=\partial_\mu A_\nu-\Gamma^\alpha{}_{\mu\nu}A_\alpha-iG_\mu A_\nu . \tag{6.2}$$

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:

$$g_{\alpha\nu}D_{;\mu}A^\alpha=\left(g_{\alpha\nu}\partial_\mu-\Gamma_{\alpha\mu\nu}-ig_{\alpha\nu}G_\mu\right)A^\alpha . \tag{6.3}$$

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

$\Gamma_{\alpha\,\mu\nu} \rightarrow \Gamma_{\alpha\,\mu\nu} + ig_{\alpha\nu}G_\mu$ 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 (6.1), 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 (6.1), this is absorbed in the operation on A^α which sums out the α index. In contrast, however, the curvature operator $R_{\tau\sigma\mu\nu} + g_{\tau\sigma}V_{[\mu\nu]}$ in (6.1) is a tensor, but the parallel transport operator $\Gamma_{\alpha\,\mu\nu} + ig_{\alpha\nu}G_\mu$ in (6.3) 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 (6.1), 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 (3.4) 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 (3.6) 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 toward the later part of section 4. The latter Bianchi identity, 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 locally-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 $R_{\tau(\sigma\mu\nu)} = R_{\tau\sigma\mu\nu} + R_{\tau\mu\nu\sigma} + R_{\tau\nu\sigma\mu} = 0$. Because (6.1) contains $R_{\tau\sigma\mu\nu}$ which is the first term of this identity, let use rewrite (6.1) 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:

$$\begin{aligned}
&\left(g_{\tau\sigma}\left[D_{;\mu}, D_{;\nu}\right] + g_{\tau\mu}\left[D_{;\nu}, D_{;\sigma}\right] + g_{\tau\nu}\left[D_{;\sigma}, D_{;\mu}\right]\right)A^\tau \\
&= \left(R_{\tau\sigma\mu\nu} + R_{\tau\mu\nu\sigma} + R_{\tau\nu\sigma\mu} + g_{\tau\sigma}V_{[\mu\nu]} + g_{\tau\mu}V_{[\nu\sigma]} + g_{\tau\nu}V_{[\sigma\mu]}\right)A^\tau \\
&= \mathbf{0} + \left(g_{\tau\sigma}V_{[\mu\nu]} + g_{\tau\mu}V_{[\nu\sigma]} + g_{\tau\nu}V_{[\sigma\mu]}\right)A^\tau \\
&= g_{\tau(\sigma}\left[D_{;\mu}, D_{;\nu)}\right]A^\tau = g_{\tau(\sigma}V_{[\mu\nu])}A^\tau \\
&= \left[D_{;(\mu}, D_{;\nu}\right]A_{\sigma)} = V_{([\mu\nu]}A_{\sigma)}
\end{aligned} \quad . \tag{6.4}$$

Above we have applied $R_{\tau(\sigma\mu\nu)} = 0$ to zero out the terms that contain the Riemann tensor, so (6.4) now incorporates this first Bianchi identity. Once again the perturbation and the curvature views converge together. In fact, here, in contrast to (6.1) and (6.3), we *can* slice off the A^τ operand from the next-to-last line above and simply write the operator equation:

$$g_{\tau(\sigma}\left[D_{;\mu}, D_{;\nu)}\right] = g_{\tau(\sigma}V_{[\mu\nu])} . \tag{6.5}$$

This is allowed because the spacetime index symmetries on the left and right side of the above are fully matched, and so we do not need to sum out index the τ index to obtain matching spacetime symmetries. Contrasting to (5.8) written as $\left[D_{;\mu}, D_{;\nu}\right] = \left[\partial_{;\mu}, \partial_{;\nu}\right] + V_{[\mu\nu]}$, we see that (6.5) is an alternative way of stating the Bianchi identity $R_{\tau(\sigma\mu\nu)} = 0$ using Yang-Mills theory.

Let us now absorb the spacetime indexes in (6.4) to lower the indexes on the generalized vector A^τ, and then rename this into the *specific* vector $A_\mu \rightarrow G_\mu = \lambda^i G^i_\mu$ with represents the Yang-Mills gauge field. With this, also combining in (5.9), equation (6.4) becomes:

$$\begin{aligned} P_{\mu\nu\sigma} &= \left[D_{;\mu}, D_{;\nu}\right]G_\sigma + \left[D_{;\nu}, D_{;\sigma}\right]G_\mu + \left[D_{;\sigma}, D_{;\mu}\right]G_\nu = V_{[\mu\nu]}G_\sigma + V_{[\nu\sigma]}G_\mu + V_{[\sigma\mu]}G_\nu \\ &= \left[D_{;(\mu}, D_{;\nu}\right]G_{\sigma)} = V_{([\mu\nu]}G_{\sigma)} \end{aligned} . \tag{6.6}$$

Contrasting, this is *totally identical* to equation (5.9) for the Yang-Mills monopole, simply with covariant rather than contravariant indexes. Here we see a stark convergence of the perturbative and curvature views: 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 (6.5) – 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^τ.

Next, because (6.5) is valid standing alone as an operator equation, let us multiply this (in the expanded form of (6.4)) from the *left* by a general vector A^τ. Thus we now write:

$$A^\tau\left(g_{\tau\sigma}\left[D_{;\mu}, D_{;\nu}\right] + g_{\tau\mu}\left[D_{;\nu}, D_{;\sigma}\right] + g_{\tau\nu}\left[D_{;\sigma}, D_{;\mu}\right]\right) = A^\tau\left(g_{\tau\sigma}V_{[\mu\nu]} + g_{\tau\mu}V_{[\nu\sigma]} + g_{\tau\nu}V_{[\sigma\mu]}\right). \tag{6.7}$$

Upon lowering indexes this becomes:

$$\begin{aligned} &A_\sigma\left[D_{;\mu}, D_{;\nu}\right] + A_\mu\left[D_{;\nu}, D_{;\sigma}\right] + A_\nu\left[D_{;\sigma}, D_{;\mu}\right] = A_\sigma V_{[\mu\nu]} + A_\mu V_{[\nu\sigma]} + A_\nu V_{[\sigma\mu]} \\ &= A_{(\sigma}\left[D_{;\mu}, D_{;\nu)}\right] = A_{(\sigma}V_{[\mu\nu])} \end{aligned} . \tag{6.8}$$

Contrasting to the identity (6.4) written as $\left[D_{;(\mu}, D_{;\nu}\right]A_{\sigma)} = V_{([\mu\nu]}A_{\sigma)}$, we see that *any* vector $A_{\sigma)}$ may be commuted with $V_{[\mu\nu]}$ to obtain the "twin" identity $A_{(\sigma}\left[D_{;\mu}, D_{;\nu)}\right] = A_{(\sigma}V_{[\mu\nu])}$ when the spacetime indexes are cycled with $(\sigma\mu\nu)$. This will lead us to a "twin" of the Einstein equation in (7.6) infra, and is an important commutativity relationship to have in mind when we regard A_σ as an NxN matrix of vectors in Yang Mills theory, such as the gauge fields G_σ.

Speaking of which, let us do just that. If we again set $A_\mu \to G_\mu = \lambda^i G^i_\mu$ as we did for (6.6), then (6.8) becomes $G_{(\sigma}\left[D_{;\mu}, D_{;\nu)}\right] = G_{(\sigma} V_{[\mu\nu])}$, which is a “twin” of the magnetic monopole equation (6.6) 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 the bottom line of (6.4) and in (6.8), 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 the twins (6.4) and (6.8) into the single relationship:

$$\left[D_{;(\mu}, D_{;\nu}\right] D_{;\sigma)} = V_{([\mu\nu]} D_{;\sigma)} = D_{;(\sigma}\left[D_{;\mu}, D_{;\nu)}\right] = D_{;(\sigma} V_{[\mu\nu])} . \tag{6.9}$$

Because this commutes $D_{;(\sigma}$ to the left of the commutator $\left[D_{;\mu}, D_{;\nu)}\right]$ in $D_{;(\sigma}\left[D_{;\mu}, D_{;\nu)}\right]$, 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 this second Bianchi identity $\partial_{;(\alpha} R_{|\tau\sigma|\mu\nu)} = 0$, we define the notation $|\tau\sigma|$ as a “wall” to seal off the $\tau\sigma$ indexes (this is *not* an absolute value symbol as used here) from the $(\sigma\mu\nu)$ cycling of the remaining free indexes. But before we do this, let us work from the final expression in (6.6), use $iD_\sigma = i\partial_{;\sigma} + G_\sigma$ inverted into $G_{\sigma)} = iD_{\sigma)} - i\partial_{;\sigma)}$ to replace $G_{\sigma)}$, and then the final line apply the Jacobian identity (6.9). The result is:

$$\begin{aligned} P_{\mu\nu\sigma} &= \left[D_{;(\mu}, D_{;\nu}\right] G_{\sigma)} = V_{([\mu\nu]} G_{\sigma)} = \left[D_{;(\mu}, D_{;\nu}\right]\left(iD_{\sigma)} - i\partial_{;\sigma)}\right) = V_{([\mu\nu]}\left(iD_{\sigma)} - i\partial_{;\sigma)}\right) \\ &= i\left(\left[D_{;(\mu}, D_{;\nu}\right] D_{\sigma)} - \left[D_{;(\mu}, D_{;\nu}\right]\partial_{;\sigma)}\right) = i\left(V_{([\mu\nu]} D_{\sigma)} - V_{([\mu\nu]}\partial_{;\sigma)}\right) \quad . \\ &= i\left(D_{;(\sigma}\left[D_{;\mu}, D_{;\nu)}\right] - \left[D_{;(\mu}, D_{;\nu}\right]\partial_{;\sigma)}\right) = i\left(D_{;(\sigma} V_{[\mu\nu])} - V_{([\mu\nu]}\partial_{;\sigma)}\right) \end{aligned} \tag{6.10}$$

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

Now, let’s move on to the second Bianchi identity $\partial_{;(\alpha} R_{|\tau\sigma|\mu\nu)} = 0$. We start with (6.1) 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, and in the final line consolidate with the $(\alpha\mu\nu)$ cyclator to fashion:

$$
\begin{aligned}
&D_{;\alpha}\left(\left[D_{;\mu}, D_{;\nu}\right] A_\sigma\right)+D_{;\mu}\left(\left[D_{;\nu}, D_{;\alpha}\right] A_\sigma\right)+D_{;\nu}\left(\left[D_{;\alpha}, D_{;\mu}\right] A_\sigma\right) \\
&=D_{;\alpha}\left(R_{\tau\sigma\mu\nu} A^\tau\right)+D_{;\mu}\left(R_{\tau\sigma\nu\alpha} A^\tau\right)+D_{;\nu}\left(R_{\tau\sigma\alpha\mu} A^\tau\right)+D_{;\alpha}\left(V_{[\mu\nu]} A_\sigma\right)+D_{;\mu}\left(V_{[\nu\alpha]} A_\sigma\right)+D_{;\nu}\left(V_{[\alpha\mu]} A_\sigma\right). \quad (6.11) \\
&=D_{;(\alpha}\left(\left[D_{;\mu}, D_{;\nu)}\right] A_\sigma\right)=D_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)} A^\tau\right)+D_{;(\alpha}\left(V_{[\mu\nu])} A_\sigma\right)
\end{aligned}
$$

It should be clear how the term $D_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)} A^\tau\right)$ sets up the ability to apply and thereby embed the second Bianchi identity $\partial_{;(\alpha} R_{|\tau\sigma|\mu\nu)}=0$ into Yang-Mills theory. So now let's proceed.

We can slightly expand the compacted form in the bottom line of (6.11) 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 $\partial_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)} A^\tau\right)=\mathbf{0}+R_{\tau\sigma(\mu\nu}\partial_{;\alpha)} A^\tau$, thus obtaining:

$$
\begin{aligned}
D_{;(\alpha}\left(\left[D_{;\mu}, D_{;\nu)}\right] A_\sigma\right)&=\partial_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)} A^\tau\right)-iG_{(\alpha}\left(R_{|\tau\sigma|\mu\nu)} A^\tau\right)+D_{;(\alpha}\left(V_{[\mu\nu])} A_\sigma\right) \\
&=\partial_{;(\alpha} R_{|\tau\sigma|\mu\nu)} A^\tau+R_{\tau\sigma(\mu\nu}\partial_{;\alpha)} A^\tau-iG_{(\alpha}\left(R_{|\tau\sigma|\mu\nu)} A^\tau\right)+D_{;(\alpha}\left(V_{[\mu\nu])} A_\sigma\right). \qquad (6.12) \\
&=\mathbf{0}+R_{\tau\sigma(\mu\nu}\partial_{;\alpha)} A^\tau-iG_{(\alpha} R_{|\tau\sigma|\mu\nu)} A^\tau+D_{;(\alpha}\left(V_{[\mu\nu])} A_\sigma\right)
\end{aligned}
$$

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 (6.12) to make some of its meanings more transparent.

Continuing with (6.12), 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:

$$
\begin{aligned}
D_{;(\alpha}\left(\left[D_{;\mu}, D_{;\nu)}\right] A_\sigma\right) &= \partial_{;(\alpha}\left(R_{|\tau\sigma|\mu\nu)} A^\tau\right) - iG_{(\alpha}\left(R_{|\tau\sigma|\mu\nu)} A^\tau\right) + D_{;(\alpha}\left(V_{[\mu\nu])} A_\sigma\right) \\
&= R_{\tau\sigma(\mu\nu}\partial_{;\alpha)} A^\tau - iG_{(\alpha} R_{|\tau\sigma|\mu\nu)} A^\tau + D_{;(\alpha}\left(V_{[\mu\nu])} A_\sigma\right) \\
&= R_{\tau\sigma(\mu\nu}\partial_{;\alpha)} A^\tau - iR_{\tau\sigma(\mu\nu} G_{\alpha)} A^\tau + D_{;(\alpha}\left(V_{[\mu\nu])} A_\sigma\right) \\
&= R_{\tau\sigma(\mu\nu} D_{;\alpha)} A^\tau + D_{;(\alpha}\left(V_{[\mu\nu])} A_\sigma\right) \\
&= R_{\tau\sigma(\mu\nu} D_{;\alpha)} A^\tau + \partial_{;(\alpha}\left(V_{[\mu\nu])} A_\sigma\right) - iG_{(\alpha} V_{[\mu\nu])} A_\sigma \\
&= R_{\tau\sigma(\mu\nu} D_{;\alpha)} A^\tau + \partial_{;(\alpha} V_{[\mu\nu])} A_\sigma + V_{([\mu\nu]}\partial_{;\alpha)} A_\sigma - iG_{(\alpha} V_{[\mu\nu])} A_\sigma \\
&= R_{\tau\sigma(\mu\nu} D_{;\alpha)} A^\tau + D_{;(\alpha} V_{[\mu\nu])} A_\sigma + V_{([\mu\nu]}\partial_{;\alpha)} A_\sigma
\end{aligned}
\quad . \tag{6.13}
$$

Now the "odd duck" is the $V_{([\mu\nu]}\partial_{;\alpha)} A_\sigma$ which contains the only remaining ordinary covariant derivative $\partial_{;\alpha)}$ amidst all the other $D_{;(\alpha}$. But from (6.10) rearranged and right-multiplied by A_σ:

$$V_{([\mu\nu]}\partial_{;\alpha)} A_\sigma = D_{;(\alpha} V_{[\mu\nu])} A_\sigma + iP_{\mu\nu\alpha} A_\sigma \,, \tag{6.14}$$

which is why we wanted to make the one final connection in (6.10) before turning to $\partial_{;(\alpha} R_{|\tau\sigma|\mu\nu)} = 0$. So we use (6.14) in (6.13) to finally write (6.13) in terms of $P_{\mu\nu\alpha}$ as:

$$P_{\mu\nu\alpha} A_\sigma = iR_{\tau\sigma(\mu\nu} D_{;\alpha)} A^\tau - iD_{;(\alpha}\left(\left[D_{;\mu}, D_{;\nu)}\right] A_\sigma\right) + 2iD_{;(\alpha} V_{[\mu\nu])} A_\sigma \,. \tag{6.15}$$

This is our final result for the magnetic source density written as an operator operating on any vector A_σ*, and it embeds both of the Bianchi identities as well as the Jacobian identity*. We also manipulate indexes (implicitly using $\partial_{;\mu} g_{\nu\sigma} = 0$) to clearly display the spacetime symmetries:

$$g_{\sigma\tau} P_{\mu\nu\alpha} A^\tau = iR_{\tau\sigma(\mu\nu} D_{;\alpha)} A^\tau - ig_{\sigma\tau} D_{;(\alpha}\left(\left[D_{;\mu}, D_{;\nu)}\right] A^\tau\right) + 2ig_{\sigma\tau} D_{;(\alpha} V_{[\mu\nu])} A^\tau \,. \tag{6.16}$$

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 rewrite and reorder (6.15) with $A^\mu \to D^{;\mu}$ to form:

$$D_{;(\alpha}\left(\left[D_{;\mu}, D_{;\nu)}\right] D_{;\sigma}\right) = R_{\tau\sigma(\mu\nu} D_{;\alpha)} D^{;\tau} + iP_{\mu\nu\alpha} D_{;\sigma} + 2D_{;(\alpha} V_{[\mu\nu])} D_{;\sigma} \,. \tag{6.17}$$

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, as a function of the Riemann curvature, the monopole density, and the perturbation!

Finally, in flat spacetime, where $R_{\tau\sigma\mu\nu}=0$ and $D_{;\mu}\to D_{\mu}$, (6.15) reduces in view of (5.7), namely $V_{[\mu\nu]}=\left[D_{\mu},D_{\nu}\right]$, see also (6.10), simply to:

$$\begin{aligned}P_{\mu\nu\alpha}A_{\sigma}&=-iD_{(\alpha}\left(\left[D_{\mu},D_{\nu)}\right]A_{\sigma}\right)+2iD_{(\alpha}V_{[\mu\nu])}A_{\sigma}=-iD_{(\alpha}\left(\left[D_{\mu},D_{\nu)}\right]A_{\sigma}\right)+2iD_{(\alpha}\left[D_{\mu},D_{\nu)}\right]A_{\sigma}\\&=i\left(D_{(\alpha}\left[D_{\mu},D_{\nu)}\right]-\left[D_{(\mu},D_{\nu}\right]\partial_{\alpha)}\right)A_{\sigma}\end{aligned}\quad. \tag{6.18}$$

For $A_{\sigma}\to D_{\sigma}$, contrast (6.10), this becomes:

$$\begin{aligned}P_{\mu\nu\alpha}D_{\sigma}&=-iD_{(\alpha}\left(\left[D_{\mu},D_{\nu)}\right]D_{\sigma}\right)+2iD_{(\alpha}V_{[\mu\nu])}D_{\sigma}=-iD_{(\alpha}\left(\left[D_{\mu},D_{\nu)}\right]D_{\sigma}\right)+2iD_{(\alpha}\left[D_{\mu},D_{\nu)}\right]D_{\sigma}\\&=iD_{(\alpha}\left[D_{\mu},D_{\nu)}\right]D_{\sigma}-i\left[D_{(\mu},D_{\nu}\right]\partial_{\alpha)}D_{\sigma}\end{aligned}. \tag{6.19}$$

Alternatively and equivalently, explicitly showing a succession of four gauge-covariant derivatives in flat spacetime, the above becomes (contrast (6.17) for curved spacetime):

$$D_{(\alpha}\left[D_{\mu},D_{\nu)}\right]D_{\sigma}=\left(\left[D_{(\mu},D_{\nu}\right]\partial_{\alpha)}-iP_{\mu\nu\alpha}\right)D_{\sigma}\,. \tag{6.20}$$

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 identity (6.15), i.e., (6.16). Based on this embedding, however, we can go even further, to fully unify classical Yang-Mills gauge theory with classical gravitation.

7. The Classical 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 (6.15) aka (6.16), 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 reconfiguring (6.16) according to the following sequence of steps which apply $D_{;(\alpha}=\partial_{;(\alpha}-iG_{(\alpha}$ and the product rule for differentiation. The bottom line consolidates the second and fourth terms in the next-to-last line:

$$\begin{aligned}
&g_{\sigma\tau}P_{\mu\nu\alpha}A^{\tau}=iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^{\tau}-ig_{\sigma\tau}D_{;(\alpha}\left(\left[D_{;\mu},D_{;\nu)}\right]A^{\tau}\right)+2ig_{\sigma\tau}D_{;(\alpha}V_{[\mu\nu])}A^{\tau}\\
&=iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^{\tau}-ig_{\sigma\tau}\partial_{;(\alpha}\left(\left[D_{;\mu},D_{;\nu)}\right]A^{\tau}\right)-g_{\sigma\tau}G_{(\alpha}\left[D_{;\mu},D_{;\nu)}\right]A^{\tau}+2ig_{\sigma\tau}D_{;(\alpha}V_{[\mu\nu])}A^{\tau}\\
&=iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^{\tau}-ig_{\sigma\tau}\partial_{;(\alpha}\left[D_{;\mu},D_{;\nu)}\right]A^{\tau}-ig_{\sigma\tau}\left[D_{;(\mu},D_{;\nu)}\right]\partial_{;\alpha)}A^{\tau}-g_{\sigma\tau}G_{(\alpha}\left[D_{;\mu},D_{;\nu)}\right]A^{\tau}+2ig_{\sigma\tau}D_{;(\alpha}V_{[\mu\nu])}A^{\tau}\\
&=iR_{\tau\sigma(\mu\nu}D_{;\alpha)}A^{\tau}-ig_{\sigma\tau}D_{;(\alpha}\left[D_{;\mu},D_{;\nu)}\right]A^{\tau}-ig_{\sigma\tau}\left[D_{;(\mu},D_{;\nu}\right]\partial_{;\alpha)}A^{\tau}+2ig_{\sigma\tau}D_{;(\alpha}V_{[\mu\nu])}A^{\tau}
\end{aligned} \quad .(7.1)$$

Now, because A^{τ} is just a dummy operand which can be any four-vector, let us just lop it off of (7.1) 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 to designate the equality of the left and right sides of (7.1) when operating on A^{τ} which acquires a mismatched symmetry when the operand A^{τ} is removed. Thus, we now write:

$$g_{\sigma\tau}P_{\mu\nu\alpha}\text{"="}iR_{\tau\sigma(\mu\nu}D_{;\alpha)}-ig_{\sigma\tau}D_{;(\alpha}\left[D_{;\mu},D_{;\nu)}\right]-ig_{\sigma\tau}\left[D_{;(\mu},D_{;\nu}\right]\partial_{;\alpha)}+2ig_{\sigma\tau}D_{;(\alpha}V_{[\mu\nu])}\,. \quad (7.2)$$

The two sides of this equation are only equal when they operate on a vector A^{τ} as in (7.1), or if the symmetries can be restored in some other way. So we will need to now manipulate this such that the symmetries on both sides once again become matching and the equality is restored.

First, we fully expand the cyclators in (7.2) to obtain:

$$\begin{aligned}
g_{\sigma\tau}P_{\mu\nu\alpha}\text{"="}&iR_{\tau\sigma\mu\nu}D_{;\alpha}+iR_{\tau\sigma\nu\alpha}D_{;\mu}+iR_{\tau\sigma\alpha\mu}D_{;\nu}\\
&-ig_{\sigma\tau}D_{;\alpha}\left[D_{;\mu},D_{;\nu}\right]-ig_{\sigma\tau}D_{;\mu}\left[D_{;\nu},D_{;\alpha}\right]-ig_{\sigma\tau}D_{;\nu}\left[D_{;\alpha},D_{;\mu}\right]\\
&-ig_{\sigma\tau}\left[D_{;\mu},D_{;\nu}\right]\partial_{;\alpha}-ig_{\sigma\tau}\left[D_{;\nu},D_{;\alpha}\right]\partial_{;\mu}-ig_{\sigma\tau}\left[D_{;\alpha},D_{;\mu}\right]\partial_{;\nu}\\
&+2ig_{\sigma\tau}D_{;\alpha}V_{[\mu\nu]}+2ig_{\sigma\tau}D_{;\mu}V_{[\nu\alpha]}+2ig_{\sigma\tau}D_{;\nu}V_{[\alpha\mu]}
\end{aligned}\,. \quad (7.3)$$

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 via $R^{\tau\sigma}{}_{\tau\alpha}=-R^{\sigma}{}_{\alpha}$. This yields the intermediate result:

$$\begin{aligned}
g^{\sigma\tau}P_{\mu\tau\alpha}\text{"="}&iR^{\sigma}{}_{\mu}D_{;\alpha}-iR^{\sigma}{}_{\alpha}D_{;\mu}+iR^{\tau\sigma}{}_{\alpha\mu}D_{;\tau}\\
&-ig^{\sigma\tau}D_{;\alpha}\left[D_{;\mu},D_{;\tau}\right]-ig^{\sigma\tau}D_{;\mu}\left[D_{;\tau},D_{;\alpha}\right]-ig^{\sigma\tau}D_{;\tau}\left[D_{;\alpha},D_{;\mu}\right]\\
&-ig^{\sigma\tau}\left[D_{;\mu},D_{;\tau}\right]\partial_{;\alpha}-ig^{\sigma\tau}\left[D_{;\tau},D_{;\alpha}\right]\partial_{;\mu}-ig^{\sigma\tau}\left[D_{;\alpha},D_{;\mu}\right]\partial_{;\tau}\\
&+2ig^{\sigma\tau}D_{;\alpha}V_{[\mu\tau]}+2ig^{\sigma\tau}D_{;\mu}V_{[\tau\alpha]}+2ig^{\sigma\tau}D_{;\tau}V_{[\alpha\mu]}
\end{aligned}\,. \quad (7.4)$$

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:

$$\begin{aligned} P^{\tau}{}_{\tau\alpha} = iRD_{;\alpha} - iR^{\sigma}{}_{\alpha}D_{;\sigma} - iR^{\tau}{}_{\alpha}D_{;\tau} \\ - iD_{;\alpha}\left[D^{;\tau}, D_{;\tau}\right] - iD^{;\tau}\left[D_{;\tau}, D_{;\alpha}\right] - iD_{;\tau}\left[D_{;\alpha}, D^{;\tau}\right] \\ - i\left[D^{;\tau}, D_{;\tau}\right]\partial_{;\alpha} - i\left[D_{;\tau}, D_{;\alpha}\right]\partial^{;\tau} - i\left[D_{;\alpha}, D^{;\tau}\right]\partial_{;\tau} \\ + 2iD_{;\alpha}V^{[\tau}{}_{\tau]} + 2iD^{;\tau}V_{[\tau\alpha]} + 2iD_{;\tau}V_{[\alpha}{}^{\tau]} \end{aligned} . \tag{7.5}$$

Above, 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 (7.1) and it became restored when we contracted down to (7.5) 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 second, third and fourth lines also cancel out by inspection because of the various antisymmetries. So all that we have left in (7.5) 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:

$$-\kappa T^{\mu\nu}D_{;\nu} = \left(R^{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}R\right)D_{;\nu} = 0 . \tag{7.6}$$

This is the gravitational field equation of Yang-Mills theory! 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 so is a free derivative), *and it is a gauge-covariant derivative*. This is a "twin" of the Einstein equation. If we want to highlight the nexus to Yang-Mills theory in the clearest way possible, we may expand the above into the form:

$$-\kappa T^{\mu\nu}\left(\partial_{;\nu} - iG_{\nu}\right) = \left(R^{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}R\right)\left(\partial_{;\nu} - iG_{\nu}\right) = 0 . \tag{7.7}$$

The latter expression $\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right)\left(\partial_{;\nu} - iG_{\nu}\right) = 0$ fully marries Einstein's curvature view of spacetime with Weyl's curvature view of gauge theory, and is a geometric identity of Yang-Mills (and even Abelian) gauge theory arising from incorporating both Bianchi identities $R_{\tau(\sigma\mu\nu)} = 0$ and $\partial_{;(\alpha}R_{|\tau\sigma|\mu\nu)} = 0$ and the Jacobian $\left[a,\left[b,c\right]\right] + \left[b,\left[c,a\right]\right] + \left[c,\left[a,b\right]\right] = 0$ into the development of section 6 to arrive at (6.15). And, if we then use this to operate on some arbitrary vector A_{σ}, we may further expand this to:

$$\begin{aligned}0 &= -\kappa T^{\mu\nu}\left(\partial_{;\nu}A_\sigma - iG_\nu A_\sigma\right) = \left(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R\right)\left(\partial_{;\nu}A_\sigma - iG_\nu A_\sigma\right)\\ &= -\kappa T^{\mu\nu}\left(\partial_\nu A_\sigma - \Gamma^\tau{}_{\nu\sigma}A_\tau - iG_\nu A_\sigma\right) = \left(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R\right)\left(\partial_\nu A_\sigma - \Gamma^\tau{}_{\nu\sigma}A_\tau - iG_\nu A_\sigma\right)\\ &= -\kappa T^{\mu\nu}\left(\delta^\tau{}_\sigma\partial_\nu - \Gamma^\tau{}_{\nu\sigma} - i\delta^\tau{}_\sigma G_\nu\right)A_\tau = \left(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R\right)\left(\partial_\nu A_\sigma - \Gamma^\tau{}_{\nu\sigma}A_\tau - iG_\nu A_\sigma\right)\\ &= -\kappa T^{\mu\nu}\left(g_{\tau\sigma}\partial_\nu - \Gamma_{\tau\nu\sigma} - ig_{\tau\sigma}G_\nu\right)A^\tau = \left(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R\right)\left(g_{\tau\sigma}\partial_\nu - \Gamma_{\tau\nu\sigma} - ig_{\tau\sigma}G_\nu\right)A^\tau\end{aligned} . \tag{7.8}$$

By the connection $-\kappa T^{\mu\nu} = R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R$ to $T^{\mu\nu}$ (sans cosmological constant, which one can also inject into the development if desired), 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 than 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 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 (here we are *not* taking A_ν to be arbitrary *but making a <u>specific association</u> 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*. And because $\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)\left(\partial_{;\nu}-iG_\nu\right)=0$ can be applied to $SU(2)_W$ and $SU(3)_C$, we now have a complete classical unification of the field equations for all four known interactions: electromagnetic, weak, strong and gravitational! All of classical field theory is geometry! While recognizing the challenges of *tractable calculation*, unified quantum field theory then emerges, *in principle*, from the functional path integration $Z = \int D\phi \exp i\int \mathfrak{L} d^4x = \int D\phi \exp iS$ of the action $S = \int \mathfrak{L} d^4x$ for with the classical field equation $\left(R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R\right)\left(\partial_{;\nu}-iG_\nu\right)=0$ over all possible configurations $D\phi$ of the classical fields $\phi = g_{\mu\nu}, G^\mu$.

8. The Configuration Space Inverse of the Electric Charge Field Equation of Classical Yang-Mills Theory

Much of the focus in the last two sections was centered on the classical Yang-Mills 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 $\partial_{;\alpha}R_{\tau\sigma\mu\nu}+\partial_{;\mu}R_{\tau\sigma\nu\alpha}+\partial_{;\nu}R_{\tau\sigma\alpha\mu}=0$ which along with 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 (7.6) to (7.8). Now it is time to return our focus largely to the field equation (3.3) of the classical Yang-Mills 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 electric charge density field equation (3.3) 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 magnetic charge density field equation (3.6) 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 precise, 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 magnetic density $P^{\sigma\mu\nu}=-i\left(\partial^{;(\sigma}\left[G^\mu,G^{\nu)}\right]+G^{(\sigma}D^{;[\mu}G^{\nu])}\right)$ of (3.6) aka (4.1) 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 fields always originate from some source, then the natural progression from (3.6), (4.1) 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 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, and 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 (3.3). 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 4 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 fermion eigenstates exhibit the same symmetries as the quarks which 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 $G_\mu \equiv I_{\tau\mu} J^\tau = I_{\tau\mu}\overline{\Psi}\gamma^\tau\Psi$.

Specifically, we now *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. As we shall now do, it helps to review how this inversion is done in electrodynamics, to prepare for the more complicated calculation required for Yang-Mills theory.

In $U(1)_{em}$ electrodynamics, we use the classical field equation mentioned between (3.3) and (3.4) to specify this inverse $G_\mu \equiv I_{L\tau\mu} J^\tau$, namely:

$$J^\nu = \left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\mu}\partial^{;\nu}\right) G_\mu = \delta^\nu{}_\tau J^\tau \equiv \left(g^{\mu\nu}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\mu}\partial^{;\nu}\right) I_{L\tau\mu} J^\tau \quad . \tag{8.1}$$

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 a 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 $I_{YM\,\tau\mu}$ arising 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 (5.10) where we found that the perturbative-only contribution to the current density is $J_P^\nu = \left(g^{\mu\nu}V - V^{\mu\nu}\right) G_\mu$. So now, we are effectively seeking the inverse of this.

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

$$\delta^\mu{}_\nu \equiv \left(g^{\mu\tau}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\tau}\partial^{;\mu}\right) I_{L\nu\tau} \quad . \tag{8.2}$$

Looking at the configuration space operator $g^{\mu\tau}\left(\partial_{;\sigma}\partial^{;\sigma} + m^2\right) - \partial^{;\tau}\partial^{;\mu}$, we see that in flat spacetime this is symmetric in its μ,τ indexes, but in curved spacetime it is 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 (3.4), and so left-right ordering matters. Especially since the non-Abelian $g^{\mu\tau}\left(D_{;\sigma}D^{;\sigma} + m^2\right) - D^{;\mu}D^{;\tau}$ in (3.3) 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* because of $V^{[\mu\nu]} = \left[D^\mu, D^\nu\right]$ in (5.7), 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, the 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 (8.2), 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 (8.2) to specify:

$$\delta^\mu{}_\nu \equiv \left(g^{\mu\tau}\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)-\partial^{;\tau}\partial^{;\mu}\right)\left(g_{\nu\tau}A+\partial_{;\nu}\partial_{;\tau}B\right) , \tag{8.3}$$

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 right at the outset, because this will be especially important when we progress to Yang-Mills theory.

Now we expand (8.3) to obtain:

$$\delta^\mu{}_\nu = \delta^\mu{}_\nu\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)A - \partial_{;\nu}\partial^{;\mu}A + \left(\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)\partial_{;\nu}\partial^{;\mu} - \partial^{;\tau}\partial^{;\mu}\partial_{;\nu}\partial_{;\tau}\right)B , \tag{8.4}$$

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 precisely, by left-multiplying with $\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)^{-1}$ to write:

$$A = \left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)^{-1}\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)A = \left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)^{-1} , \tag{8.5}$$

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 (8.5) back into (8.4) while maintaining all the "overkill" of ordering and taking inverses yields, with some rearrangement:

$$\partial^{;\nu}\partial^{;\mu}\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)^{-1} = \left(\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)\partial^{;\nu}\partial^{;\mu} - \partial_{;\tau}\partial^{;\mu}\partial^{;\nu}\partial^{;\tau}\right)B . \tag{8.6}$$

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:

$$B = \left(\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)\partial^{;\nu}\partial^{;\mu} - \partial_{;\tau}\partial^{;\mu}\partial^{;\nu}\partial^{;\tau}\right)^{-1}\partial^{;\nu}\partial^{;\mu}\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)^{-1} . \tag{8.7}$$

Now using (8.5) and (8.7) in $I_{L\nu\tau} \equiv g_{\nu\tau}A + \partial_{;\nu}\partial_{;\tau}B$ we obtain:

$$I_{\nu\tau} = \left[g_{\nu\tau} + \partial_{;\nu}\partial_{;\tau} \left(\left(\partial_{;\sigma}\partial^{;\sigma} + m^2 \right) \partial^{;\alpha}\partial^{;\beta} - \partial_{;\sigma}\partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma} \right)^{-1} \partial^{;\alpha}\partial^{;\beta} \right] \left(\partial_{;\sigma}\partial^{;\sigma} + m^2 \right)^{-1} . \tag{8.8}$$

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 (8.8) becomes:

$$I_{L\nu\tau} = \frac{g_{\nu\tau} + \dfrac{\partial_{;\nu}\partial_{;\tau}\partial^{;\alpha}\partial^{;\beta}}{\left(\partial_{;\sigma}\partial^{;\sigma} + m^2 \right)\partial^{;\alpha}\partial^{;\beta} - \partial_{;\sigma}\partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma}}}{\partial_{;\sigma}\partial^{;\sigma} + m^2} . \tag{8.9}$$

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

$$I_{L\nu\tau} = \frac{g_{\nu\tau} + \dfrac{\partial_\nu\partial_\tau}{m^2}}{\partial_\sigma\partial^\sigma + m^2} = \frac{-g_{\nu\tau} + \dfrac{k_\nu k_\tau}{m^2}}{k_\sigma k^\sigma - m^2} \overset{+i\varepsilon}{\Rightarrow} \frac{-g_{\nu\tau} + \dfrac{k_\nu k_\tau}{m^2}}{k_\sigma k^\sigma - m^2 + i\varepsilon} . \tag{8.10}$$

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 (8.10) traceable to (8.2)), which yields:

$$G_\nu = \frac{-g_{\tau\nu} + \dfrac{k_\tau k_\nu}{m^2}}{k_\sigma k^\sigma - m^2 + i\varepsilon} J^\tau = -\frac{1}{k_\sigma k^\sigma - m^2 + i\varepsilon} J_\nu \overset{m=0}{\Rightarrow} -\frac{1}{k_\sigma k^\sigma + i\varepsilon} J_\nu . \tag{8.11}$$

After a final flat spacetime commutation $[\partial_\nu, \partial_\tau] = -[k_\nu, k_\tau] = 0$, 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., [7] after I.5(4)).

Now, it is easy to see from (8.10) as $m \to 0$, via $k_\nu k_\tau / m^2 \to \infty$, that $I_{L\nu\tau} \to \infty$. This 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., [7] section 3.4). But what happens in curved spacetime when we use $+i\varepsilon$, and set $m \to 0$? This will be instructive for our momentary consideration of Yang-Mills. In this circumstance, using (8.9) in $G_\nu = I_{L\tau\nu}J^\tau$, the inverse equation corresponding to (8.11), becomes:

$$G_\nu \equiv \frac{g_{\tau\nu} + \dfrac{\partial_{;\tau}\partial_{;\nu}\partial^{;\alpha}\partial^{;\beta}}{\left(\partial_{;\sigma}\partial^{;\sigma}+m^2\right)\partial^{;\alpha}\partial^{;\beta} - \partial_{;\sigma}\partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma}}}{\partial_{;\sigma}\partial^{;\sigma}+m^2} J^\tau \underset{+i\varepsilon}{\overset{m=0}{\Rightarrow}} \frac{g_{\tau\nu} + \dfrac{\partial_{;\tau}\partial_{;\nu}\partial^{;\alpha}\partial^{;\beta}}{\partial_{;\sigma}\left(\partial^{;\sigma}\partial^{;\alpha}\partial^{;\beta} - \partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma}\right)}}{\partial_{;\sigma}\partial^{;\sigma}+i\varepsilon} J^\tau \ . \tag{8.12}$$

Even with $m=0$, none of the reductions of (8.10) or (8.11) occur. To obtain $\partial_{;\nu}\partial^{;\alpha}\partial^{;\beta}\partial_{;\tau}J^\tau = 0$ from $\partial_{;\tau}\partial_{;\nu}\partial^{;\alpha}\partial^{;\beta}J^\tau$ using $\partial_{;\tau}J^\tau = 0$, 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 add $+i\varepsilon$), as we have done on the rightmost expression above. This, of course, describes the photon. Even with $m=0$, so long as we use $+i\varepsilon$, the inverse is only singular in the circumstance where $\partial^{;[\sigma}\partial^{;\alpha}\partial^{;\beta]} = \partial^{;\sigma}\partial^{;\alpha}\partial^{;\beta} - \partial^{;\beta}\partial^{;\alpha}\partial^{;\sigma} = 0$, 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 (8.12) 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 non-infinite behaviors due to non-commuting derivatives will 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 Yang-Mills electric field strength (3.3) which we cast in a form analogous to (8.1), namely:

$$J^\nu = \left(g^{\mu\nu}\left(D_{;\sigma}D^{;\sigma}+m^2\right) - D^{;\mu}D^{;\nu}\right)G_\mu = \delta^\nu{}_\tau J^\tau \equiv \left(g^{\mu\nu}\left(D_{;\sigma}D^{;\sigma}+m^2\right) - D^{;\mu}D^{;\nu}\right)I_{YM\,\tau\mu}J^\tau , \tag{8.13}$$

where $I_{YM\,\tau\mu}$ is now the Yang-Mills inverse and we *define* $G_\mu \equiv I_{YM\,\tau\mu}J^\tau$ 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 (8.2) to (8.8), but now the "overkill" of being very careful about inverses and left-right ordering is essential. Completely analogously to (8.8), but with the Yang-Mills "minimal coupling" discussed in relation to the "gauge theory on steroids" view of (2.6), with the simple replacement of $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$, we obtain:

$$I_{YM\,\nu\tau} = \left[g_{\nu\tau} + D_{;\nu}D_{;\tau}\left(m^2 D^{;\alpha}D^{;\beta} + D_{;\sigma}D^{;\sigma}D^{;\alpha}D^{;\beta} - D_{;\sigma}D^{;\beta}D^{;\alpha}D^{;\sigma}\right)^{-1} D^{;\alpha}D^{;\beta}\right]\left(D_{;\sigma}D^{;\sigma}+m^2\right)^{-1} . \tag{8.14}$$

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 (8.14) in a way that facilitates visual comparison to (8.9) for Abelian gauge theory, we shall now adopt a "quoted denominator" notation 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, (8.14) now is written as:

$$I_{YM\,\nu\tau}=\frac{g_{\nu\tau}+\dfrac{D_{;\nu}D_{;\tau}\ {}_{\vee}\ D^{;\alpha}D^{;\beta}}{"m^2D^{;\alpha}D^{;\beta}+D_{;\sigma}D^{;\sigma}D^{;\alpha}D^{;\beta}-D_{;\sigma}D^{;\beta}D^{;\alpha}D^{;\sigma}"\ {}_{\vee}}}{"D_{;\sigma}D^{;\sigma}+m^2"}\ . \tag{8.15}$$

By comparison to (8.9), we see in stark relief the manner in which classical Yang-Mills gauge theory is simply Gauge theory on steroids with the minimal coupling principle $\partial^\mu\to D^\mu=\partial^\mu-iG^\mu$. One should note two factorizations which are available in the upper denominator of (8.15). 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 (8.9) to (8.10) 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 (8.12), this helps avert a singular numerator even if we set $m=0$, because this will remain finite to the degree that $D^{;\sigma}D^{;\alpha}D^{;\beta}-D^{;\beta}D^{;\alpha}D^{;\sigma}=D^{;[\sigma}D^{;\alpha}D^{;\beta]}\neq 0$. In section 10, this elimination of the Proca mass, $m\to 0$, will be of particular interest for filling the mass gap.

We note finally, referring back to sections 6 and 7, 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 (7.6). Along the way, beginning with (6.9), we obtained several useful identities involving the commutativity properties of taking three or four successive covariant derivatives. Clearly, based on these identities, as a general rule, $D^{;[\sigma}D^{;\alpha}D^{;\beta]}\neq 0$. Thus, *(8.15) 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 (8.15) is the inverse are placed on shell without $+i\varepsilon$*. These properties of (8.15) will become essential for filling the mass gap.

9. Populating Yang-Mills Monopoles with Fermions, and the Recursive Nature of the Yang-Mills: A Sixth View of Yang-Mills which may Aid in the Quantum Path Integration of Yang-Mills Theory

We will examine (8.14) and (8.15) much more closely in the next section when we finally turn directly to the mass gap solution. 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 (3.6) to populate the magnetic monopole density (3.6) 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:

$$\begin{aligned}P^{\sigma\mu\nu} &= -i\left(\partial^{;(\sigma}\left[I_{YM}^{\alpha\mu}J_\alpha, I_{YM}^{\beta\nu)}J_\beta\right] + I_{YM}^{\tau(\sigma}J_\tau D^{;[\mu}I_{YM}^{\beta\nu])}J_\beta\right)\\ &= -i\left(\partial^{;(\sigma}\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_{YM}^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{;[\mu}I_{YM}^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right)\\ &= -i\begin{pmatrix}\partial^{;\sigma}\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu}\overline{\Psi}\gamma_\beta\Psi\right] + \partial^{;\mu}\left[I_{YM}^{\alpha\nu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\sigma}\overline{\Psi}\gamma_\beta\Psi\right] + \partial^{;\nu}\left[I_{YM}^{\alpha\sigma}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\mu}\overline{\Psi}\gamma_\beta\Psi\right]\\ +I_{YM}^{\tau\sigma}\overline{\Psi}\gamma_\tau\Psi D^{;[\mu}I_{YM}^{\beta\nu]}\overline{\Psi}\gamma_\beta\Psi + I_{YM}^{\tau\mu}\overline{\Psi}\gamma_\tau\Psi D^{;[\nu}I_{YM}^{\beta\sigma]}\overline{\Psi}\gamma_\beta\Psi + I_{YM}^{\tau\nu}\overline{\Psi}\gamma_\tau\Psi D^{;[\sigma}I_{YM}^{\beta\mu]}\overline{\Psi}\gamma_\beta\Psi\end{pmatrix}\end{aligned} \quad . \quad (9.1)$$

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 linear plus non-linear/perturbative (L+P) aspects of Yang-Mills gauge theory are fully included in the above. *This is the complete Yang-Mills monopole with all non-linearity included.* Now we shall study this monopole from a range of viewpoints.

First, it is critically-important to observe that if we wish to do so, to obtain an even more detailed expression we may explicitly substitute into (9.1), the $I_{YM\nu\tau}$ of (8.14) with a renaming and raising of some indexes. 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 additional inverses 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 (9.1) represents this Yang-Mills monopole in its most compact form, this is a <u>*recursive*</u> expression because of the fact that if we use (8.14) in $G_\mu \equiv I_{YM\,\tau\mu}J^\tau$ to write gauge field G_τ in terms of the current density J^ν via (contrast the Abelian (8.12)):

$$G_\tau = \left[g_{\nu\tau} + D_{;\nu}D_{;\tau}\left(m^2 D^{;\alpha}D^{;\beta} + D_{;\sigma}D^{;\sigma}D^{;\alpha}D^{;\beta} - D_{;\sigma}D^{;\beta}D^{;\alpha}D^{;\sigma}\right)^{-1}D^{;\alpha}D^{;\beta}\right]\left(D_{;\sigma}D^{;\sigma} + m^2\right)^{-1}J^\nu, \quad (9.2)$$

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.* Then, via $G_\mu \equiv I_{YM\,\tau\mu}J^\tau$, we may generate a similar recursion embedding the current densities J^τ and more gauge fields.

In other words, it is very important to observe that (9.2), and so (9.1), 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*. And via $G_\mu \equiv I_{YM\,\tau\mu}J^\tau$, this in turn cascades into an infinite nesting of current densities and thus fermion wavefunctions. It may well be possible to discern the patterns and develop a closed form of (9.2), but for the moment, we simply note that this recursion is yet a *sixth view* of Yang-Mills gauge theory. To summarize: Yang Mills field theory is *1) non-commuting, 2) non-linear, 3) steroidal, 4) perturbative, 5) geometrically-curved and now 6), based on (9.2),* <u>*recursive*</u>. 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 (9.1), 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 (6.16) used in section 7, starting with (7.1), 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 (7.6).

Having found this recursive aspect to Yang-Mills theory, we now return to Jaffe and Witten who on page 7 of [1], 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 (if it is convergent or divergent) in the limit of infinite recursive nesting.

It will of course be of great interest to examine the behavior of (8.14) a.k.a. (8.15) to see if it is exhibits suitable convergence under infinite recursive nesting, and how this relates to expressions obtained during efforts to quantize Yang-Mills. If we look at the numerator *N* in (8.15) 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 noting 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 (9.2) into closed form. 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:

$$x = 1 + \frac{Bx}{1} \to 1 + \frac{B(1+\frac{Bx}{2})}{1} \to 1 + \frac{B(1+\frac{B(1+\frac{Bx}{3})}{2})}{1} \to 1 + \frac{B(1+\frac{B(1+\frac{B(1+\frac{Bx}{4})}{3})}{2})}{1}\ldots \tag{9.3}$$
$$= 1 + B + \frac{1}{2!}B^2 + \frac{1}{3!}B^3 + \frac{1}{4!}B^4 x \overset{n\to\infty}{\to} e^B$$

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 (8.15) 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 Maclaurin series*, and then, armed with those definitions, to take a fresh look at (9.2) 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 (9.2), in light of (9.3), which is a mathematical insight with possible physical implications is this: In (9.3) x is a "dummy" variable that gets stripped away in the infinite application of recursion. This means that in (9.2) the gauge field G_τ is the dummy variable that will get stripped away by the recursion as the nesting reaches infinity, and that what will remain behind is the single G_τ on the left of (9.2) expressed as an infinite recursive series with up to infinite powers of the source currents J^ν. Possibly analogously, when we take a path integral, such as in QED:

$$\begin{aligned} Z &= \int DG_\mu \exp i \int d^4x \left(\tfrac{1}{2} G_\mu \left(g^{\mu\nu} \left(\partial_\sigma \partial^\sigma + m^2 \right) - \partial^\mu \partial^\nu \right) G_\nu - J^\mu G_\mu \right) \\ &\equiv \mathcal{C} \exp\left(iW\left(J\right)\right) = \mathcal{C} \exp\left(-\frac{1}{2} i \int \frac{d^4k}{\left(2\pi\right)^4} J^\mu \frac{-g_{\mu\nu} + \dfrac{k_\mu k_\nu}{m^2}}{k_\sigma k^\sigma - m^2} J^\nu \right) \end{aligned} \quad , \tag{9.4}$$

the gauge field G_τ is the "dummy" variable of integration, it also gets stripped away as the integration takes place, and what is left behind is an amplitude expression with up to infinite powers of the source currents J^ν.

With this in mind, using what Zee [7] 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):

$$\int D\phi \exp\left(-\tfrac{1}{2} \phi \cdot K \cdot \phi - V\left(\phi\right) - J \cdot \phi \right) = \mathcal{C} \exp\left(V\left(\delta / \delta J \right)\right) \exp\left(\tfrac{1}{2} J \cdot K^{-1} \cdot J \right), \tag{9.5}$$

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

$$\int dx \exp\left(-\tfrac{1}{2} Ax^2 - Jx \right) = \left(-2\pi / A \right)^{.5} \exp\left(J^2 / 2A \right) \tag{9.6}$$

can be fully reformulated in terms of a recursive function. As a 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 (9.3)) in the form $((x))$. Thus, in this (possibly new) notation, we may write (9.3) in compact form as:

$$e^B \equiv \left(\left(\;\right)\right)_{n=1}^{\infty}\left(1+B\left(\left(x\right)\right)/n\right) \tag{9.7}$$

This means that the Gaussian integral (9.6) may be *recursively* written as:

$$\int dx \exp\left(-\tfrac{1}{2}Ax^2-Jx\right)=\left(2\pi/A\right)^{.5}\exp\left(J^2/2A\right)=\left(2\pi/A\right)^{.5}\left(\left(\;\right)\right)_{n=1}^{\infty}\left(1+\frac{J^2\left(\left(x\right)\right)}{2An}\right) \tag{9.8}$$

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 (9.1) and (9.2), 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. One should amidst this analysis, be looking for ways to analytically calculate the exact Yang-Mills path integral with the aid of the recursive kernel in (9.2) which does mirror the types of terms that get fed into the Yang-Mills path integral. In section 13, we shall do exactly that. All of the forgoing also applies to gravitation theory, which from a "gravitation gravitates" view possesses a similar sort or recursive non-linearity.

It is also worth observing that the magnetic monopole (9.1), now populated with fermions (which in section 11 we will show are quarks) is really, at bottom, a non-Abelian combination of *both* of Maxwell's classical equations (3.1) and (3.2) into a *single equation*. Specifically, the Yang-Mills 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 (9.2) and then inserted into the monopole density (3.6) to arrive at (9.1). Einstein, in his final paper [21] 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 that (9.1), 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 into one*, will also have a strength $z_1=12$ interrelating its $i=1,2,3...N^2-1$ Abelian sources $J^{i\nu}$, $P^{i\sigma\mu\nu}$ and fields $G^{i\mu}$, and so also have the same strength as $R_{\mu\nu}=0$.

The final, very important point to note is that because of its origin in (3.2) and (3.6) as a Yang-Mills monopole, (9.1) contains *three additive terms* in index-cyclic $(\sigma\mu\nu)$ 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 Yang-Mills fermions Ψ. This was the source of the "three-ness" discussed at some length toward the end of section 4. 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 (9.1) via (9.2) 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 (9.1) will be teeming with non-linear physics and many additional quarks and antiquarks and amidst a sea of gluons that arise at the first, second, thousandth, and millionth recursive order. This will all be developed in detail in section 11.

In this light, as stated in the introduction, and as we shall detail in the forthcoming development of section 11, QCD is not a theory of first principle, it is *corollary theory*. 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 make three specific showings: First, in section 10, we shall show how the relationship (8.14) which via $G_\mu \equiv I_{YM\,\tau\mu} J^\tau$ is contained to infinite recursive order in monopole (9.1) via (9.2), fills the mass gap. To preview: if we set $m = 0$ in (8.14), 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 $\left(D_{;\sigma} D^{;\sigma}\right)^{-1}$ in (8.14), yield a mass eigenvalue spectrum, which one expects will come to be associated with the non-zero masses of the observed mesons such as those catalogued in [22]. Second, in section 11, as has already been developed to some degree in section 4, we shall show from a more formal standpoint how and why (9.1) contains all of the expected color symmetries of a baryon, and at the same time confines its fermions (which we shall identify with quarks) and its gauge fields (which will identify with gluons), while permitting the flux of colorless quark combinations that we observe in the form of mesons. It is by this means that we shall identify the magnetic monopoles of

Yang-Mills gauge theory as baryons, which naturally possess three colored quarks at the lowest recursive order and only permit a flow of mesons across their closed surfaces. Finally, in section 12 we shall examine the natural origin of chiral symmetry breaking, and particularly, of the various vector (V) and axial (A) meson states which are phenomenologically-evidenced in strong interactions.

10. The Mass Gap Solution

Let us now show how the solution to the mass gap that is embedded in equation (8.14) which in the form (9.2) yields an infinite recursion. We shall develop this solution using the more "user-friendly" representation (8.15).

The configuration space inverse $I_{YM\,\nu\tau}$ in (8.15), upon expansion of each $D^\mu = \partial^\mu - iG^\mu$ followed by reapplication of (9.2), 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 which represent the "difference" between Yang-Mills gauge theory and an Abelian gauge theory such as Maxwell's electrodynamics. As we did earlier with (5.10), we use 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 ([7], page 356) inverse (8.15) and the linear inverse (8.9). 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 that we can study *only* the impact of Yang-Mills theory separated from any impact due to spacetime curvature, we represent both of (8.9) and (8.15) in flat spacetime, and so turn the gravitationally-covariant derivatives $\partial^{;\sigma}$ into ordinary ones ∂^σ. Thus, we form:

$$
\begin{aligned}
&I_{P\nu\tau} = I_{YM\,\nu\tau} - I_{L\nu\tau} \\
&= \frac{g_{\nu\tau} + \dfrac{D_\nu D_{\tau\,\vee}\, D^\alpha D^\beta}{"m^2 D^\alpha D^\beta + D_\sigma D^\sigma D^\alpha D^\beta - D_\sigma D^\beta D^\alpha D^\sigma"_{\vee}}}{"D_\sigma D^\sigma + m^2"} - \frac{g_{\nu\tau} + \dfrac{\partial_\nu \partial_\tau \partial^\alpha \partial^\beta}{m^2 \partial^\alpha \partial^\beta + \partial_\sigma \partial^\sigma \partial^\alpha \partial^\beta - \partial_\sigma \partial^\beta \partial^\alpha \partial^\sigma}}{\partial_\sigma \partial^\sigma + m^2}
\end{aligned}
\quad .(10.1)
$$

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 (8.9) to (8.10) where we did the same reduction earlier):

$$
\begin{aligned}
&I_{P\nu\tau} = I_{YM\,\nu\tau} - I_{L\nu\tau} \\
&= \frac{g_{\nu\tau} + \dfrac{D_\nu D_{\tau\,\vee}\, D^\alpha D^\beta}{"m^2 D^\alpha D^\beta + D_\sigma D^\sigma D^\alpha D^\beta - D_\sigma D^\beta D^\alpha D^\sigma"_{\vee}}}{"D_\sigma D^\sigma + m^2"} - \frac{g_{\nu\tau} + \dfrac{\partial_\nu \partial_\tau}{m^2}}{\partial_\sigma \partial^\sigma + m^2}\ .
\end{aligned}
\quad (10.2)
$$

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 (8.10), 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 (9.4). 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 6 and especially (6.9) and (6.17) to (6.20), this term $D_\sigma D^{[\sigma} D^\alpha D^{\beta]}$ is non-vanishing everywhere there are non-zero perturbations as defined in (5.3) to (5.5)..

Now let us return to (5.6) for two successive gauge-covariant derivatives, and write this perturbatively in momentum space in flat spacetime via $i\partial^\mu \to k^\mu$, as

$$D^{;\mu} D^\nu = -k^\mu k^\nu + V^{\mu\nu} = -k^\mu k^\nu - k^\mu G^\nu - G^\mu k^\nu - G^\mu G^\nu \,, \tag{10.3}$$

which also means that:

$$V^{\mu\nu} \equiv -k^\mu G^\nu - G^\mu k^\nu - G^\mu G^\nu \,. \tag{10.4}$$

So we expand the various $D^\mu D^\nu = -k^\mu k^\nu + V^{\mu\nu}$ in (10.2) and convert into momentum space, to obtain:

$$\begin{aligned} I_{P\nu\tau} &= I_{YM\,\nu\tau} - I_{L\nu\tau} \\ &= \frac{-g_{\nu\tau} + \dfrac{\left(k_\nu k_\tau - V_{\nu\tau}\right)_{\vee}\left(-k^\alpha k^\beta + V^{\alpha\beta}\right)}{\text{"}m^2\left(-k^\alpha k^\beta + V^{\alpha\beta}\right) + \left(k_\sigma k^\sigma - V_\sigma{}^\sigma\right)\left(k^\alpha k^\beta - V^{\alpha\beta}\right) - \left(k_\sigma k^\beta - V_\sigma{}^\beta\right)\left(k^\alpha k^\sigma - V^{\alpha\sigma}\right)\text{"}_{\vee}}}{\text{"}k_\sigma k^\sigma - V_\sigma{}^\sigma - m^2\text{"}} \\ &\quad - \frac{-g_{\nu\tau} + \dfrac{k_\nu k_\tau}{m^2}}{k_\sigma k^\sigma - m^2} \end{aligned} \,. \tag{10.5}$$

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 (10.5), the m^2 is the Proca mass of the Yang-Mills gauge bosons, introduced by hand back in (3.3). 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 in the physics we are describing –be able to reintroduce in some other way without ruining the gauge invariance and renormalizability of the theory. So now, the 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 $I_{L\nu\tau}$ to keep one "red flag" in place as will be momentarily discussed. With setting $m^2 = 0$ in $I_{YM\,\nu\tau}$, the above now reduces to:

$$I_{P\nu\tau} = I_{YM\,\nu\tau} - I_{L\nu\tau}$$
$$= \frac{-g_{\nu\tau} + \dfrac{\left(k_\nu k_\tau - V_{\nu\tau}\right)_{\vee}\left(-k^\alpha k^\beta + V^{\alpha\beta}\right)}{"\left(k_\sigma k^\sigma - V_\sigma{}^\sigma\right)\left(k^\alpha k^\beta - V^{\alpha\beta}\right) - \left(k_\sigma k^\beta - V_\sigma{}^\beta\right)\left(k^\alpha k^\sigma - V^{\alpha\sigma}\right)"_{\vee}}}{"k_\sigma k^\sigma - V_\sigma{}^\sigma"} - \frac{-g_{\nu\tau} + \dfrac{k_\nu k_\tau}{m^2}}{k_\sigma k^\sigma - m^2} \,. \qquad (10.6)$$

This means that (3.3) is now reverted to $J^\nu = \left(g^{\mu\nu} D_{;\sigma} D^{;\sigma} - D^{;\mu} D^{;\nu}\right) G_\mu$, so that the Yang-Mills gauge bosons are now massless. This means, for example, that if our gauge group is SU(3)$_C$, then these gauge bosons will be massless gluons.

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 itself), 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 the Faddeev-Popov method and was originally developed by Feynman, see, e.g., [7] section III.4. The latter $\xi = 0$ is the gauge of (10.6). Let us also raise the free ν index everywhere. Thus, (10.6) now becomes:

$$I_P{}^\nu{}_\tau = I_{YM}{}^\nu{}_\tau - I_L{}^\nu{}_\tau$$
$$= \frac{-\delta^\nu{}_\tau + \dfrac{\left(k^\nu k_\tau - V_{\nu\tau}\right)_{\vee}\left(-k^\alpha k^\beta + V^{\alpha\beta}\right)}{"V_\sigma{}^\sigma\left(V^{\alpha\beta} - k^\alpha k^\beta\right) + k_\sigma k^\beta V^{\alpha\sigma} - V_\sigma{}^\beta V^{\alpha\sigma} + V_\sigma{}^\beta k^\alpha k^\sigma"_{\vee}}}{"-V_\sigma{}^\sigma"} - \frac{-\delta^\nu{}_\tau + \left(1-\xi\right)\dfrac{k^\nu k_\tau}{m^2}}{k_\sigma k^\sigma - m^2 + i\varepsilon} \,. \qquad (10.7)$$

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 (8.11). Thus, (10.7) now becomes:

$$I_P{}^\nu{}_\tau = I_{YM}{}^\nu{}_\tau - I_L{}^\nu{}_\tau = \frac{-\delta^\nu{}_\tau + \dfrac{\left(k^\nu k_\tau - V^\nu{}_\tau\right)_{\vee}\left(-k^\alpha k^\beta + V^{\alpha\beta}\right)}{"V_\sigma{}^\sigma\left(V^{\alpha\beta} - k^\alpha k^\beta\right) + k_\sigma k^\beta V^{\alpha\sigma} - V_\sigma{}^\beta V^{\alpha\sigma} + V_\sigma{}^\beta k^\alpha k^\sigma"_{\vee}}}{"-V_\sigma{}^\sigma"} - \frac{-\delta^\nu{}_\tau}{k_\sigma k^\sigma - m^2 + i\varepsilon} \,. \qquad (10.8)$$

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 to be 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 $m^2 = 0$). Referring to (10.4), 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 the linear inverse $I_L{}^\nu{}_\tau$, in the denominator $k_\sigma k^\sigma - m^2 + i\varepsilon$, 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 make the inverse / propagator $\pi_{\nu\tau} = iI_{L\nu\tau}$ non-infinite. 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 (8.12), 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 (10.7), $\left(V_\sigma{}^\sigma\left(V^{\alpha\beta} - k^\alpha k^\beta\right) + k_\sigma k^\beta V^{\alpha\sigma} - V_\sigma{}^\beta V^{\alpha\sigma} + V_\sigma{}^\beta k^\alpha k^\sigma\right)^{-1}$, 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 (5.1) 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 (2.6)):

$$\begin{aligned}\mathfrak{L} &= \left(D_\mu\phi\right)^\dagger\left(D^\mu\phi\right) - m^2 G_\mu G^\mu = \phi\left(\overleftarrow{\partial}_\mu - iG_\mu\right)\left(\partial^\mu - iG^\mu\right)\phi - m^2 G_\mu G^\mu \\ &= \phi\overleftarrow{\partial}_\mu\partial^\mu\phi - i\phi G^\mu\partial_\mu\phi - i\phi\overleftarrow{\partial}_\mu G^\mu\phi - \phi G_\mu G^\mu\phi - m^2 G_\mu G^\mu\end{aligned} \tag{10.9}$$

Above, we use $\left(\left(\partial_\mu - iG_\mu\right)\phi\right)^\dagger = \phi\left(\overleftarrow{\partial}_\mu - iG_\mu\right)$ due to the hermicity of the gauge fields $G_\mu = \lambda^i G^i_\mu$ which is in turn due to $\lambda^i = \lambda^{i\dagger}$ for the Yang-Mills generators, and we also use a left-operating $\overleftarrow{\partial}_\mu$. (While we are here, contrast $\left(D_\mu\phi\right)^\dagger D^\mu\phi$ above to one possible use of the Einstein-Weyl equation (7.6) 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 start with in (10.9) are a scalar ϕ for which we take the gauge-covariant derivative $D^\mu\phi$, we end 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:

$$-g^2\phi G_\mu G^\mu \phi = -g^2(v+h+...)G_\mu G^\mu (v+h+...) = -(vg)^2 G_\mu G^\mu - g^2\left(2vh+h^2+...\right)G_\mu G^\mu \quad (10.10)$$

reveals the term $-(vg)^2 G_\mu G^\mu$. So now (10.9) contains $-(vg)^2 G_\mu G^\mu - m^2 G_\mu G^\mu$. But the term $-m^2 G_\mu G^\mu$ was introduced by hand with a Proca mass and it ruins the gauge symmetry. The term $-(vg)^2 G_\mu G^\mu$, 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 $-(vg)^2 G_\mu G^\mu$ 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 $-(vg)^2 G_\mu G^\mu$ rather than $m^2 G_\mu G^\mu$ as the boson mass term, we keep the gauge theory remains renormalizable. The benefit of having $-m^2 G_\mu G^\mu$ in (10.9) is that it represents an "anticipated" mass against which we compare the emergent $-(vg)^2 G_\mu G^\mu$ to identify the renormalizable mass vg in lieu of the Proca mass *m*.

The exact same sort of thing is happening in (10.8). 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 (10.8) is completely analogous to the term $m^2 G_\mu G^\mu$ in (10.9). Each contains a hand-added, renormalization-destroying, "anticipated" Proca mass. And (10.8) does (10.9) 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 (10.7) before we gauged out this term with $\xi = 1$ in (10.8), everything would blow up. Were we to set the boson on-shell in $I_{L\nu\tau}$ and not use $+i\varepsilon$ added in (10.7), 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. By keeping *m* explicitly in $I_{L\nu\tau}$, we keep the "red flag" of what is "anticipated" so that we can see how this anticipated mass arises from $I_{YM\nu\tau}$, just as when we kept the Proca mass in (10.9). 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 (10.7) 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 (10.7)? 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 with matrix indexes *AB*, (10.8) really says:

$$
\begin{aligned}
&I_P{}^{\nu}{}_{\tau AB} = I_{YM}{}^{\nu}{}_{\tau AB} - I_L{}^{\nu}{}_{\tau}\delta_{AB} = \\
&\left(-\delta^{\nu}{}_{\tau} + \left(k^{\nu}k_{\tau} - V^{\nu}{}_{\tau}\right)\left(V_{\sigma}{}^{\sigma}\left(V^{\alpha\beta} - k^{\alpha}k^{\beta}\right) + k_{\sigma}k^{\beta}V^{\alpha\sigma} - V_{\sigma}{}^{\beta}V^{\alpha\sigma} + V_{\sigma}{}^{\beta}k^{\alpha}k^{\sigma}\right)^{-1}\left(-k^{\alpha}k^{\beta} + V^{\alpha\beta}\right)\right)\left(-V_{\sigma}{}^{\sigma}\right)^{-1}{}_{AB} \\
&-\left(-\delta^{\nu}{}_{\tau}\right)/\left(k_{\sigma}k^{\sigma} - m^2 + i\varepsilon\right)\delta_{AB}
\end{aligned}
\qquad .(10.11)
$$

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 $I_L{}^{\nu}{}_{\tau}$ (which is related to the linear propagator by $\pi_{\nu\tau} = iI_{L\nu\tau}$) into the diagonal positions of the Yang-Mills unit matrix δ_{AB}, thus forming $I_L{}^{\nu}{}_{\tau}\delta_{AB}$.

But (10.11) is in the form of an eigenvalue equation for the matrix $I_{YM}{}^{\nu}{}_{\tau AB}$, with $I_L{}^{\nu}{}_{\tau} = -i\pi^{\nu}{}_{\tau}$ representing the eigenvalues of $I_{YM}{}^{\nu}{}_{\tau AB}$. So if we use this to operate on any Yang-Mills column vector V_B, then $I_L{}^{\nu}{}_{\tau} = -i\pi^{\nu}{}_{\tau}$ will represent the eigenvalues, i.e., the propagator *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 (10.11) 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. For (10.11) this eigenvalue equation takes the form:

$$\boxed{\left|I_P{}^{\nu}{}_{\tau AB}\right| = \left|I_{YM}{}^{\nu}{}_{\tau AB} - I_L{}^{\nu}{}_{\tau}\delta_{AB}\right| = 0} \, . \qquad (10.12)$$

That is it! This is the mass gap solution! Once we deduce a non-zero eigenvalue $I_L{}^{\nu}{}_{\tau} = -i\pi^{\nu}{}_{\tau}$ 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 observable, anticipated mass *m* will be related to this by:

$$\frac{-\delta^{\nu}{}_{\tau}}{k_{\sigma}k^{\sigma} - m^2 + i\varepsilon} = I_L{}^{\nu}{}_{\tau} = -i\pi^{\nu}{}_{\tau} \; . \qquad (10.13)$$

In this way, we may deduce both the mass *m* and, if an eigenvalue $I_L{}^{\nu}{}_{\tau}$ is a *complex number* with an 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 $+i\varepsilon$ corresponds not to the mass – but to a *half-life*. (See, e.g., [23], page 150.)

So, we now turn directly 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 $\mathbb{R}^4$ 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 $N \geq 2$ 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 (10.1) generally, and included in (10.2) in flat spacetime. As pointed out already, the term $D_\sigma D^{[\sigma} D^\alpha D^{\beta]}$ in (10.2) is non-vanishing. To maintain the renormalizability of gauge group G, we *must* set this Proca mass to zero, as we do in (10.6). 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. These results apply to "*any* compact simple gauge group G." For good measure, though not essential, we even place the gauge bosons on-shell as in (10.7).

Now that the gauge bosons are massless, the question becomes how "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 (10.4). For every such perturbation / excitation, $V^{\mu\nu} \neq 0$ and $V_\sigma{}^\sigma \neq 0$, 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 (10.12) will be finite and non-zero and given by $I_L{}^\nu{}_\tau$. These eigenvalues, which are physical observables related to the linear propagator by $I_L{}^\nu{}_\tau = -i\pi^\nu{}_\tau$, may, in the process, also be complex. These eigenvalues in turn, are related to boson masses and lifetimes via (10.13). This means that the "anticipated" mass *m* in (10.13) 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 (10.6).* 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 (10.13). This works for <u>*any*</u> 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 (10.1), 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* which we did maintain as a red flag in $I_L{}^\nu{}_\tau$ in (10.13) is similar to $m^2 G_\mu G^\mu$ which we maintained as a red flag in (10.9). 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 (10.12), this mass which fills the mass gap is revealed in the theory because the excitations in (10.11) give this mass non-zero eigenvalues via (10.13) and the non-zero eigenvalues $I_L{}^{\nu}{}_{\tau}$ are the reciprocals of what then becomes a finite, non-zero, possibly complex, well-behaved $k_\sigma k^\sigma - m^2 + i\varepsilon$, even though the gauge boson masses have been set to zero. If we set $k_\sigma k^\sigma = 0$, 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 (stable particle) for real eigenvalues, and a pure mass number and finite lifetime (unstable particle) 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 *vector bosons* and at the same time have *gauge bosons* which are massless. The mass gap is filled because (10.12) "reveals" a non-zero mass in the inverse (10.13) without ever introducing that mass by hand, in exactly the same way that (10.10) reveals a non-zero mass in the Lagrangian density (10.9) 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 (9.1) have all the chromodynamic color symmetries required of a baryon, and at the same time confine their quarks and its gauge fields, while permitting the flux of colorless quark combinations that we observe in the form of mesons. Given that the mass gap is now filled, this in turn would mean that the nuclear forces associated with these monopole baryons have short range. And, as we shall see, the specific massive particles which emerge in the mass gap solution (10.12), physically, are the mesons observed to be the mediators of strong interactions.

11. Populating Yang-Mills Monopoles with Fermions to Reveal that Yang-Mills Monopoles have the Chromodynamic and Confinement Symmetries of Baryons and Emit and Absorb Objects with the Chromodynamic Symmetries of Mesons

Let us return to the monopole (9.1) 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, we stay in flat spacetime and thereby set all spacetime-covariant derivatives to ordinary derivatives, $\partial_{;\mu} \to \partial_\mu$. And, we keep in mind that $P^{\sigma\mu\nu} = P^{\sigma\mu\nu}{}_{AB}$ is an NxN matrix for SU(N). So, substituting $I_{YM\,\nu\tau} = I_{L\nu\tau} + I_{P\nu\tau}$ into (9.1) yields:

$$\begin{aligned}
P^{\sigma\mu\nu} &= -i\left(\partial^{(\sigma}\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_{YM}^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_{YM}^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&= -i\left(\partial^{(\sigma}\left[\left(I_L^{\alpha\mu}+I_P^{\alpha\mu}\right)\overline{\Psi}\gamma_\alpha\Psi, \left(I_L^{\beta\nu)}+I_P^{\beta\nu)}\right)\overline{\Psi}\gamma_\beta\Psi\right] + \left(I_L^{\tau(\sigma}+I_P^{\tau(\sigma}\right)\overline{\Psi}\gamma_\tau\Psi D^{[\mu}\left(I_L^{\beta\nu])}+I_P^{\beta\nu])}\right)\overline{\Psi}\gamma_\beta\Psi\right) \\
&= -i\left(\partial^{(\sigma}\left[I_L^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_L^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_L^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_L^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&\quad -i\left(\partial^{(\sigma}\left[I_L^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_P^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_L^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_P^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&\quad -i\left(\partial^{(\sigma}\left[I_P^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_L^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_P^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_L^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&\quad -i\left(\partial^{(\sigma}\left[I_P^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_P^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_P^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_P^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \\
&\equiv P_{LL}^{\sigma\mu\nu} + P_{LP}^{\sigma\mu\nu} + P_{PL}^{\sigma\mu\nu} + P_{PP}^{\sigma\mu\nu}
\end{aligned} \quad .(11.1)$$

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:

$$P_{LL}^{\sigma\mu\nu} = -i\left(\partial^{(\sigma}\left[I_L^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_L^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + I_L^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}I_L^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right) \equiv P_{LL1}^{\sigma\mu\nu} + P_{LL2}^{\sigma\mu\nu} . \quad (11.2)$$

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:

$$P_{LL1}^{\sigma\mu\nu} = -i\partial^{(\sigma}\left[I_L^{\alpha\mu}J_\alpha, I_L^{\beta\nu)}J_\beta\right] = -i\partial^{(\sigma}\left[I_L^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_L^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right], \quad (11.3)$$

where we have also used $J_\alpha = \overline{\Psi}\gamma_\alpha\Psi$ to consolidate back to show a source density. Now, let us substitute the linear inverse derived in (8.10) sans $+i\varepsilon$ into the above, to obtain:

$$\begin{aligned}
P_{LL1}^{\sigma\mu\nu} &= -i\partial^{(\sigma}\left[I_L^{\alpha\mu}J_\alpha, I_L^{\beta\nu)}J_\beta\right] = -i\partial^{(\sigma}\left[\frac{-g^{\alpha\mu}+k^\alpha k^\mu/m^2}{k_\sigma k^\sigma - m^2}J_\alpha, \frac{-g^{\beta\nu)}+k^\beta k^{\nu)}/m^2}{k_\sigma k^\sigma - m^2}J_\beta\right] \\
&= i\partial^{(\sigma}\left[\frac{J^\mu}{k_\sigma k^\sigma - m^2}, \frac{J^{\nu)}}{k_\sigma k^\sigma - m^2}\right] = i\partial^{(\sigma}\left[\frac{\overline{\Psi}\gamma^\mu\Psi}{k_\sigma k^\sigma - m^2}, \frac{\overline{\Psi}\gamma^{\nu)}\Psi}{k_\sigma k^\sigma - m^2}\right]
\end{aligned}, \quad (11.4)$$

The terms $k^\alpha k^\mu / m^2$ etc. are eliminated via the conserved current $k^\alpha J_\alpha = 0$, see (8.11), 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 cyclator in the final expression of (11.4) as such:

$$\begin{aligned}P_{LL1}^{\sigma\mu\nu} &= i\left(\partial^{\sigma}\left[\frac{\overline{\Psi}\gamma^{\mu}\Psi}{k_{\sigma}k^{\sigma}-m^{2}},\frac{\overline{\Psi}\gamma^{\nu}\Psi}{k_{\sigma}k^{\sigma}-m^{2}}\right]+\partial^{\mu}\left[\frac{\overline{\Psi}\gamma^{\nu}\Psi}{k_{\sigma}k^{\sigma}-m^{2}},\frac{\overline{\Psi}\gamma^{\sigma}\Psi}{k_{\sigma}k^{\sigma}-m^{2}}\right]+\partial^{\nu}\left[\frac{\overline{\Psi}\gamma^{\sigma}\Psi}{k_{\sigma}k^{\sigma}-m^{2}},\frac{\overline{\Psi}\gamma^{\mu}\Psi}{k_{\sigma}k^{\sigma}-m^{2}}\right]\right)\\ &= i\frac{1}{k_{\sigma}k^{\sigma}-m^{2}}\left(\partial^{\sigma}\frac{\overline{\Psi}\gamma^{[\mu}\Psi\overline{\Psi}\gamma^{\nu]}\Psi}{k_{\sigma}k^{\sigma}-m^{2}}+\partial^{\mu}\frac{\overline{\Psi}\gamma^{[\nu}\Psi\overline{\Psi}\gamma^{\sigma]}\Psi}{k_{\sigma}k^{\sigma}-m^{2}}+\partial^{\nu}\frac{\overline{\Psi}\gamma^{[\sigma}\Psi\overline{\Psi}\gamma^{\mu]}\Psi}{k_{\sigma}k^{\sigma}-m^{2}}\right)\end{aligned},\quad(11.5)$$

where, for example, we compact $\overline{\Psi}\gamma^{[\mu}\Psi\overline{\Psi}\gamma^{\nu]}\Psi=\overline{\Psi}\gamma^{\mu}\Psi\overline{\Psi}\gamma^{\nu}\Psi-\overline{\Psi}\gamma^{\nu}\Psi\overline{\Psi}\gamma^{\mu}\Psi$. Now, let's develop the above in some depth. The development to follow parallels sections 2, 3 and 5 of [13], but streamlines and simplifies that development considerably and, perhaps more importantly, puts that development in the overall context of the complete set of non-linear behaviors which are the hallmark of Yang-Mills gauge theory.

To start, we note the spin sum relationship which is often normalized such that $N^{2}=\mathrm{E}+m$. Here, we shall *not* use this normalization but will use the spin sum *prior to normalization* which is (see, e.g., [23] exercise 5.9):

$$\sum\nolimits_{\text{spins}}U\overline{U}=\frac{N^{2}}{E+m}\left(\not{p}+m\right).\qquad(11.6)$$

Also seeing the emergent $\Psi\overline{\Psi}=U\overline{U}$ in each of the three terms in (11.5), we take the $\Psi\overline{\Psi}=U\overline{U}$ in all three of these terms in (11.5), and then use (11.6) to write:

$$P_{LL1}^{\sigma\mu\nu}=i\frac{1}{k_{\sigma}k^{\sigma}-m^{2}}\frac{N^{2}}{E+m}\left(\partial^{\sigma}\frac{\overline{\Psi}\gamma^{[\mu}\left(\not{p}+m\right)\gamma^{\nu]}\Psi}{k_{\sigma}k^{\sigma}-m^{2}}+\partial^{\mu}\frac{\overline{\Psi}\gamma^{[\nu}\left(\not{p}+m\right)\gamma^{\sigma]}\Psi}{k_{\sigma}k^{\sigma}-m^{2}}+\partial^{\nu}\frac{\overline{\Psi}\gamma^{[\sigma}\left(\not{p}+m\right)\gamma^{\mu]}\Psi}{k_{\sigma}k^{\sigma}-m^{2}}\right)\qquad(11.7)$$

Next, we keep in mind that the fermion propagator

$$\frac{\not{p}+m}{p^{\tau}p_{\tau}-m^{2}}=\frac{\not{p}+m}{\left(\not{p}+m\right)\left(\not{p}-m\right)}=\left(\not{p}-m\right)^{-1},\qquad(11.8)$$

while also noting the appearance of $\left(\not{p}+m\right)/\left(k_{\tau}k^{\tau}-m^{2}\right)$ throughout (11.7) which is very similar in form to the first expression in (11.8). So, if we can find some rationale (see section 3 of [13]) 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 term $P_{LL1}^{\sigma\mu\nu}$. Observing that $1/\left(k_{\tau}k^{\tau}-m^{2}\right)$ 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 (11.7), we shift:

$$\frac{1}{k_{\tau}k^{\tau}-m^{2}}\partial^{\gamma}\frac{\overline{\psi}\gamma^{[\alpha}\left(\not{p}+m\right)\gamma^{\beta]}\psi}{k_{\tau}k^{\tau}-m^{2}}\Rightarrow\frac{1}{k_{\tau}k^{\tau}}\partial^{\gamma}\frac{\overline{\psi}\gamma^{[\alpha}\left(\not{p}+m\right)\gamma^{\beta]}\psi}{p_{\tau}p^{\tau}-m^{2}}.\qquad(11.9)$$

and now take p^τ to represent the fermion four-momentum. It should be clear that both parts of (11.9) 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, (11.7) becomes:

$$P_{LL1}^{\sigma\mu\nu} = i\frac{1}{k_\sigma k^\sigma}\frac{N^2}{E+m}\left(\partial^\sigma \frac{\overline{\Psi}\gamma^{[\mu}\left(\not{p}+m\right)\gamma^{\nu]}\Psi}{p_\sigma p^\sigma - m^2} + \partial^\mu \frac{\overline{\Psi}\gamma^{[\nu}\left(\not{p}+m\right)\gamma^{\sigma]}\Psi}{p_\sigma p^\sigma - m^2} + \partial^\nu \frac{\overline{\Psi}\gamma^{[\sigma}\left(\not{p}+m\right)\gamma^{\mu]}\Psi}{p_\sigma p^\sigma - m^2}\right) \tag{11.10}$$

If we now normalize such that $N^2 = \left(E+m\right)k_\sigma k^\sigma$, then via (11.8) we can reduce (11.10) to:

$$\begin{aligned} P_{LL1}^{\sigma\mu\nu} &= i\left(\partial^\sigma\left(\overline{\Psi}\gamma^{[\mu}\left(\not{p}-m\right)^{-1}\gamma^{\nu]}\Psi\right) + \partial^\mu\left(\overline{\Psi}\gamma^{[\nu}\left(\not{p}-m\right)^{-1}\gamma^{\sigma]}\Psi\right) + \partial^\nu\left(\overline{\Psi}\gamma^{[\sigma}\left(\not{p}-m\right)^{-1}\gamma^{\mu]}\Psi\right)\right) \\ &\equiv i\left(\partial^\sigma\left(\overline{\Psi}_1\gamma^{[\mu}\left(\not{p}-m\right)^{-1}\gamma^{\nu]}\Psi_1\right) + \partial^\mu\left(\overline{\Psi}_2\gamma^{[\nu}\left(\not{p}-m\right)^{-1}\gamma^{\sigma]}\Psi_2\right) + \partial^\nu\left(\overline{\Psi}_3\gamma^{[\sigma}\left(\not{p}-m\right)^{-1}\gamma^{\mu]}\Psi_3\right)\right) \end{aligned} \tag{11.11}$$

which contains three additive terms each containing a propagating fermion wavefunction. But in the bottom line above, we resume the development toward the end of section 9 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 $\Psi = \Psi_A$, with $A = 1...N$ for SU(N). Since each of the three terms in (11.11) represents a fermion propagating within the $P_{LL1}^{\sigma\mu\nu}$ system, in an important step, we designate (define) each term as containing a distinct eigenstate Ψ_1, Ψ_2, Ψ_3 of the SU(N) wavefunction $\Psi = \Psi_A$, $A = 1...N$. Specifically, Dirac-Fermi-Pauli exclusion tells us to make certain that the fermions in each of these three 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 9, 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 in (11.11) as:

$$\Psi_1 \equiv \left|\lambda^8 = \tfrac{1}{\sqrt{3}}; \lambda^3 = 0\right\rangle = \begin{pmatrix}\psi_R\\0\\0\end{pmatrix}; \Psi_2 \equiv \left|\lambda^8 = -\tfrac{1}{2\sqrt{3}}; \lambda^3 = \tfrac{1}{2}\right\rangle = \begin{pmatrix}0\\\psi_G\\0\end{pmatrix}; \Psi_3 \equiv \left|\lambda^8 = -\tfrac{1}{2\sqrt{3}}; \lambda^3 = -\tfrac{1}{2}\right\rangle = \begin{pmatrix}0\\0\\\psi_B\end{pmatrix}. \tag{11.12}$$

This, together with having set *m*=0 in (10.6), means that the $G^\mu = \lambda^i G^{i\mu}$ may now be interpreted not just as generalized gauge bosons, but specifically, as the bi-colored massless gluons of chromodynamics. It also means that we may construct:

$$\Psi_1\overline{\Psi}_1 = \begin{pmatrix} \psi_R\overline{\psi}_R & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad \Psi_2\overline{\Psi}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \psi_G\overline{\psi}_G & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad \Psi_3\overline{\Psi}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \psi_B\overline{\psi}_B \end{pmatrix}. \tag{11.13}$$

We then use (11.13) to display the explicit 3x3 matrix character of $P^{\sigma\mu\nu}_{LL1} = P^{\sigma\mu\nu}_{LL1\ AB}$ of (11.5) with successive Ψ_1, Ψ_2, Ψ_3 assigned as in (11.11) to each of the three terms as such:

$$P^{\sigma\mu\nu}_{LL1\ AB} = i\frac{1}{k_\tau k^\tau - m^2}\begin{pmatrix} \partial^\sigma \dfrac{\overline{\psi}_R\gamma^{[\mu}\psi_R\overline{\psi}_R\gamma^{\nu]}\psi_R}{k_\sigma k^\sigma - m^2} & 0 & 0 \\ 0 & \partial^\mu \dfrac{\overline{\psi}_G\gamma^{[\nu}\psi_G\overline{\psi}_G\gamma^{\sigma]}\psi_G}{k_\sigma k^\sigma - m^2} & 0 \\ 0 & 0 & \partial^\nu \dfrac{\overline{\psi}_B\gamma^{[\sigma}\psi_B\overline{\psi}_B\gamma^{\mu]}\psi_B}{k_\sigma k^\sigma - m^2} \end{pmatrix}. \tag{11.14}$$

Then, repeating the same steps that brought us from (11.5) to (11.11), we may turn this into:

$$P^{\sigma\mu\nu}_{LL1\ AB} = i\begin{pmatrix} \partial^\sigma\left(\overline{\psi}_R\gamma^{[\mu}\left(\not{p}-m\right)^{-1}\gamma^{\nu]}\psi_R\right) & 0 & 0 \\ 0 & \partial^\mu\left(\overline{\psi}_G\gamma^{[\nu}\left(\not{p}-m\right)^{-1}\gamma^{\sigma]}\psi_G\right) & 0 \\ 0 & 0 & \partial^\nu\left(\overline{\psi}_B\gamma^{[\sigma}\left(\not{p}-m\right)^{-1}\gamma^{\mu]}\psi_B\right) \end{pmatrix}. \tag{11.15}$$

The trace equation $\mathrm{Tr}P^{\sigma\mu\nu}_{LL1} = P^{\sigma\mu\nu}_{LL1\ AA}$ is then easily deduced to be:

$$\mathrm{Tr}P^{\sigma\mu\nu}_{LL1} = i\left(\partial^\sigma\left(\overline{\psi}_R\gamma^{[\mu}\left(\not{p}-m\right)^{-1}\gamma^{\nu]}\psi_R\right) + \partial^\mu\left(\overline{\psi}_G\gamma^{[\nu}\left(\not{p}-m\right)^{-1}\gamma^{\sigma]}\psi_G\right) + \partial^\nu\left(\overline{\psi}_B\gamma^{[\sigma}\left(\not{p}-m\right)^{-1}\gamma^{\mu]}\psi_B\right)\right). \tag{11.16}$$

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 [13]. There are of course other terms that we see in (11.1) and (11.2), but we are working with this specific term because it most clearly displays the chromodynamic symmetries of the monopole $P^{\sigma\mu\nu}$. And, although we are working with the one term $\mathrm{Tr}P^{\sigma\mu\nu}_{LL1\ AB}$ out of the eight terms in (11.1), the assignment (11.12) is systemic: with (11.12), every single Ψ in the complete monopole $P^{\sigma\mu\nu}$ of (11.1) 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 (11.16), i.e., $\sigma \sim R$, $\mu \sim G$ and $\nu \sim B$, and keeping in mind that $\mathrm{Tr}P_{LL1}^{\sigma\mu\nu}$ is antisymmetric in all spacetime indexes, we may 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 antisymmetric QCD symmetry required of a baryon.

Furthermore, if we apply Gauss' / Stokes' theorem to (11.16) and also include from (4.3) 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:

$$\begin{aligned} &\iiint \mathrm{Tr}P_{LL1} = \oiint \mathrm{Tr}F_{LL1} = -i\oiint \mathrm{Tr}G^2{}_{LL1} = -3i\oiint \mathrm{Tr}\left[G^{\mu}, G^{\nu}\right]_{LL1} dx_{\mu} dx_{\nu} \\ &= i\oiint \left(\overline{\psi}_R \gamma^{[\mu} \left(\not{p} - m\right)^{-1} \gamma^{\nu]} \psi_R + \overline{\psi}_G \gamma^{[\mu} \left(\not{p} - m\right)^{-1} \gamma^{\nu]} \psi_G + \overline{\psi}_B \gamma^{[\mu} \left(\not{p} - m\right)^{-1} \gamma^{\nu]} \psi_B \right) dx_{\mu} dx_{\nu} \end{aligned} . \quad (11.17)$$

What is the color wavefunction for the $-3i\left[G^{\mu}, G^{\nu}\right]$ entities? By inspection, $\overline{R}R + \overline{G}G + \overline{B}B$. But this is the colorless symmetric wavefunction of a meson! So *quarks do <u>not</u> net flow in and out of closed two-dimensional surfaces surrounding P_{LL1}, except in the colorless $\overline{R}R + \overline{G}G + \overline{B}B$ combination of a meson!* In this way, (11.17) validates the suspicion expressed at the end of section 4 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 with interactions mediated by mesons.

Of course, (11.17) does beg the question of what flows in and out of the complete monopole (11.1), because (11.17) only considers the term P_{LL1}. So if we go back to (11.1) to apply Gauss'/Stokes' theorem, we obtain:

$$i\iiint P = \oiint \left[I_{YM}^{\alpha\mu} \overline{\Psi} \gamma_{\alpha} \Psi, I_{YM}^{\beta\nu} \overline{\Psi} \gamma_{\beta} \Psi \right] dx_{\mu} dx_{\nu} + \iiint I_{YM}^{\tau(\sigma} \overline{\Psi} \gamma_{\tau} \Psi D^{[\mu} I_{YM}^{\beta\nu])} \overline{\Psi} \gamma_{\beta} \Psi dx_{\sigma} dx_{\mu} dx_{\nu} . \quad (11.18)$$

The first term in (11.1), because of the lead $\partial^{(\sigma}$ in (11.1), is fully integrable via Gauss'/Stokes theorem. The second term in (11.1) 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 (11.17), 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 (11.18), 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 4, we now see that for the magnetic monopole (11.1) with surface flux shown in the $\oiint I_{YM}^{\alpha\mu} \overline{\Psi} \gamma_{\alpha} \Psi, I_{YM}^{\beta\nu)} \overline{\Psi} \gamma_{\beta} \Psi dx_{\mu} dx_{\nu}$ term in (11.18), 1) the color wavefunction is that of a baryon, namely $R[G,B] + G[B,R] + B[R,G]$; 2) from (4.4) and (4.5), $\oiint \mathrm{Gluons} = 0$; 3) from (11.17), $\oiint \mathrm{Mesons} \neq 0$ 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 and does not flow across the surface of any theoretical entity proposed to be a baryon, and we see that the Yang-Mills magnetic monopole has *precisely the formal color symmetries and boundary flows required for a baryon.*

Again, 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 (11.16) 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 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 (11.17) 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. Consequently, the physical particle states which the spacetime geometry *does* permit to net flow across closed surfaces are well-suited to represent mesons including the pion. And in the process, QCD itself is fully reproduced. But again, QCD is not a theory of first principle, but rather a corollary theory derived by deduction from Maxwell's electrodynamics as extended into non-Abelian domains by Yang-Mills gauge theory. But in the process, we solve confinement and the mass gap and come to understand symmetric colorless meson flow.

Of course, if we wish to associate these magnetic monopoles with physical baryons, we still need to make them 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 [13], and we need to calculate their energies to see if they make sense in relation to empirical data, see sections 11 and 12 of [13] which shows how the energies calculated from the linear-linear field strength F_{LL1} in $\oiint \mathrm{Tr}F_{LL1}$ in (11.17) appear to track very closely with empirical nuclear binding energies, see also [24]. Insofar as topological stability, we simply note that the trace equation (11.16) 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, *if we assume the <u>simple</u> group SU(3)*, the left and right sides of (11.16) do not match up because one side is traceless and the other is not. It is on this basis that we introduce the product group $SU(3)_C\times U(1)_{B-L}$, and then obtain the monopole (11.16) (and generally, (11.1)) from the spontaneous symmetry breaking of larger SU(4) gauge groups with a $B-L$ (baryon minus lepton number) generator along the lines laid out by Weinberg in [25] at 442 and [12] at 472-473, which in view of [26] Section 12.2.2 and [27] yields the quantum numbers required to turn these monopole baryons into proton and neutrons and ensure that they are topologically stable. These details of all of this are in sections 6 through 8 of [13], and fully apply to the development here with little if any elaboration or modification needed.

12. Chiral Symmetry Breaking

Referring back to Jaffe and Witten at page 3 of [1], in section 10 we showed how Yang-Mills theory leads to a "mass gap" notwithstanding having massless gauge gluons, and in section 11 we demonstrated "quark confinement" of all but the color-neutral meson combinations $\overline{R}R+\overline{G}G+\overline{B}B$. Now let us briefly explore the origins of "chiral symmetry breaking," which is the third leg of the mass gap problem

In (11.17) we identified the mesons which flow in and out of the magnetic monopoles *P*. And in (11.16) we showed how these *P*, by virtue of their $R\left[G,B\right]+G\left[B,R\right]+B\left[R,G\right]$ color wavefunctions and the net flow only of mesons and nothing else, may be interpreted as baryons. Let us now rewrite (11.17) for the meson flow in and out of the monopole baryons with $\overline{C}C\equiv\overline{R}R+\overline{G}G+\overline{B}B$ representing a compacting (*C*=Color, not charge conjugation) of the three additive terms into a single shorthand term, as:

$$\begin{aligned}\iiint \mathrm{Tr}P_{LL1} &= i\oint\!\!\oint\left(\overline{\psi}_C\gamma^{[\mu}\left(\not{p}-m\right)^{-1}\gamma^{\nu]}\psi_C\right)dx_\mu dx_\nu = i\oint\!\!\oint\left(\frac{\overline{\psi}_C\gamma^{[\mu}\left(\not{p}+m\right)\gamma^{\nu]}\psi_C}{p^2-m^2}\right)dx_\mu dx_\nu \\ &= i\oint\!\!\oint\left(p_\alpha\frac{\overline{\psi}_C\gamma^{[\mu}\gamma^\alpha\gamma^{\nu]}\psi_C}{p^2-m^2}\right)dx_\mu dx_\nu + 2\oint\!\!\oint\left(m\frac{\overline{\psi}_C\sigma^{\mu\nu}\psi_C}{p^2-m^2}\right)dx_\mu dx_\nu\end{aligned}\quad . \tag{12.1}$$

This also makes use of $\left(\not{p}-m\right)^{-1}=\left(\not{p}+m\right)/\left(p^2-m^2\right)$. In the second line we separate the two additive terms that emanate from $\not{p}+m$ while applying $\not{p}=p_\alpha\gamma^\alpha$ and expressly introducing the Dirac bilinear $\sigma^{\mu\nu}=\frac{1}{2}i\left[\gamma^\mu,\gamma^\nu\right]$. Now let's look at what these two terms represent.

The latter term for which the core structure is $\oint\!\!\oint\overline{\psi}_C\sigma^{\mu\nu}\psi_C dx_\mu dx_\nu$, contains the second-rank antisymmetric tensor $\overline{\psi}_C\sigma^{\mu\nu}\psi_C$ which, because $\overline{C}C\equiv\overline{R}R+\overline{G}G+\overline{B}B$, is understood to represent a spin-2 vector (V) meson. So this latter term represents the flow of a spin-2 tensor (as opposed to *axial tensor*) meson across the closed monopole / baryon surface, that is, it represents the flow of a $\overline{C}C\equiv\overline{R}R+\overline{G}G+\overline{B}B$ meson with spin 2 and positive parity. In particle parlance, this has $J^P=2^+$, see, e.g., [28] pages 2-4. But what about the other term with the $\gamma^{[\mu}\gamma^\alpha\gamma^{\nu]}$ combination? For this, we expand the main structural term into:

$$p_\alpha\overline{\psi}_C\gamma^{[\mu}\gamma^\alpha\gamma^{\nu]}\psi_C = p_0\overline{\psi}_C\gamma^{[\mu}\gamma^0\gamma^{\nu]}\psi_C + p_1\overline{\psi}_C\gamma^{[\mu}\gamma^1\gamma^{\nu]}\psi_C + p_2\overline{\psi}_C\gamma^{[\mu}\gamma^2\gamma^{\nu]}\psi_C + p_3\overline{\psi}_C\gamma^{[\mu}\gamma^3\gamma^{\nu]}\psi_C\,. \tag{12.2}$$

Then, we evaluate each of the six independent components for $\mu\nu=01,02,03,12,23,31$. The terms where either the μ or ν index is equal to the middle α index drop out because of the μ,ν antisymmetry. Applying the Dirac relation $\gamma^5=i\gamma^0\gamma^1\gamma^2\gamma^3$ in various combinations to the remaining terms while using $g_{\mu\nu}=\eta_{\mu\nu}$ to lower indexes, the result can be covariantlysummarized via the Levi-Civita tensor in a basis where $\varepsilon_{0123}=\sqrt{-g}$ and in flat spacetime where $\varepsilon^{0123}=-1$, by the expression:

$$\overline{\psi}_C\gamma^{[\mu}\not{p}\gamma^{\nu]}\psi_C = p_\alpha\overline{\psi}_C\gamma^{[\mu}\gamma^\alpha\gamma^{\nu]}\psi_C = 2i\varepsilon^{\mu\nu\sigma\tau}p_\sigma\overline{\psi}_C\gamma_\tau\gamma^5\psi_C\,. \tag{12.3}$$

This means that the first term in (12.1) has a core structure $-2\varepsilon^{\mu\nu\sigma\tau}\oiint\left(p_\sigma\overline{\psi}_C\gamma_\tau\gamma^5\psi_C\right)dx_\mu dx_\nu$. Because $\overline{\psi}_C\gamma_\tau\gamma^5\psi_C$ has a single vector index in γ_τ together with a γ^5, this represents a spin-1 *axial vector (A) meson* flowing in and out of the monopole baryon. This is a $J^P=1^+$ meson! So in (11.17) we established that nothing *other than* mesons with $\overline{C}C\equiv\overline{R}R+\overline{G}G+\overline{B}B$ net flow across closed surfaces of the monopole baryons. Now in (12.1) and (12.3) we see that the spin-parity characteristics of the particular mesons in (12.1) are $J^P=2^+$ and $J^P=1^+$. But what about other mesons, such as the pseudoscalar (axial scalar) mesons with $J^P=0^-$ which includes the π mesons which play a central role in strong interactions between nuclei, as well as the whole panoply of mesons catalogued by [28], [29]?

Now we keep in mind that $\iiint\mathrm{Tr}P_{LL1}$ in (11.17) only draws from the $P_{LL}^{\sigma\mu\nu}$ term in (11.1), which is the linear-linear term for flow across a closed monopole / baryon surface. More generally, the meson flow across the surface is given by the term $\oiint\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu}\overline{\Psi}\gamma_\beta\Psi\right]dx_\mu dx_\nu$ in (11.18) which contains all of the non-linear aspects of Yang-Mills theory. But look at what is contained in this term: this tem contains the full inverses $I_{YM}^{\alpha\mu}$ of (8.14), (8.15) which we showed in (9.2) themselves bring in additional gauge bosons / gluons in an infinitely recursive, non-linear fashion via the fact that $D^\mu=\partial^\mu-iG^\mu$. So, if we take the G^μ which enter (8.14), (8.15) via $D^\mu=\partial^\mu-iG^\mu$ and then use $G_\mu\equiv I_{YM\,\tau\mu}J^\tau$ to introduce current densities J^τ and inverses $I_{YM\,\tau\mu}$ as we did in (9.1) and then use these to in turn populate the monopole baryons with fermions via $J^\mu=\overline{\Psi}\gamma^\mu\Psi$ as we also did in (9.1), then in the process, given the infinite recursion, we will now have terms involving $J^N; N=2...\infty$. That is, (9.1) can be recursively expanded to contain $J^\mu=\overline{\Psi}\gamma^\mu\Psi$, multiplied by a like-current density to up to infinite order. We also keep in mind the discussion from (9.4) to (9.8) and note that path integration also is expected to introduce higher powers $J^N; N=2...\infty$ of J^μ. This is what we use Green's functions and Wick contractions to keep track of when we do path integrals.

However, as we saw in (11.6), each time we are able to suitably-commute the $J^\mu=\overline{\Psi}\gamma^\mu\Psi$ in $\oiint\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu}\overline{\Psi}\gamma_\beta\Psi\right]dx_\mu dx_\nu$ of (11.8) to a position where we have two spinors adjacent to one another in the form $\Psi\overline{\Psi}$, we may set $\Psi\overline{\Psi}\to U\overline{U}$ and then use (11.6) to remove those spinors and introduce a $\left(\not{p}-m\right)^{-1}=\left(\not{p}+m\right)/\left(p^2-m^2\right)$ in their place. And we then saw in (12.1) and (12.3) how this yields the spin/parity characteristics of these mesons. But what we learn more generally from (12.1) and (12.3) is that each time we have a current density for which we do this sequence of operations, we are adding Dirac vertexes γ^μ to $\oiint\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu}\overline{\Psi}\gamma_\beta\Psi\right]dx_\mu dx_\nu$, and as $J^N; N=2...\infty$, we will simultaneously be creating $\left(\gamma^\mu\right)^N; N=2...\infty$ combinations of self-multiplied Dirac gammas which emerge following suitable commutation operations and then the application of (11.6).

But, of course, $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, so even though there may be a very large (up to infinite) sequence of γ^μ, we have a closed group consisting of only $\gamma^0,\gamma^1,\gamma^2,\gamma^3,\gamma^5$, and so the terms with up to infinite multiplicative combinations of γ^μ will nonetheless cycle in a closed manner via $\left(\gamma^0\gamma^1\gamma^2\gamma^3\gamma^5\right)^N = \left(-i\right)^N$. So depending on the particular order (power) of *J* of any given term in $\oiint\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu}\overline{\Psi}\gamma_\beta\Psi\right]dx_\mu dx_\nu$, one will find V and A meson terms of the forms $\overline{\psi}_C\psi_C$ (scalar 0^+), $\overline{\psi}_C\gamma^5\psi_C$ (pseudoscalar 0^-), $\overline{\psi}_C\gamma^\mu\psi_C$ (vector 1^-), $\overline{\psi}_C\gamma^\mu\gamma^5\psi_C$ (axial vector 1^+), $\overline{\psi}_C\left[\gamma^\mu,\gamma^\nu\right]\psi_C$ (tensor 2^+), $\overline{\psi}_C\left[\gamma^\mu,\gamma^\nu\right]\gamma^5\psi_C$ (axial tensor 2^-), as well as spin 3 and spin 4 vector and axial mesons which can always be recast as a spin 0, 1 or 2 meson via $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. Any higher powers of the γ^μ will recycle to one of these V or A mesons with spin 0, 1, 2, 3 or 4.

So we now see that because of the infinite recursive nesting of the full inverses $I_{YM}^{\alpha\mu}$ of (8.14), (8.15), and also because path integration results in principle in similar $J^N; N=2...\infty$ powers of current densities, that Yang-Mills theory is accompanied by an infinite $\left(\gamma^\mu\right)^N; N=2...\infty$ range of vertex multiplications which will recycle via $\left(\gamma^0\gamma^1\gamma^2\gamma^3\gamma^5\right)^N = \left(-i\right)^N$, and so via the term $\oiint\left[I_{YM}^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, I_{YM}^{\beta\nu}\overline{\Psi}\gamma_\beta\Psi\right]dx_\mu dx_\nu$ in (11.18), will yield a flux of mesons with the full set of J^P characteristics that are observed in the meson spectrum as catalogued, for example, by [28], [29].

So the third and final leg of the mass gap problem [1], namely the "chiral symmetry breaking" which is "needed to account for the 'current algebra' theory of soft pions that was developed in the 1960s," is accounted for and explained by the presence in (11.18) of terms which contain products $\left(\gamma^\mu\right)^N$, $N=2...\infty$ of Dirac gamma matrices which are then evaluated and reduced with $i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^5 = 1$ to yield the entire observed J^P meson spectrum, in terms of their spin / parity characteristics. (We do not in this paper attempt to explain meson flavors, which is a function of the quark generations u,d; c,s; t,b.)

This brings us full circle back to the discussion at the start of section 3, in which we observed that Yang-Mills theory is rooted in the Hamiltonian quaternions $i^2 = j^2 = k^2 = ijk = -1$ dating back to 1843. The modern representation of Hamilton's quaternions is of course embodied in the 2x2 Pauli spin matrices $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = -i\sigma_1\sigma_2\sigma_3 = I$ developed circa 1925, which are Hermitian, which have the commutation relationship $\left[\sigma_i,\sigma_j\right] = 2i\varepsilon_{ijk}\sigma_k$, and which form the basis for Yang-Mills theory in which $\left[\lambda_i,\lambda_j\right]=if_{ijk}\lambda_k$ with $Tr\left(\lambda^i\lambda^j\right)=\frac{1}{2}\delta^{ij}$. But these quaternions and spin matrices are also embedded in well-known fashion into Dirac's γ^μ defined to reproduce the Minkowski metric tensor $\mathrm{diag}\left(\eta^{\mu\nu}\right)=\left(1,-1,-1,-1\right)$ via $\frac{1}{2}\left(\gamma^\mu\gamma^\nu-\gamma^\nu\gamma^\mu\right)\equiv\eta^{\mu\nu}$. And, of course, $i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^5=1$. So if one wished to represent the Dirac gamma matrices in the form of Hamilton's original quaternions and carve them into a bridge somewhere, one would use the penknife to carve:

$$-\gamma^{0^2}=\gamma^{1^2}=\gamma^{2^2}=\gamma^{3^2}=-\gamma^{5^2}=-i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^5=-1, \tag{12.4}$$

with $-i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^5=-1$ being the spacetime generalization of Hamilton's $ijk=-1$.

So if one desires to take some of the mystery or consternation out of vector/axial and left/right chiral relationships involving γ^5 and the "chiral symmetry breaking" of strong interactions, it is sufficient to note that $i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^5=1$ is simply the Dirac form of Hamilton's quaternions, and that in any theory where one has a product of current densities J^N, $N=2...\infty$ one will likewise have a similar product $\left(\gamma^\mu\right)^N$ of vertices which, via the Dirac quaternion relationship (12.4), will recycle itself and in the process produce particles over an entire spectrum of spin 0, 1, 2, 3 and 4 with both odd and even parity. When this is then understood in the context of (11.17) and (11.18) which describes a flow of color-neutral $\overline{R}R+\overline{G}G+\overline{B}B$ mesons across a closed monopole / baryon surface, and in the context of (8.14) and (8.15) wherein $I_{YM}^{\alpha\mu}$ introduces an infinite order of recursive nesting, it then becomes evident that this stands at the root of "chiral symmetry breaking" and "the 'current algebra' theory of soft pions" which is one of the three main aspects to understanding and solving the mass gap problem.

13. Quantum Yang-Mills Theory

It is worth remarking at this point that in sections 10 through 12, we have been able to solve the mass gap, confinement and chiral symmetry breaking problems entirely on the basis of the *classical* Maxwell field equations extended into non-Abelian gauge fields in the form of (3.1) and (3.2). However, we have relied to some degree on the "recursive" inverses developed and elaborated in sections 8 and 9, which we noted at the time might be useful for Yang-Mills functional path integration. Now, it is time to focus directly on *Quantum* Yang-Mills Theory in order to better understand these solutions in the context of relativistic, nonlinear quantum field theories, see [1] at page 7.

We begin by turning to the third view of Yang-Mills elaborated in (2.5) and (2.6) and used in much of the development since, in which we regard Yang-Mills gauge theory as Maxwell's electrodynamics "on steroids," and specifically, as a theory in which all of the ordinary spacetime derivatives in are replaced with $\partial^\mu\rightarrow D^\mu=\partial^\mu-iG^\mu$ according to a "minimal coupling" principle analogous to that used in gravitational theory to go from $\partial_\mu G^\nu$ in flat spacetime to $\partial_{;\mu}G^\nu \equiv \partial_\mu G^\nu + \Gamma^\nu_{\mu\sigma}G^\sigma$ in gravitationally-curved spacetime. We later saw in sections 6 and 7 how by marrying together both gravitational curvature and gauge curvature, it was possible in (7.7) to derive a classical "Einstein-Weyl" gravitational field equation for Yang-Mills gauge theory.

We now approach Quantum Yang-Mills Theory by posing a very simple question: Does this view of Yang-Mills gauge theory as a minimally-coupled gauge theory on steroids, which clearly applies to the key *classical* equations (2.5), (2.6), (3.1), (3.2), and which even carried forward from the Abelian inverse (8.8), (8.9) to the Yang-Mills inverses (8.14), (8.15), also carry forward even further, into *Quantum* Yang-Mills Theory? That is, might it be possible to simply take the path integral of linear quantum electrodynamics, replace all $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$ throughout wherever ∂^μ appears in a configuration space operator, make the analogous replacement $k^\mu \to \pi^\mu = k^\mu + G^\mu$ of the canonical momentum k^μ with the kinetic momentum π^μ wherever we have performed a $i\partial^\mu \to k^\mu$ transformation into momentum space, and by this simple injection of "steroids," arrive at an analytically-exact expression for the non-linear Yang-Mills path integral? If this does turn out to be possible, then when specifically used for $SU(3)_C$ Chromodynamics as uncovered in section 11, this path integral would become the analytically-exact path integral for Quantum Chromodynamics (QCD).

A priori, without being aware of the recursive view of Yang-Mills theory which we started to develop in section 9, one might be inclined to answer this question in the negative. But as we shall now demonstrate, when one takes account for the recursive view of Yang-Mills, it turns out that the answer to this question is *yes!* The minimal coupling $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$ used to go from Abelian to non-Abelian gauge theory for classical Yang-Mills theory, when supplemented with the analogous minimal coupling $k^\mu \to \pi^\mu = k^\mu + G^\mu$ in momentum space, does carry over from *classical* Yang-Mills theory "to the other side of the river," and works just as well to help us arrive, exactly and analytically, via a recursive kernel, at *Quantum* Yang-Mills Theory in Riemann / Minkowski space $\mathbb{R}^4$. With such a showing, we address a final, key aspect of the mass gap problem [1] at page 6, which is to "prove that for any compact simple gauge group G, *a non-trivial quantum Yang–Mills theory exists* on $\mathbb{R}^4$."

If the steroidal principle of minimal coupling carries $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$ and $k^\mu \to \pi^\mu = k^\mu + G^\mu$ through unscathed from classical to quantum Yang-Mills theory, then this would mean that we may start with the QED path integral (9.4), back up a few steps so as to employ (8.9) in flat spacetime rather than (8.10) in the final term of (9.4), as such:

$$\begin{aligned}Z &= \int DG_\mu \exp i\int d^4x\left(\tfrac{1}{2}G_\mu\left(g^{\mu\nu}\left(\partial_\sigma\partial^\sigma + m^2\right) - \partial^\mu\partial^\nu\right)G_\nu - J^\mu G_\mu\right) \\ &\equiv \mathcal{C}\exp\left(iW\left(J\right)\right) = \mathcal{C}\exp\left(-\frac{1}{2}i\int\frac{d^4k}{\left(2\pi\right)^4}J^\mu\frac{g_{\nu\tau} + \dfrac{\partial_\nu\partial_\tau\partial^\alpha\partial^\beta}{\left(\partial_\sigma\partial^\sigma + m^2\right)\partial^\alpha\partial^\beta - \partial_\sigma\partial^\beta\partial^\alpha\partial^\sigma}}{\partial_\sigma\partial^\sigma + m^2}J^\nu\right).\end{aligned} \tag{13.1}$$

Then we may substitute $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$ into the configuration space operator while simultaneously substituting $\partial^\mu \to D^\mu = \left(\partial^\mu - iG^\mu\right) \to -i\pi^\mu = -i\left(k^\mu + G^\mu\right)$ into the $W(J)$, while also using the placement markers and quoted denominators of (8.15), and also including a Trace (Tr) and multiplying through by 2 (see e.g., (2.6)) which is required when using Yang-Mills matrices with the normalization $Tr\left(\lambda^i\lambda^j\right) = \frac{1}{2}\delta^{ij}$ in a Lagrangian. All of this yields:

$$
\begin{aligned}
Z &= \int DG_\mu \exp i \int d^4x \mathrm{Tr}\left(G_\mu\left(g^{\mu\nu}\left(D_\sigma D^\sigma + m^2\right) - D^\mu D^\nu\right)G_\nu - 2J^\mu G_\mu\right) \equiv \mathcal{C}\exp\left(iW(J)\right) \\
&= \mathcal{C}\exp\left(-i\int \frac{d^4k}{(2\pi)^4} \mathrm{Tr}\left(J^\mu \frac{-g_{\mu\nu} + \dfrac{\pi_\mu \pi_{\nu\vee} \pi^\alpha \pi^\beta}{"m^2\pi^\alpha\pi^\beta - \pi_\sigma\pi^\sigma\pi^\alpha\pi^\beta + \pi_\sigma\pi^\beta\pi^\alpha\pi^\sigma"_\vee}}{"\pi_\sigma\pi^\sigma - m^2"} J^\nu\right)\right)
\end{aligned} \quad .(13.2)
$$

Then, in accordance with the mass gap solution of section 10 in which we set the Proca mass $m \to 0$ and the uncovering of SU(3)$_C$ in section 11, we set $m \to 0$ and regard the gauge group of (13.2) to be SU(3)$_C$ and so write (13.2) *specifically for QCD* as:

$$
\begin{aligned}
Z &= \int DG_\mu \exp i \int d^4x \mathrm{Tr}\left(G_\mu\left(g^{\mu\nu} D_\sigma D^\sigma - D^\mu D^\nu\right)G_\nu - 2J^\mu G_\mu\right) \\
&\equiv \mathcal{C}\exp\left(iW(J)\right) = \mathcal{C}\exp\left(-i\int \frac{d^4k}{(2\pi)^4} \mathrm{Tr}\left(J^\mu \frac{-g_{\mu\nu} + \dfrac{\pi_\mu \pi_{\nu\vee} \pi^\alpha \pi^\beta}{"\pi_\sigma \pi^{[\beta}\pi^\alpha\pi^{\sigma]}"_\vee}}{"\pi_\sigma\pi^\sigma"} J^\nu\right)\right).
\end{aligned} \quad (13.3)
$$

So now the question is simply this: is (13.2) in fact a mathematically correct result, i.e., is the Gaussian integral properly formulated and then evaluated? We now prove that (13.2) is correct, and specifies the *exact analytical form of the Quantum Yang-Mills path integral using a recursive kernel*, in a manner that we stated after (9.2) might be possible. To prove that (13.2) is correct requires three steps: 1) obtaining the product rule for the classical Yang-Mills Lagrangian density $\mathfrak{L}$; 2) obtaining the classical Yang-Mills action $S = \int \mathfrak{L} d^4x$; and 3) showing that (13.2) correctly evaluates $Z = \int D\phi \exp iS$ for the Yang-Mills gauge field $\phi = G_\mu$, which will rely upon the recursive view developed in section 9. We work in flat spacetime.

First, as to the product rule, for any product *ab* of *a*, *b* operated on by the gauge-covariant derivative $D^\mu = \partial^\mu - iG^\mu$, we may write:

$$
D^\mu(ab) = \left(\partial^\mu - iG^\mu\right)(ab) = \partial^\mu ab + a\partial^\mu b - iG^\mu ab \ . \quad (13.4)
$$

The extra term $-iG^\mu ab$ is wholly a creature of the gauge-covariant derivative, and does not exist for an ordinary derivative. So with the assignments $a = G^\nu$, $b = D_{[\mu}G_{\nu]}$, (13.4) becomes:

$$D^{\mu}(G^{\nu}D_{[\mu}G_{\nu]})=\partial^{\mu}G^{\nu}D_{[\mu}G_{\nu]}+G^{\nu}\partial^{\mu}D_{[\mu}G_{\nu]}-iG^{\mu}G^{\nu}D_{[\mu}G_{\nu]}=D^{\mu}G^{\nu}D_{[\mu}G_{\nu]}+G^{\nu}\partial^{\mu}D_{[\mu}G_{\nu]}. \quad (13.5)$$

Noting that the Lagrangian density (2.6) for a pure Yang-Mills gauge field contains a term $-\frac{1}{2}D^{[\mu}G^{\nu]}D_{[\mu}G_{\nu]}=-D^{\mu}G^{\nu}D_{[\mu}G_{\nu]}$, we now restructure (13.5) in terms of $D^{\mu}G^{\nu}D_{[\mu}G_{\nu]}$. The full calculation is instructive, with index gymnastics starting on the fifth line:

$$\begin{aligned}
D^{\mu}G^{\nu}D_{[\mu}G_{\nu]} &= D^{\mu}(G^{\nu}D_{[\mu}G_{\nu]})-G^{\nu}\partial^{\mu}D_{[\mu}G_{\nu]} \\
&=\left(\partial^{\mu}-iG^{\mu}\right)(G^{\nu}D_{[\mu}G_{\nu]})-G^{\nu}\partial^{\mu}D_{[\mu}G_{\nu]} \\
&=\partial^{\mu}\left(G^{\nu}D_{[\mu}G_{\nu]}\right)-iG^{\mu}G^{\nu}D_{[\mu}G_{\nu]}-G^{\nu}\partial^{\mu}D_{[\mu}G_{\nu]} \\
&=\partial^{\mu}\left(G^{\nu}D_{[\mu}G_{\nu]}\right)-iG^{\mu}G^{\nu}D_{\mu}G_{\nu}+iG^{\mu}G^{\nu}D_{\nu}G_{\mu}-G^{\nu}\partial^{\mu}D_{\mu}G_{\nu}+G^{\nu}\partial^{\mu}D_{\nu}G_{\mu}\ . \\
&=\partial^{\mu}\left(G^{\nu}D_{[\mu}G_{\nu]}\right)+\left(-iG^{\sigma}G^{\nu}D_{\sigma}+iG^{\nu}G^{\sigma}D_{\sigma}-G^{\nu}\partial^{\sigma}D_{\sigma}+G^{\sigma}\partial^{\nu}D_{\sigma}\right)G_{\nu} \\
&=\partial^{\mu}\left(G^{\nu}D_{[\mu}G_{\nu]}\right)+\left(G^{\sigma}D^{\nu}D_{\sigma}-G^{\nu}D^{\sigma}D_{\sigma}\right)G_{\nu} \\
&=\partial^{\mu}\left(G^{\nu}D_{[\mu}G_{\nu]}\right)-G_{\mu}\left(g^{\mu\nu}D_{\sigma}D^{\sigma}-D^{\nu}D^{\mu}\right)G_{\nu}
\end{aligned} \quad (13.6)$$

We see in the final line, the emergence of the Yang-Mills configuration space operator sans Proca mass, $g^{\mu\nu}D_{\sigma}D^{\sigma}-D^{\nu}D^{\mu}$, contrast (3.3). This is the minimally-coupled, steroidal Yang-Mills configuration space operator. The only place in which the minimal coupling does *not* carry through, is in the term $\partial^{\mu}\left(G^{\nu}D_{[\mu}G_{\nu]}\right)$. But as we shall shortly see, this is exactly what we need in order to eliminate this term with a boundary condition when calculating the action.

Working from (2.6) and applying (13.6), let us now form the Yang-Mills Lagrangian density including a current source $J^{\mu}G_{\mu}$, to obtain:

$$\begin{aligned}
\mathfrak{L} &=\mathrm{Tr}\left(-\tfrac{1}{2}F^{\mu\nu}F_{\mu\nu}-2J^{\mu}G_{\mu}\right)=\mathrm{Tr}\left(-\tfrac{1}{2}D^{[\mu}G^{\nu]}D_{[\mu}G_{\nu]}-2J^{\mu}G_{\mu}\right)=\mathrm{Tr}\left(-D^{\mu}G^{\nu}D_{[\mu}G_{\nu]}-2J^{\mu}G_{\mu}\right) \\
&=\mathrm{Tr}\left(-\partial^{\mu}\left(G^{\nu}D_{[\mu}G_{\nu]}\right)+G_{\mu}\left(g^{\mu\nu}D_{\sigma}D^{\sigma}-D^{\nu}D^{\mu}\right)G_{\nu}-2J^{\mu}G_{\mu}\right)
\end{aligned} .(13.7)$$

This is the classical Yang-Mills Lagrangian density arrived at via the product rule (13.6).

<u>Second, as to the classical action</u> $S=\int\mathfrak{L}d^4x$, we use (13.7), and add back in the Proca mass just for the moment, to write:

$$S=\int\mathfrak{L}d^4x=\int d^4x\mathrm{Tr}\left(-\partial^{\mu}\left(G^{\nu}D_{[\mu}G_{\nu]}\right)+G_{\mu}\left(g^{\mu\nu}\left(D_{\sigma}D^{\sigma}+m^2\right)-D^{\nu}D^{\mu}\right)G_{\nu}-2J^{\mu}G_{\mu}\right). \quad (13.8)$$

Except for the additional term $\partial^\mu\left(G^\nu D_{[\mu}G_{\nu]}\right)$ noted above, this is *identical* to the action term appearing in the path integral $Z=\int D\phi \exp iS$ of (13.2). And, except for this same term $\partial^\mu\left(G^\nu D_{[\mu}G_{\nu]}\right)$, the steroidal minimal coupling $\partial^\mu \to D^\mu = \partial^\mu - iG^\mu$ has indeed carried through to the classical Yang-Mills action. But this extra term can be eliminated by boundary conditions in the usual way, and the reason we can do so, is precisely because this additional term is *not* steroidal, but simply contains an ordinary derivative ∂^μ rather than the gauge-covariant D^μ. Specifically, by imposing $G^\nu\left(x^\mu=\infty\right)=G^\nu\left(x^\mu=-\infty\right)=0$ (or even the looser condition $G^\nu\left(\infty\right)=G^\nu\left(-\infty\right)$) as a boundary condition upon the gauge potential, so that $G^\nu\Big|_{A^\nu\left(x^\mu=-\infty\right)}^{A^\nu\left(x^\mu=+\infty\right)}=0$ for each of the coordinates $x^\mu=\left(t,x,y,z\right)$, and with $d^4x=dx^0dx^1dx^2dx^3$, we may calculate that:

$$
\begin{aligned}
\int d^4x\partial^\mu (G^\nu D_{[\mu}G_{\nu]}) &= \int d^4x g^{\mu\sigma}\frac{\partial}{\partial x^\sigma}(G^\nu D_{[\mu}G_{\nu]}) \\
&= \int dx^0dx^1dx^2dx^3 g^{\mu 0}\frac{\partial}{\partial x^0}(G^\nu D_{[\mu}G_{\nu]}) + \int dx^0dx^1dx^2dx^3 g^{\mu 1}\frac{\partial}{\partial x^1}(G^\nu D_{[\mu}G_{\nu]}) \\
&\quad + \int dx^0dx^1dx^2dx^3 g^{\mu 2}\frac{\partial}{\partial x^2}(G^\nu D_{[\mu}G_{\nu]}) + \int dx^0dx^1dx^2dx^3 g^{\mu 3}\frac{\partial}{\partial x^3}(G^\nu D_{[\mu}G_{\nu]}) \\
&= \int dx^1dx^2dx^3 g^{\mu 0}\left(G^\nu\Big|_{G^\nu(t=-\infty)}^{G^\nu(t=+\infty)}\right)D_{[\mu}G_{\nu]} + \int dx^0dx^2dx^3 g^{\mu 1}\left(G^\nu\Big|_{G^\nu(x=-\infty)}^{G^\nu(x=+\infty)}\right)D_{[\mu}G_{\nu]} \\
&\quad + \int dx^0dx^1dx^3 g^{\mu 2}\left(G^\nu\Big|_{G^\nu(y=-\infty)}^{G^\nu(y=+\infty)}\right)D_{[\mu}G_{\nu]} + \int dx^0dx^1dx^2 g^{\mu 3}\left(G^\nu\Big|_{G^\nu(z=-\infty)}^{G^\nu(z=+\infty)}\right)D_{[\mu}G_{\nu]} \\
&= 0
\end{aligned}
, \qquad (13.9)
$$

So with $\int d^4x\partial^\mu (G^\nu D_{[\mu}G_{\nu]})=0$, the classical Yang-Mills action (13.8) now reduces to:

$$S=\int \mathcal{L}d^4x=\int d^4x\mathrm{Tr}\left(G_\mu\left(g^{\mu\nu}\left(D_\sigma D^\sigma+m^2\right)-D^\nu D^\mu\right)G_\nu-2J^\mu G_\mu\right). \qquad (13.10)$$

This is an important result, because it tells us that the action we have employed in the configuration space portion of the path integral (13.2) is the correct action. Specifically, given that the electrodynamic action is $S=\int d^4x\mathrm{Tr}\left(\frac{1}{2}G_\mu\left(g^{\mu\nu}\left(\partial_\sigma\partial^\sigma+m^2\right)-\partial^\nu\partial^\mu\right)G_\nu-J^\mu G_\mu\right)$, we see that the steroidal minimal coupling first elaborated in (2.5) and (2.6) does carry through all the way into the classical Yang-Mills action that feeds the path integral (13.2). Thus, the path integral (13.2) is properly formulated. Now that we know that the configuration space portion of (13.2) is correct, we now need simply to prove that the Gaussian integration of this expression is correct.

Third, as to the evaluation of the Gaussian integral, let us expand each of the $D^\mu=\partial^\mu-iG^\mu$ in (13.10) to explicitly show the gauge fields in the form:

$$\begin{aligned} S=\int \mathfrak{L} d^4x = \int d^4x \mathrm{Tr}\Big(G_\mu\Big(g^{\mu\nu}\big(\big(\partial_\sigma - iG_\sigma\big)\big(\partial^\sigma - iG^\sigma\big)+m^2\big)-\big(\partial^\nu - iG^\nu\big)\big(\partial^\mu - iG^\mu\big)\Big)G_\nu - 2J^\mu G_\mu\Big) \\ = \int d^4x \mathrm{Tr}\left(G_\mu \begin{pmatrix} g^{\mu\nu}\big(\big(\partial_\sigma\partial^\sigma - i\partial_\sigma G^\sigma - iG_\sigma\partial^\sigma - G_\sigma G^\sigma\big)+m^2\big) \\ -\big(\partial^\nu\partial^\mu - i\partial^\nu G^\mu - iG^\nu\partial^\mu - G^\nu G^\mu\big) \end{pmatrix} G_\nu - 2J^\mu G_\mu \right) \end{aligned} . \quad (13.11)$$

If we now try to use this in $Z=\int DG_\mu \exp iS$ without being aware of the recursive nature of G_σ, then we run into the brick wall that has thus far made it impossible to analytically solve the path integral using an action such as (3.11). After all, the ability to exactly solve the QED path integral is based on the Gaussian integral $\int dx \exp\left(-\frac{1}{2}Ax^2 - Jx\right) = \left(2\pi / A\right)^{.5}\exp\left(J^2/2A\right)$ from (9.8) in which the variable of integration *x*, which abstracts to the gauge field G_μ in (13.11), appears only to quadratic order, and specifically, appears as x^2 and *x*. But in (13.11), we have a polynomial in G_μ up to an abstracted x^4. Ordinarily, one turns to (9.5) to try to solve this using the variation $V\left(\delta/\delta J\right)$, and specifically, uses the fact that $G_\mu = \delta\left(J^\mu G_\mu\right)/\delta J^\mu$ to replace all occurrence of G_μ which are of higher than second order with $G_\mu \to \delta/\delta J^\mu$ and then segregate those terms from the integrand, which thereby allows the integral $\int dx \exp\left(-\frac{1}{2}Ax^2 - Jx\right)$ to be taken. Then the various $\delta/\delta J$ are used to operate on the $\exp\left(J^2/2A\right)$ that emerges following the integration. This, however, is exceptionally difficult to do in exact, closed form.

But we are now aware from (9.2) for $G_\tau = I_{YM\,\nu\tau} J^\nu$ that:

$$G_\tau = \left[g_{\nu\tau} + D_{;\nu}D_{;\tau}\left(m^2 D^{;\alpha}D^{;\beta} + D_{;\sigma}D^{;\sigma}D^{;\alpha}D^{;\beta} - D_{;\sigma}D^{;\beta}D^{;\alpha}D^{;\sigma}\right)^{-1} D^{;\alpha}D^{;\beta}\right]\left(D_{;\sigma}D^{;\sigma} + m^2\right)^{-1} J^\nu . \quad (13.12)$$

We are also now aware from the development in section 9, that each and every G^μ in $g^{\mu\nu}\left(\left(\partial_\sigma\partial^\sigma - i\partial_\sigma G^\sigma - iG_\sigma\partial^\sigma - G_\sigma G^\sigma\right)+m^2\right) - \left(\partial^\nu\partial^\mu - i\partial^\nu G^\mu - iG^\nu\partial^\mu - G^\nu G^\mu\right)$ which is the configuration space operator in (13.11) can be replaced using (13.12) with a source current via $G_\tau = I_{\nu\tau}J^\nu$, and that each D^μ in $I_{YM\,\nu\tau}$ can again be expanded using $D^\mu = \partial^\mu - iG^\mu$ and that (13.12) can be reapplied again, and that this process can be repeated, iteratively, recursively, ad infinitum. So in the limit of infinite recursion, $\lim_{N\to\infty}\left(\left(\;\;\right)\right)$ using the nesting notation developed in section 9, the action (13.11) becomes:

$$S=\int \mathfrak{L} d^4x = \int d^4x \mathrm{Tr}\left(G_\mu \lim_{N\to\infty}\left(\left(\begin{matrix} g^{\mu\nu}\left(\left(\partial_\sigma\partial^\sigma - i\partial_\sigma G^\sigma - iG_\sigma\partial^\sigma - G_\sigma G^\sigma\right)+m^2\right) \\ -\left(\partial^\nu\partial^\mu - i\partial^\nu G^\mu - iG^\nu\partial^\mu - G^\nu G^\mu\right) \end{matrix} \right)\right) G_\nu - 2J^\mu G_\mu \right) . \quad (13.13)$$

That is, in (13.13), every occurrence of G^μ is regarded via $\lim_{N\to\infty}\left(\left(\ \right)\right)$ to have been replaced with a current density J^ν *ad infinitum* using (13.12), *and so this configuration space operator is no longer a function of* G^μ *but rather is a function of* J^μ. This is exactly the same thing that we do with the functional variation $G_\mu \to \delta/\delta J^\mu$ to remove all terms which are polynomial (greater than second order) in the gauge field, but we use an infinite recursion instead! As a result, when we now seek to take $Z=\int DG_\mu \exp iS$, the expression $A\Leftrightarrow \lim_{N\to\infty}\left(\left(\ \right)\right)$ in (13.13) corresponds to the A in $\int dx \exp\left(-\frac{1}{2}Ax^2-Jx\right)=\left(2\pi/A\right)^{.5}\exp\left(J^2/2A\right)$, and the overall expression (13.13) contains *only terms quadratic in* G_μ because all of the higher order terms in the gauge field have been turned into functions of J^μ using the infinite recursion in lieu of the variation $G_\mu \to \delta/\delta J^\mu$. So to do the Gaussian integral, all we now need is the inverse 1/*A*. But going back to (13.10), $A\Leftrightarrow g^{\mu\nu}\left(D_\sigma D^\sigma+m^2\right)-D^\nu D^\mu$ is just the configuration space operator that we have been using ever since (3.2), and we know its inverse from (8.14) and (8.15). So if we just define the double nest symbol ${}_\infty\left(\left(\ \right)\right)$ to denote the infinite nesting $\lim_{N\to\infty}\left(\left(\ \right)\right)$, then we can use the action (13.10) to in fact obtain an *exact* expression for the path integral, namely:

$$\begin{aligned}Z&=\int DG_\mu \exp i\int d^4x\,\mathrm{Tr}\left(G_\mu\, {}_\infty\left(\left(g^{\mu\nu}\left(D_\sigma D^\sigma+m^2\right)-D^\mu D^\nu\right)\right)G_\nu-2J^\mu G_\mu\right)\equiv \mathcal{C}\exp\left(iW\left(J\right)\right)\\&=\mathcal{C}\exp\left(-i\int\frac{d^4k}{\left(2\pi\right)^4}\mathrm{Tr}\left(J^\mu\, {}_\infty\left(\left(\frac{-g_{\mu\nu}+\dfrac{\pi_\mu\pi_{\nu\vee}\pi^\alpha\pi^\beta}{"m^2\pi^\alpha\pi^\beta-\pi_\sigma\pi^\sigma\pi^\alpha\pi^\beta+\pi_\sigma\pi^\beta\pi^\alpha\pi^\sigma"\vee}}{"\pi_\sigma\pi^\sigma-m^2"}\right)\right)J^\nu\right)\right)\end{aligned} \qquad .(13.14)$$

This is identical to the (13.2) which was arrived at by applying the steroidal minimal coupling to the QED path integral, but for the ${}_\infty\left(\left(\ \right)\right)$ nest to indicate an infinitely-iterative recursive application of (13.12) to all appearances of G^μ occurring inside the nest. It is important that this ${}_\infty\left(\left(\ \right)\right)$ nest also appears above in what we now segregate into the Yang-Mills amplitude:

$$W\left(J\right)=-\int\frac{d^4k}{\left(2\pi\right)^4}\mathrm{Tr}\left(J^\mu\, {}_\infty\left(\left(\frac{-g_{\mu\nu}+\dfrac{\pi_\mu\pi_{\nu\vee}\pi^\alpha\pi^\beta}{"m^2\pi^\alpha\pi^\beta-\pi_\sigma\pi^\sigma\pi^\alpha\pi^\beta+\pi_\sigma\pi^\beta\pi^\alpha\pi^\sigma"\vee}}{"\pi_\sigma\pi^\sigma-m^2"}\right)\right)J^\nu\right) \qquad (13.15)$$

because $\pi^\mu=k^\mu+G^\mu$. Thus, here too, so long as we infinitely nest the G^μ using (13.12), we will have an amplitude that is exclusively a function of J^ν, from second order all the way to infinite order, while G^μ, which is both the variable of path integration and the dummy variable of recursion (contrast (9.3)), has been entirely removed in the process.

So we now establish that the steroidal minimal coupling of Yang-Mill theory, first articulated in (2.5) and (2.6) not only applies to classical field equations and the classical inverse and the classical action, but also, that it survives path integration, and applies as well, in the form $k^\mu \to \pi^\mu = k^\mu - {}_\infty\left(\left(G^\mu\right)\right)$, with an infinitely-recursively expanded G^μ, to the amplitude functions of Quantum Yang Mills Theory. Put plainly, classical and quantum Yang-Mills theory are simply classical and quantum electrodynamics on minimally-coupled steroids, with an infinitely-recursive expansion of the gauge fields.

For QCD, the amplitude is based on (13.3) for massless gluons, and is simply (13.15) with $m=0$ and the gauge group regarded to be SU(3)$_C$, namely:

$$W(J) = -\int \frac{d^4k}{(2\pi)^4}\mathrm{Tr}\left(J^\mu \,{}_\infty\left(\left(\frac{-g_{\mu\nu} + \dfrac{\pi_\mu \pi_{\nu\vee}\pi^\alpha\pi^\beta}{"\pi_\sigma \pi^{[\beta}\pi^\alpha\pi^{\sigma]}"\,\vee}}{"\pi_\sigma\pi^\sigma"}\right)\right) J^\nu \right). \tag{13.16}$$

Above, $\pi^\mu = k^\mu + G^\mu = k^\mu + \lambda^i G^{i\mu}$, $i = 1...8$, and λ^i are the generators of SU(3)$_C$. *This is the exact analytical solution to the QCD path integral, specified using an exact recursive kernel in closed form.*

Having derived (13.15) generally for *any* Quantum Yang-Mills Theory and (13.16) specifically for QCD, let us briefly talk about the practical aspects of calculating with (13.15) and (13.16). Ideally, one would wish to use the approach developed in section 9 and specifically the approach illustrated by the example (9.3) and (9.7) for e^B, to obtain a closed, *non-recursive* analytical expression for what we now define as a *propagator operator*:

$${}_\infty\Pi_{\mu\nu} \equiv i\,{}_\infty\left(\left(\left(-g_{\mu\nu} + \pi_\mu\pi_\nu\left(\pi_\sigma\pi^{[\beta}\pi^\alpha\pi^{\sigma]}\right)^{-1}\pi^\alpha\pi^\beta\right)\left(\pi_\sigma\pi^\sigma\right)^{-1}\right)\right) = i\,{}_\infty\left(\left(\frac{-g_{\mu\nu} + \dfrac{\pi_\mu \pi_{\nu\vee}\pi^\alpha\pi^\beta}{"\pi_\sigma \pi^{[\beta}\pi^\alpha\pi^{\sigma]}"\,\vee}}{"\pi_\sigma\pi^\sigma"}\right)\right), \tag{13.17}$$

in (13.16), and its more general counterpart which may be similarly defined from (13.15). If a closed analytical expression for the above can be developed analogously to the example (9.3), (9.7) so as to strip off the recursion, then this operator would become closed and exact, rather than just being an exact recursive kernel. But we leave this mathematical problem for another day. We have in the above shown this operator fully specified in terms of the required matrix inversions and multiplications, as well as in its "user friendly" form based on (8.15) with place markers and quoted denominators. By writing ${}_\infty\Pi_{\mu\nu}$ we the infinity prescript, we denote that all of the kinetic momenta are to be infinitely recursively expanded via $\pi^\mu = k^\mu - {}_\infty\left(\left(G^\mu\right)\right)$.

Absent an exact mathematical evaluation of (13.17) (or the more general analogue from (13.15)) which sums out the recursion into a closed form, one can do *numerical* computations by performing the recursion up to a specified *finite* level of nesting, while recognizing that this then becomes approximate up to the levels of nesting that are left out, in the same way that "loop" calculations of less than an infinite number of loops are also approximate not exact up to all of the loops that are left out. Symbolically, we use ${}_N\left(\left(\ \right)\right)$ to designate a recursion which is *numerically* applied for *N* recursive iterations, in contrast to ${}_\infty\left(\left(\ \right)\right)$ for an infinite succession of iterations which in the absence of an infinite computing resource is necessarily an *analytical* calculation. So when (13.15) through (13.17) are approached numerically rather than analytically, one needs to do a finite recursive calculation, rather than an infinite one. But for *finite* recursions, one generally needs two inputs: first, a recursive kernel; second, a terminal condition. A good example is the recursive definition of the factorial function: The recursive kernel says that $n! = n \times (n-1)!$. The terminal condition says that $0! = 1$.

So working off of (13.17), the recursive kernel for a *finite* nesting is:

$$ {}_N\Pi_{\mu\nu} \equiv i\ {}_N\left(\left(\left(-g_{\mu\nu} + \pi_\mu \pi_\nu \left(\pi_\sigma \pi^{[\beta} \pi^\alpha \pi^{\sigma]}\right)^{-1} \pi^\alpha \pi^\beta\right)\left(\pi_\sigma \pi^\sigma\right)^{-1}\right)\right). \tag{13.18} $$

But, what is the terminal condition which we denote as ${}_0\Pi_{\mu\nu}$? That would simply be the linear propagator $\pi_{\mu\nu} = iI_{L\mu\nu}$ of Abelian gauge theory which was developed in (8.9) and (8.10), which includes the Proca mass and $+i\varepsilon$, and which is give by:

$$ {}_0\Pi_{\mu\nu} \equiv \pi_{\mu\nu} = i\frac{-g_{\mu\nu} + \dfrac{k_\mu k_\nu}{m^2}}{k_\sigma k^\sigma - m^2 + i\varepsilon} = iI_{L\mu\nu}. \tag{13.19} $$

In other words, when doing a *finite* recursive calculation numerically up to *N* levels of nesting, one substitutes (13.12) for the gauge fields through *N* iterations, and then on the next iteration, one terminates by using (13.19) as the terminal condition analogous to $0! = 1$, which is just (13.18) in which $\pi^\mu \to k^\mu$ and in which the Proca mass *m* and $+i\varepsilon$ are included.

Going back to the mass gap solution (10.12), we now recognize that the Yang-Mills inverse $I_{YM\,\mu\nu}$ is related to the Yang-Mills propagator matrix *operator* $\Pi_{\mu\nu} = \Pi_{\mu\nu AB}$ according to $\Pi_{\mu\nu} = iI_{YM\,\mu\nu}$ while at the same time the linear inverse and the Abelian (e.g. QED) propagator are related by $\pi_{\mu\nu} = iI_{L\mu\nu}$. We may use this insight to rewrite the mass gap solution (10.12) in terms of these quantum propagators as:

$$ \boxed{\left|\Pi_{\mu\nu AB} - \pi_{\mu\nu}\delta_{AB}\right| = 0} \tag{13.20} $$

with $\pi_{\mu\nu}$ defined as in (13.19). In this light, the mass gap solution has a very simple interpretation: A propagator $\pi_{\mu\nu}$ is simply a second-rank spacetime tensor eigenvalue of a propagator operator matrix $\Pi_{\mu\nu AB}$. In this regard, it is interesting to note that for a numerical calculation, the eigenvalues ${}_{(N)}\pi_{\mu\nu}$ of the operators ${}_N\Pi_{\mu\nu}$ will be different for different nesting levels. In other words, because ${}_{N=n}\Pi_{\mu\nu} \neq {}_{N=m}\Pi_{\mu\nu}$ for different nesting levels $n \neq m$, one will also have different eigenvalues ${}_{(n)}\pi_{\mu\nu} \neq {}_{(m)}\pi_{\mu\nu}$ for these different nesting levels.

The foregoing also allows us to go back to (11.1) and explicitly include the propagator operator $\Pi_{\mu\nu} = iI_{YM\,\mu\nu}$ in the expression (11.1) for the complete Yang-Mills monopole baryon, via the substitution $\Pi_{\mu\nu} = iI_{YM\,\mu\nu}$. Thus, we now write the monopole baryon as:

$$P^{\sigma\mu\nu} = i\left(\partial^{(\sigma}\left[\Pi^{\alpha\mu}\overline{\Psi}\gamma_\alpha\Psi, \Pi^{\beta\nu)}\overline{\Psi}\gamma_\beta\Psi\right] + \Pi^{\tau(\sigma}\overline{\Psi}\gamma_\tau\Psi D^{[\mu}\Pi^{\beta\nu])}\overline{\Psi}\gamma_\beta\Psi\right). \tag{13.21}$$

In this way, the monopole baryon now reflects the quantum result that *i* times the classical Yang-Mills inverse (8.14), (8.15) is in fact equal to the Yang-Mills propagator operator obtained via a path integration (13.14) that takes advantage of a recursive understanding (13.12) of the gauge fields $G_\tau = I_{YM\,\nu\tau}J^\nu = -i\Pi_{\nu\tau}J^\nu$.

The final questions which arise, now that in (13.14), we have proved the *existence* of a non-trivial quantum Yang–Mills theory on $\mathbb{R}^4$ for *any* compact simple gauge group G, are questions as to the circumstances under which the Yang-Mills amplitude (13.15) and the specific QCD amplitude (13.16) will converge or diverge. Certainly, per (13.19), there is convergence at the zero recursive order ${}_0\Pi_{\mu\nu} \equiv \pi_{\mu\nu}$, because this is just the Abelian propagator. (Note, however, that this zero-order convergence still depends upon a Proca mass and / or $+i\varepsilon$.) But what happens for infinite nesting, or for various finite levels of nesting?

If one had a closed analytical expression for ${}_\infty\Pi_{\mu\nu}$ in (13.17) which "cashes out" the recursion out to infinite nesting, then one could simply use (10.3) in the form of:

$$\pi^\mu\pi^\nu = k^\mu k^\nu + k^\mu G^\nu + G^\mu k^\nu + G^\mu G^\nu = k^\mu k^\nu - V^{\mu\nu} \tag{13.22}$$

to write and evaluate (13.17) with $-{}_\infty V^{\mu\nu} = {}_\infty\left(k^\mu G^\nu + G^\mu k^\nu + G^\mu G^\nu\right)$ in the perturbative form:

$$\begin{aligned} {}_\infty\Pi_{\mu\nu} &= i\,{}_\infty\left(\left(\left(-g_{\mu\nu} + \pi_\mu\pi_\nu\left(\pi_\sigma\pi^{[\beta}\pi^\alpha\pi^{\sigma]}\right)^{-1}\pi^\alpha\pi^\beta\right)\left(\pi_\sigma\pi^\sigma\right)^{-1}\right)\right) \\ &= i\,{}_\infty\left(\left(\left(-g_{\mu\nu} + \left(k_\mu k_\nu - V_{\mu\nu}\right)\left(\left(k_\sigma k^{[\beta} - V_\sigma{}^{[\beta}\right)\left(k^\alpha k^{\sigma]} - V^{\alpha\sigma]}\right)\right)^{-1}\left(k^\alpha k^\beta - V^{\alpha\beta}\right)\right)\left(k_\sigma k^\sigma - V\right)^{-1}\right)\right) \end{aligned} . \tag{13.23}$$

In the limit $V_{\mu\nu} \to 0$ there is no recursion because all of the gauge fields $G^{\mu} \to 0$ and this becomes (13.19), with Proca mass and $+i\varepsilon$ as already noted. So it is to be anticipated that for small $V_{\mu\nu}$ this infinite recursion will continue to converge, with the inverses preventing any "catastrophe" by taking over the role of the Proca mass and $+i\varepsilon$ as elaborated in section 10. And, it ought not be surprising if for $_{\infty}V_{\mu\nu}$ above a certain threshold, this expression goes over from convergence to divergence, because a main point of the infinite recursion is that we now obtain an expression with up to infinite powers of the current density J^{ν}. But we also know from sections 4 and 11 that gauge field and quark confinement are a built-in aspect of the magnetic monopoles of Yang-Mills gauge theory, such that there is no net flux of any color across any closed surface of a Yang-Mills monopole. Thus, it is fair to anticipate that the threshold between convergence and divergence may also reveal itself to be related to the manner in which color is confined as established in sections 4 and 11. This too, however, we leave for another day.

In conclusion, we have now in (13.14) proved the *existence* of a non-trivial quantum Yang–Mills theory on $\mathbb{R}^4$ for *any* simple gauge group G, and in (13.15) we have applied this specifically to QCD. Thereafter, we have remarked as to how these findings may be used to approach doing Quantum Yang Mills including QCD calculations both analytically and numerically, and we have discussed how one might approach trying to understand the ranges of convergence and divergence of this Quantum Yang-Mills Theory. Coupled with the findings of section 10 for the mass gap solution, section 11 for the emergence of $SU(3)_C$ Chromodynamics, quark and gluon confinement, and meson interaction from the magnetic monopolies of Yang-Mills gauge theory, and section 12 for chiral symmetry breaking based on the same recursion that was central to developing Quantum Yang-Mills Theory in the present section 13 , this provides a substantially complete solution to the Yang-Mills and Mass Gap problem [1].

Finally, having shown how to obtain a non-linear Yang-Mills Quantum Field Theory using a recursive approach, we now have a first example courtesy of Yang-Mills, of how to develop non-linear quantum field theory in $\mathbb{R}^4$. It would certainly be of great interest to see what can be achieved if one applies a similar recursive analysis to gravitational theory and the Einstein-Hilbert action, which from the non-linear viewpoint that "gravitation gravitates," may well be the quintessential example of a recursive field theory.

14. Conclusion

In all of the foregoing, we have now shown how $SU(3)_C$ chromodynamics, which is the theory of strong interactions, is a corollary theory emerging naturally from the combination of nothing other than Maxwell / Weyl gauge theory with Yang-Mills theory. In the process, we have shown not only the emergence from the Maxwell / Yang-Mills combination of all that is to be expected from $SU(3)_C$ chromodynamics, but additionally, we have shown how the observed baryons containing three colored quarks in the ground state are the magnetic charges of Yang-Mills gauge theory and how these magnetic charges naturally confine their quarks and gluons but do pass mesons in order to interact. That is, we have explained quark and gluon confinement and how it is that strong interactions are mediated by mesons but not gauge fields. The main components of this understanding are in sections 4 and 11 and the key resultant equations are (11.1) and (11.18).

Additionally, we have demonstrated in section 10 based mainly on the development in section 8 how the inherent non-linearity of Yang-Mills theory may be used to solve the "mass gap" problem and yield a nuclear interaction that is short range notwithstanding its being based on massless gluon gauge fields, see specifically, equations (10.12) and (10.13). In section 12 we have shown the origin of "chiral symmetry breaking" in strong interactions. In section 9 we found that the non-linear nature of Yang-Mills theory contains a recursive aspect which later, in section 13, provides a useful tool for solving the Yang-Mills path integral in order to exactly, analytically arrive at quantum Yang-Mills theory. As a result of further developing Weyl's original geometric view of gauge theory, we in section 7 we uncovered a classical field equation (7.6) unifying gravitational theory with Weyl's gauge theory including both its Maxwell / Abelian and Yang-Mills variants, at the level of the Einstein equation for gravitation. Finally, in section 13, we use the recursive aspects of Yang-Mills theory from section 9 to develop and solve an exact, closed recursive path integral for Quantum Yang-Mills Theory and thereby prove the existence of a non-trivial quantum Yang–Mills theory on R^4 for any simple gauge group G.

References

[1] A. Jaffe and E. Witten, *Quantum Yang–Mills Theory*, Clay Mathematics Institute (2000)

[2] 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.

[3] R. Feynman, *The quantitative behavior of Yang–Mills theory in 2+1 dimensions*, Nucl. Phys. B188 (1981), 479-512.

[4] I. M. Singer, *The geometry of the orbit space for nonabelian gauge theories*, Phys. Scripta 24 (1981), 817-820.

[5] 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

[6] 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

[7] Zee, A., *Quantum Field Theory in a Nutshell*, Princeton (2003)

[8] H.C. Ohanian, *Gravitation and Spacetime*, Norton (1976)

[9] Misner, C. W., Thorne, K. S., and Wheeler, J. A., *Gravitation*, W. H. Freeman & Co. (1973)

[10] t'Hooft, G., *Magnetic Monopoles in Unified Gauge Theories*, Nuclear Physics B79, 276-284, (1974)

[11] A.M. Polyakov, *Particle Spectrum in the Quantum Field Theory*, JETP Lett. 20, 194-195 (1974)

[12] Cheng, T-P and Li, L-F, *Gauge theory of elementary particle physics*, Oxford (1984) (2011 reprint)

[13] 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

[14] Close, F. E., *An Introduction to Quarks and Partons*, Academic Press (1979)

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

[16] H. Weyl, *Gravitation and Electricity*, Sitzungsber. Preuss. Akad.Wiss., 465-480. (1918).

[17] H. Weyl, *Space-Time-Matter* (1918)

[18] H. Weyl, Electron und Gravitation, Zeit. f. Physik, 56 (1929), 330

[19] Minkowski, Hermann (1908/9), *Space and Time*, Physikalische Zeitschrift 10 (1909) 75–88

[20] A. Einstein, *On the Electrodynamics of Moving Bodies*, *Annalen der Physik* 17 (1905) 891

[21] A. Einstein, *Relativistic Theory of the Non-Symmetric Field*, in *The Meaning of Relativity*, Princeton University Press, December 1954

[22] http://pdg.lbl.gov/2013/tables/contents_tables_mesons.html

[23] Halzen, F., and Martin A. D., *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, J. Wiley & Sons (1984)

[24] J. Yablon, *Predicting the Binding Energies of the 1s Nuclides with High Precision, Based on Baryons which Are Yang-Mills Magnetic Monopoles*, Journal of Modern Physics, Vol. 4 No. 4A, 2013, pp. 70-93. doi: 10.4236/jmp.2013.44A010.
[25] Weinberg, S., *The Quantum Theory of Fields, Volume II, Modern Applications*, Cambridge (1996)
[26] Volovok, G. E., *The Universe in a Helium Droplet*, Clarendon Press – Oxford (2003)
[27] Howard Georgi and Sheldon Glashow, *Unity of All Elementary-Particle Forces*, Physical Review Letters, 32 at 438 (1974)
[28] http://pdg.lbl.gov/2013/reviews/rpp2012-rev-quark-model.pdf
[29] http://pdg.lbl.gov/2013/tables/rpp2013-sum-mesons.pdf