B.4: GAUGE FIELD THEORY

It is a striking fact about Nature that there exist gauge fields which play a key role in mediating interactions. At the "fundamental level" of particle physics one has the EM field, the various fields involved in the standard model, and the gravitational field. In various areas of condensed matter physics it has also been found useful to introduce gauge field descriptions of certain kinds of collective mode.

It is not possible to survey all of these. In what follows I will (i) discuss some of the underlying general features of all gauge fields, and why we have them, then (ii) discuss how they are described in a path integral formulation, using the ideas first developed by Fadeev & Popov. I will show how this works in both Abelian gauge theories (like QED) and non-Abelian theories (the Yang-Mills model). I will then give a hint of how these models are used in the real world. It is not possible to go into great detail here - there is no space - but we can at least see how these applications come about

B.4(a) GAUGE FIELDS - GENERAL FEATURES

(i) SOME HISTORY: We have known since the discovery of the EM field the General Relativistic description of spacetime that gauge fields are here to stay. The discovery of QM and slow development of QFT made it clear that the quantized versions of such fields mediated interactions in physics - the EM and spacetime fields being bosonic, with spin-1 and spin 2 respectively. Once the spin-statistics theorem was first given (by Pauli) it was also clear that only bosonic fields could be associated with classical macroscopic force fields. By the early 1960's (notably in the work of Weinberg) it was known that that the only consistent theories of massless bosonic fields had to be spin-0 (Higgs field - given a mass by the "Higgs" mechanism, actually first discussed by PW Anderson and N. Bogolubov), spin-1 (EM field), or spin-2 (gravitational field). In condensed matter theory, the key mechanism needed to give these gauge fields a mass (ie., a finite energy gap) was found - the appearance of an "order parameter", a concept first defined by Landau & Lifshitz in 1935, and developed with great effect by Landau, and later by others, including London (1938) BCS and Bogolubov (1956-1960), and Anderson (1958-1962). Anderson's paper, with the hypothesis of the Higgs boson, appeared in 1963. The papers of Higgs, and of Englert & Brout, were all written and published in 1964, and very quickly followed by papers of Kibble, (1964) and Guralnik & Hagen (1964). The appelation "Higgs boson" is a misnomer.

However there were serious mathematical problems, already noted in the 1950's in the context of QED. These concerned the renormalizability and the practical task of renormalizing these theories. The very small coupling constant α in QED made its practical use quite easy, but the same was not true of the weak or strong interactions - by the late 1950's many physicists were in despair over the application of QFT to these interactions, and Landau and others led the way to alternative formulations, such as the S-matrix theory (or "bootstrap" theory), reminiscent of behaviourist "black box" psychology. This detour wasted the time of many physicists (although it produced some useful mathematics), until the tide began to turn in 1966-67. This happened in a curious fashion. In the late 1950's very few physicists (with the exception of astrophysicists in the UK, Dirac, the Russians surrounding E.M. Lifshitz, and those in the former circle of Einstein in the USA) paid any attention to gravity or to GR. This astonishing neglect showed how close-minded some communities in science can be; but it also arose because GR seemed to have little relevance to earth-bound physics, and moreover, seemed to be quite irreconcilable with QM. At that time, only Einstein was talking publicly about his ideas of a "unified field theory".

Curiously it was Feynman who first broke away from this, at the Chapel Hill conference in 1957, where he argued, using thought experiments, that gravity had to be quantized. In a period of intense work between 1957-63 he tried to do this, but ran into a fundamental problem - gravity is non-renormalizable. It is a matter of some mystery why, after Feynman had had such success with both superfluid ^4He and the polaron problem using path integral methods, in the period 1952-1956 (diagrams are almost useless for these systems), that in the study of gravity he should abandon path integrals for the "Shut Up And Calculate" (SHUAC) methods of diagrams - and this is what he did. His approach led to the discovery of ghosts in gauge field theory, and was pursued to a successful but almost unreadable conclusion by B. de Witt (1964-1967); but it was completely unusable.

All this changed in the period 1967-1970 with 3 key developments. First, in 1966-67 Fadeev & Popov succeeded in formulating gauge field theory in path integral language - not just for QED, but also for Yang-Mills theory, & even, in principle, for gravity. The result was equivalent to deWitt's (published a week later), but in contrast, was simple to understand & use. Second, in 1967-68, Salam & Weinberg separately published theories unifying the weak & EM interactions into one "electroweak" gauge theory. Nobody paid any attention to this until 1970, when 't Hooft, a beginning PhD student in Utrecht, showed that the Salam-Weinberg theory was renormalizable - to everyone's astonishment - and then, in a tour de force, showed how to do calculations with it, using a combination of path integral methods and "dimensional regularization", a technique introduced by 't Hooft & his supervisor Veltman (& found by others independently at around the same time).

It is hard to imagine Feynman's reaction when he saw what he might have been able to achieve had he stuck with his own path integral methods!

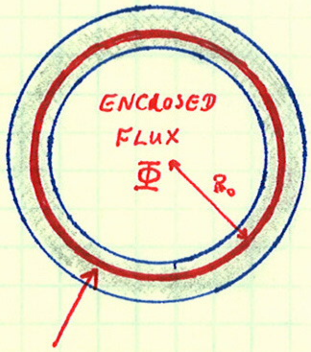
All of this work was subsequently developed into what is now the "standard model". The key further element to be found was the discovery of asymptotic freedom in the strong interactions (mediated by gluons, with a role also played by other bosonic fields). This discovery was actually made by 't Hooft in 1972, but he was discouraged from publishing it by Veltman; it was rediscovered by Gross & Wilczek in 1973, and also by Politzer in 1973 (whose supervisor, Coleman, was in contact with 't Hooft).

At the same time in condensed matter physics, the idea of the "gauge principle" was being applied to various systems. Here the impact was less clear, because we always have a "more microscopic" theory which does not require the gauge formulation; and gauge theories are hard to work with. Thus, eg., the gauge theories of high-T_c superconductivity and of the FQH (Fractional Hall liquid) have had little practical impact. However they have motivated many interesting new developments, and the idea of "spontaneously broken gauge symmetry", with the appearance of an order parameter, is central to all of condensed matter physics. Indeed, in focussing on this, CM physicists are really returning to the basic original of gauge theory, which goes back to Weyl in the 1920's, viz., that there be an apparent arbitrariness in the way that we parametrize physical variables such as phase, or even length and time in spacetime. In classical physics these variables really are redundant - they are eliminated by fixing a "standard of measurement". But this is not so in a quantum theory. Let us now see how this works, in the context of one of the simplest gauge theories, viz., QED and the quantized EM field. The remarks we make now arose from the analysis of Yang & Mills in 1954, of Aharonov & Bohm in 1957-59, and subsequent elaborations by many authors.

(ii) <u>THE AHARONOV - BOHM EFFECT</u> : When first proposed in the late 1950's, this analysis by Aharonov & Bohm (the mathematical analysis was actually done by MHL Pryce in Bristol, who later spent the years 1968-2001 at UBC), caused huge controversy. This was not least because Bohm, a brilliant young protegé of Oppenheimer during the war years, had been expelled from Princeton & from the USA in the early 1950's, accused of being a communist during the McCarthy years. With the help of Einstein & Pryce, then both at Princeton, Bohm went first to Brazil and then to Israel; Pryce then recruited him to Bristol in the UK in the late 1950's.

There are several key works by Aharonov & Bohm; the one we will be looking at concerns gauge fields. You may also find it interesting to look at Feynman's lectures in Physics, vol. 3, on this - Feynman was one of the early supporters of Bohm & his work.

Consider a situation where a single non-relativistic particle is forced to move on a 2-dimensional circle of radius R_0 - such a situation is now easy to organize with mesoscopic rings, or superconducting SQUIDS, or quantum wells, or even optically with photons. However when Chambers did the first experiment in 1961 it was not so simple.



Allowed Region for Particle.

The quantum mechanics of this problem is quite straightforward, & easily done using the Schrodinger eqtn. The Hamiltonian is assumed <u>STATIC</u>; then

$$\boxed{\mathcal{H} = \frac{1}{2m}\left(\underline{p} + q\underline{A}(\underline{r})\right)^2 + q\phi(\underline{r})} \tag{1}$$

and we will assume that a flux Φ is enclosed inside the ring, ie., that

$$\oint_C d\underline{l} \cdot \underline{A}(\ell) = R_0 \int d\theta\, A(\theta) = \Phi \tag{2}$$

This is not actually the problem solved by Aharonov & Bohm, and it is easy. Notice that it is possible to replace the full EM Lagrangian, derived from $\mathcal{H}$ by canonical transformation, viz.,

$$L = \frac{1}{2} m \dot{r}^2 + q \underline{A}(r) \cdot \dot{\underline{r}} - q\phi(\underline{r}) \tag{3}$$

where the momentum is $\quad p = \partial L / \partial \dot{r} = m\dot{r} + qA(\underline{r}) \quad (4)$

by a truncated version only valid on the 2-d ring of radius R_0. Let us assume that the electrostatic potential $\phi(r) = \phi_0$, a constant (no electric field). Then, ON THE RING, we have

$$\left.\begin{aligned} L &\rightarrow \frac{1}{2} m \dot{r}^2 + q \dot{\underline{r}} \cdot \underline{A}(\underline{r}) \\ &\underset{|\underline{r}| = R_0}{\longrightarrow} \frac{1}{2} I_0 \dot{\theta}^2 + q R_0 \dot{\theta} A(\theta) \end{aligned}\right\} \tag{5}$$

where the "moment of inertia" $I_0 = m R_0^2 \quad (6)$

(i) Consider first this problem when $\Phi = 0$; we just have a particle circulating on a ring, with

$$\mathcal{H}_0 = L_0 = \frac{1}{2} I_0 \dot{\theta}^2 \tag{7}$$

Then the normalized solutions to the Schrodinger eqtn: $\mathcal{H} \psi_\ell(\theta) = \epsilon_\ell \psi_\ell(\theta) \quad (8)$ are given by

$$\psi_\ell(\theta) = \frac{1}{\sqrt{2\pi}} e^{i\ell\theta} \tag{9}$$

where $\ell = 0, \pm 1, \pm 2, \ldots$ etc., is an integer (the angular momentum quantum number), and θ is, as above the angular coordinate. Here we can treat it as COMPACT, ie., $0 \leq \theta \leq 2\pi$.

The propagator for the particle is also easily found; we have

$$\left.\begin{aligned} G(\theta_1 \theta_2 ; t_1 t_2) \equiv G(\theta, t) &= \sum_\ell \psi_\ell(\theta) \psi^*_\ell(0) e^{\frac{i}{\hbar}\epsilon^0_\ell t} \\ &\equiv \sum_{\ell} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{\psi_\ell(\theta) \psi^*_\ell(0)}{\omega - \epsilon^0_\ell} \end{aligned}\right\} \tag{10}$$

where the eigenvalues are just $\quad \epsilon^0_\ell = \frac{\hbar^2}{2 I_0} \ell^2 \quad (11)$

Using either form in (10), we get the result

$$\boxed{G(\theta, t) = \left(\frac{I_0}{2\pi i \hbar t} \right)^{1/2} \sum_{\ell=-\infty}^{\infty} \exp \left\{ \frac{i}{\hbar} I_0 \frac{(\theta + 2\pi \ell)^2}{2t} \right\}} \tag{12}$$

which we can write in more compact form as

$$G(\Theta,t) = \left(\frac{I_0}{2\pi i\hbar t}\right)^{1/2} e^{\frac{i}{\hbar} I_0 \,\Theta^2/2t} \;\vartheta_3\!\left(\frac{\pi I_0 \Theta}{\hbar t} \,;\, \frac{2\pi I_0}{\hbar t}\right) \tag{13}$$

where $\vartheta_3(z,t)$ is the biperiodic Jacobi Θ-function, defined in the complex plane z, and with the series representation

$$\vartheta_3(z,t) = \sum_{n=-\infty}^{\infty} e^{i(\pi t n^2 + 2zn)} \tag{14}$$

ie., a discretized version of a Gaussian integral.

Why do we have such a complicated result? The first part of (13) looks just like a free particle of mass I_0; so where does the Jacobi Θ-fn. come from? In this calculation it comes from the compactness of the variable Θ; we are dealing with a particle in a box. But now we can rederive this result in a quite different way, extending the domain of Θ so that $-\infty < \Theta < \infty$. Let's rederive the result (13) using path integrals. Then we have

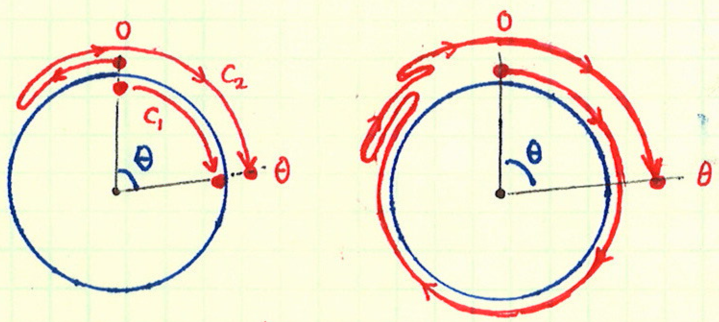
$$\left.\begin{aligned} G(\Theta,t) &= \int_{\Theta(0)=0}^{\Theta(t)=\Theta} \mathcal{D}\underline{r}(t)\, e^{\frac{i}{\hbar} S[\underline{r},\dot{\underline{r}}]} \\ &= A(t)\, e^{\frac{i}{\hbar} S_{cl}(\Theta,t)} \end{aligned}\right\} \tag{15}$$

where $A(t)$ is just the fluctuation determinant for a particle of mass I_0 in free particle motion:

$$A(t) = \left(I_0/2\pi i\hbar t\right)^{1/2} \tag{16}$$

Now, however, we must be careful with the classical action. Recall that we are supposed to sum over ALL paths. In Fig (a) at left we see 2 simple paths which begin at $\Theta(0) = 0$, and terminate at $\Theta(t) = \Theta$. Notice that in both cases the WINDING NUMBER n is zero.

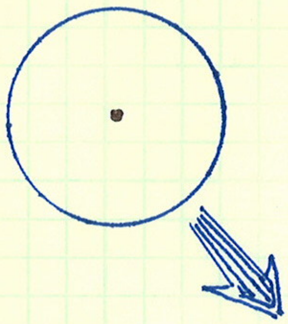
However this is not true of the path in (b) at right. This has winding number $n = 1$. And yet this path must also be included in $G(\Theta,t)$.



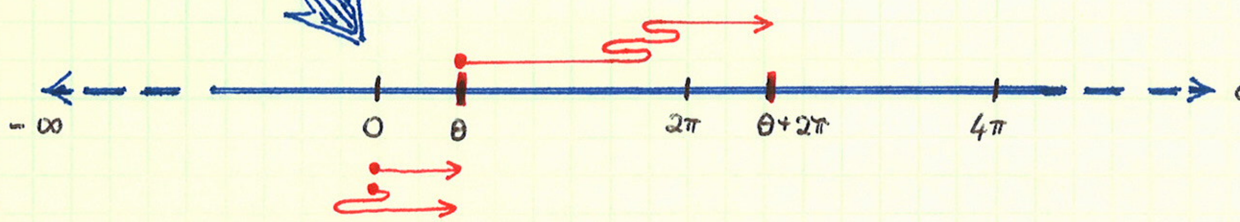
(a) (b)

The simplest way to then derive the answer for $S_{cl}(\Theta,t)$ is to either (a) note that any path beginning at $\Theta(0) = 0$ and ending at $\Theta(t) = \Theta$ can be decomposed into a path going from $\Theta(0) = 0$ to $\Theta(\tau) = 0$, where $\tau \leq t$, and then another path going to $\theta(t) = \theta$; but we now sum over all possible winding numbers for the first of these 2 paths. Or else (b), we just "unfold" the ring, as shown below. Now we see that any point on the line at

$$\theta' = \theta + 2n\pi \qquad n = 0, \pm 1, \pm 2, \ldots \tag{17}$$



must be identified with θ, and we must then sum over all paths leading to any of these points. The paths shown on the last page are shown again below.



However it is clear from this diagram that we are dealing with a free particle on the line, and so we immediately find that

$$e^{\frac{i}{\hbar} S_{cl}(\theta, t)} = \sum_{n=-\infty}^{\infty} e^{\frac{i}{\hbar} I_0 \frac{(\theta + 2\pi n)^2}{2t}} \tag{18}$$

have summed over the different winding numbers. Inserting (18) and (16) into (15), we again recover (13).

(ii) Now let's go to the finite flux case. The energy levels are shifted, and the eigenfunctions change; we have

$$\left.\begin{aligned} \psi_\ell(\theta, \tilde{\varphi}) &\quad \frac{1}{\sqrt{2\pi}} e^{i(\ell + \tilde{\varphi}/2\pi)\theta} \\ \epsilon_\ell(\tilde{\varphi}) &= \frac{\hbar^2}{2I_0} (\ell + \tilde{\varphi}/2\pi)^2 \end{aligned}\right\} \tag{19}$$

and from this we can derive the new answer. Before doing so, notice a crucial point. The answer does not depend on how the flux is distributed inside inside the ring, or even on whether the flux density (ie., the magnetic field) is finite at the ring itself - it depends only on the total flux, or rather, the "dimensionless flux" $\tilde{\varphi}$, defined as

$$\left.\begin{aligned} \tilde{\varphi} &= \frac{q}{\hbar} \oint d\underline{\ell} . A(\ell) = \frac{2\pi q}{\hbar} \Phi = \Phi / \Phi_0 \\ &\text{where} \quad \Phi_0 = h/q \end{aligned}\right\} . \tag{20}$$

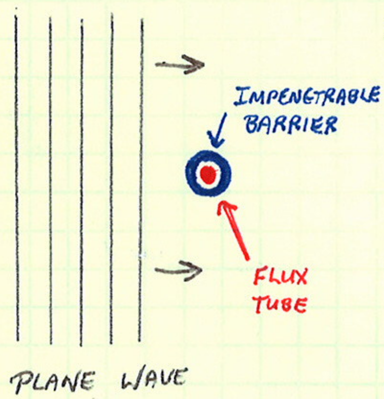
and $\Phi_0 = h/q$ is the flux quantum for charge q.

From here on it is clear that for a path with winding number n, we must add a phase $n\tilde{\varphi}$ to the action exponent, and so now we get

$$\left.\begin{aligned} G(\theta, t; \tilde{\varphi}) &= \left(\frac{I_0}{2\pi i\hbar t}\right)^{1/2} \sum_{n=-\infty}^{\infty} \exp\left\{ in\tilde{\varphi} + \frac{i}{\hbar} I_0 \frac{(\theta + 2\pi n)^2}{2t} \right\} \\ &\equiv \left(\frac{I_0}{2\pi i\hbar t}\right)^{1/2} e^{\frac{i}{\hbar} I_0 \theta^2/2t}\, \vartheta_3\!\left(\frac{\pi I_0 \theta}{\hbar t} - \frac{\tilde{\varphi}}{2} ; \frac{2\pi I_0}{\hbar t}\right) \end{aligned}\right\} \quad (21)$$

and in both forms of this answer, we see what is already obvious from the new eigenfunctions & energies in (19), viz., that the flux has shifted the answer - it is a "phase shift" operator.

Now let's go to the problem that Aharonov & Bohm actually looked at in their famous paper. Their problem is shown below - we have an INFINITESIMAL FLUX TUBE, still carrying flux Φ, but now confined to a very thin filament, which we will now treat as a δ-function. We then have



$$\mathcal{H} = \frac{1}{2m}\left(\underline{p} + q\underline{A}(\underline{r})\right)^2 \quad (22)$$

where

$$\left.\begin{aligned} \underline{A}(\underline{r}) &= \Phi \frac{\hat{z} \times \hat{r}}{2\pi r} \\ &\equiv \tilde{\varphi}\, \Phi_0 \frac{\hat{z} \times \hat{r}}{2\pi r} \equiv \hat{\theta} \frac{\Phi}{2\pi r} \end{aligned}\right\} \quad (23)$$

where $\hat{z}$ and $\hat{r}$ are unit vectors, as is $\hat{\theta}$. To make the point even more clearly, let's surround the flux tube at the origin by an infinite potential barrier OUTSIDE the flux tube, but with a radius r_0 which we will also take to be infinitesimal. Then nothing from outside can penetrate, and the flux tube is isolated from the outside world.

And yet, from what we have done above, we know that even though the electric field $\underline{E}(\underline{r}) = 0$ everywhere, and $\underline{B}(\underline{r}) = 0$ except in the flux tube, still a QUANTUM PARTICLE moving outside will feel the flux! This is utterly different from classical mechanics, where the particle dynamics is governed by the Lorentz eqtn, which is LOCAL:

$$m\ddot{\underline{r}}(t) = q\underline{E}(\underline{r}) + q\left(\dot{\underline{r}} \times \underline{B}(\underline{r})\right) \quad (24)$$

Now, we imagine a plane wave state of the particle incident on the flux tube. If we let the barrier potential radius $r_0 \to 0$, then for some finite wavelength incident wave, with wavelength $\lambda = 2\pi/k$, the scattering cross-section (in 2d) is

$$\sigma_k \sim k r_0 \ln(k r_0) \quad (25)$$

so the particle does not scatter off the potential barrier. Nevertheless it still scatters off the flux tube, even though it never sees the flux. The Schrodinger eqtn in 2-d, for the Hamiltonian (i) with $\phi(r) = 0$, is

$$\left\{ \partial_r^2 + \frac{1}{r}\partial_r + \left[k^2 - \frac{1}{r^2}\left(i\partial_\theta + \tilde{\varphi}\right)^2 \right] \right\} \Psi(r,\theta) = 0 \tag{26}$$

and for $r \neq 0$, the solution can be written in terms of the eigenfunctions of this eqtn:

$$\left.\begin{aligned} \Psi(r,\theta) &= \sum_{\ell k} c_{\ell k}\, \psi_{k\ell}(r,\theta) = \sum_\ell c_\ell\, e^{i\ell\theta} J_{|\ell+\tilde{\varphi}|}(kr) \\ &\qquad (r > r_0) \end{aligned}\right\} \tag{27}$$

with the eigenvalues:

$$\left.\begin{aligned} \mathcal{H}\, \psi_{\ell k}(r,\theta) &= \epsilon_{\ell k}\, \psi_{\ell k}(r,\theta) \\ \epsilon_{\ell k} &= \hbar^2 k^2 / 2m \end{aligned}\right\} \tag{28}$$

from which we have

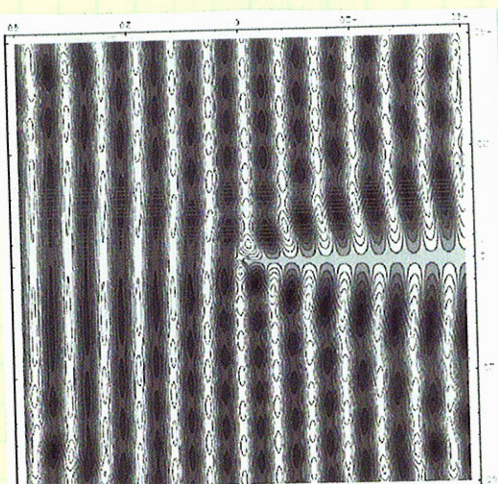
$$G(r_2 r_1; \theta_2 \theta_1; t) = \frac{1}{2\pi} \sum_\ell \int k\,dk\, J_{|\ell+\tilde{\varphi}|}(kr_1)\, J_{|\ell+\tilde{\varphi}|}(kr)\, e^{i[(\theta_2-\theta_1)\ell - \epsilon_{lk} t]} \tag{29}$$

However this part is easy. More messy is the determination of the scattering amplitude $f_k(\theta)$, defined by the asymptotic form of $\Psi(r,\theta)$, by

$$\Psi(r,\theta) = e^{ikx} + \frac{1}{\sqrt{r}} f_k(\theta)\, e^{ikr} \qquad (r \to \infty) \tag{30}$$

This was the problem that Pryce solved for Aharonov, to give the result

$$f_k(\theta) \simeq - \frac{e^{i\pi/4}}{\sqrt{2\pi k}} \sin \pi\tilde{\varphi}\; \frac{e^{i\theta/2}}{\cos \theta/2} \tag{31}$$



constant phase 'eikonal' plot

The form of this result is rather complex, & is illustrated at left. From the result in (31) we see that the scattering amplitude is periodic in the flux $\tilde{\varphi}$; indeed, when $\tilde{\varphi} = n$, an integer, it has no effect at all.

Notice that (31) diverges for $\theta \to \pi$, and in fact it breaks down for both forward & backward scattering. The figure at left gives an idea of how it actually behaves.

Of course what is so strange about this result is that it shows that the particle dynamics, in the QUANTUM THEORY, is controlled not by $\underline{B}(\underline{r})$ and or $\underline{E}(\underline{r})$, but by $\underline{A}(\underline{r})$. This means that in the quantum theory, it is $\underline{A}(\underline{r})$ (and more generally $\underline{A}(\underline{r},t)$) that is the fundamental object, and not $\underline{B}(\underline{r},t)$ or $\underline{E}(\underline{r},t)$, which are merely derived from $\underline{A}(\underline{r},t)$.

This is the exact opposite of classical EM theory. There, the fundamental physical quantities are $\underline{B}(\underline{r},t)$ and $\underline{E}(\underline{r},t)$; and from (24), we see that these 2 fields entirely control the dynamics of electric charge.

(iii) CLASSICAL EM vs. QED : What is the reason for this fundamental difference between classical EM theory, and its quantized version? For many years this was not properly understood, and yet the reason is to be found in eqtn. (21). We notice that the term in the exponential appears without an accompanying $\hbar$, which is buried in the definition of $\tilde{\varphi}$; let's rewrite it as

$$G(\theta, t; \tilde{\varphi}) = A(t) \sum_{n=-\infty}^{\infty} \exp \frac{i}{\hbar}\left\{ n\hbar\tilde{\varphi} + I\frac{(\theta+2\pi n)^2}{2t}\right\} \tag{32}$$

$$\text{where} \quad A(t) = \left(\frac{I_0}{2\pi i\hbar t}\right)^{1/2} \tag{33}$$

Now suppose we consider the limit $\hbar \to 0$. It is singular; the phase in the exponent diverges. However we notice that the term in $\tilde{\varphi}$ is independent of $\hbar$, and it DISAPPEARS FROM THE EXPONENT. This is because there are actually 2 quantum parameters here, $\hbar$ and q, and we are perfectly entitled to treat them as independent. Thus let's write (32) as

$$\left.\begin{aligned} G(\theta, t; \hbar, q) &= A(\hbar t)\sum_{n=-\infty}^{\infty} e^{\frac{i}{2} nq\frac{\Phi}{\pi}}\; e^{\frac{i}{2} I_0 \frac{(\theta+2n\pi)^2}{\hbar t}} \\ &\equiv A(\hbar t)\sum_{n=-\infty}^{\infty} e^{in\omega_c(q\Phi)}\; e^{i\psi_n(\hbar t,\theta)} \end{aligned}\right\} \tag{34}$$

where the prefactor $A(\hbar t)$ and the phase $\psi_n(\hbar t, \theta)$ both diverge as $\hbar t \to 0$:

$$\left.\begin{aligned} A(\hbar t) &= \left(\frac{I_0}{2\pi i\hbar t}\right)^{1/2} \xrightarrow[\hbar t\to 0]{} \infty \\ \psi_n(\hbar t, \theta) &= \frac{I_0}{2}\frac{(\theta+2n\pi)^2}{\hbar t} \xrightarrow[\hbar t\to 0]{} \infty \end{aligned}\right\} \tag{35}$$

whereas the TOPOLOGICAL PHASE $\omega_c(q\Phi)$ does not:

$$\omega_c = \frac{q\Phi}{2\pi} \qquad (\text{independent of } \hbar, t). \tag{36}$$

Thus we see already, from looking at a simple non-relativistic problem in the motion of a charged particle coupled to a static EM field, that there is a crucial term in the dynamics that exists in the quantum version of the theory, but not in the classical.

If we think about this a little more, it is not all that obvious why there ought to be any relationship between the quantum and classical versions of electrodynamics. After all, classical EM theory makes no reference to Dirac electrons or holes, which exist in QED, and it is not obvious why the form of the Lagrangians or actions for the 2 theories should look the same. Thus we can, if we like, start by asking - why should there be a quantum generalization of a classical gauge theory like QED, and what should it look like?

B.4(b) DERIVATION OF GAUGE FIELD THEORIES

In this sub-section we will give some of the well-known arguments that are used to derive the existence and form of gauge theories. A key part of what follows is the comparison between the quantum and classical versions of these theories. I will focus mainly on the general form of these theories, with not too much attention paid to examples. Note, as already emphasized in the previous sub-section, that physics is full of different sorts of gauge theory - why this is should become clear in the following.

(i) QUANTUM ELECTRODYNAMICS: A U(1) GAUGE THEORY

QED is the simplest sort of gauge theory - we shall see why this is in the course of the derivations to be given. It is important in what follows to compare the classical and quantum versions of electrodynamics, so we begin by reviewing classical EM theory. In what follows it will be assumed that everyone is familiar with special relativity and relativistic notation.

CLASSICAL ELECTRODYNAMICS: Classical Electrodynamics is a theory of sources coupled to fields, and it can be phrased entirely in terms of the fields $\underline{E}(x)$ and $\underline{B}(x)$, and the charge 4-current $J(x)$. However, as is well known, this is not the best way to formulate it, and is not ideal for the generalization to QED.

Let's see how we can set up classical EM theory starting with an experimental fact, viz., that there exist fields $\underline{E}(x)$ and $\underline{B}(x)$, and that they act upon electric charges according to the Lorentz force law in (24) above. In 4-vector notation we rewrite the Lorentz eqtn. as

$$\boxed{f^\mu = q F^{\mu\nu} U_\nu} \qquad (37)$$

where U_ν is the 4-velocity vector, ie.,

$$U_\nu = \frac{dx_\nu}{d\tau} \qquad (38)$$

where τ is the proper time interval (on a worldline). In what follows we will assume the following conventions; the infinitesimal interval $d\underline{s}$ is related to the metric by

$$c^2 d\tau^2 = ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu \xrightarrow[\text{FLAT SPACE}]{} \eta_{\mu\nu}\, dx^\mu dx^\nu \qquad (39)$$

where the coordinate 4-vector is $x^\mu = (x^0; x^1, x^2, x^3) \equiv (ct; x, y, z)$, and the flat space Minkowski metrix is taken to be

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix} \qquad (40)$$

with signature -2. If in some general pseudo-Riemannian spacetime we define a set of basis vectors $\underline{e}_\mu$ (contravariant) and $\underline{e}^\mu$ (covariant), then

$$\left.\begin{aligned} d\underline{s} &= \underline{e}_\mu dx^\mu = \underline{e}^\mu dx_\mu \\ \text{and} \quad g^{\mu\nu} &= \underline{e}^\mu \cdot \underline{e}^\nu \qquad g_{\mu\nu} = \underline{e}_\mu \cdot \underline{e}_\nu \qquad g^\mu_\nu = \underline{e}^\mu \cdot \underline{e}_\nu = \delta^\mu_\nu \end{aligned}\right\} \qquad (41)$$

The 4-vectors of main interest to us will be:

4-velocity: $\underline{U} = U^\nu \underline{e}_\nu$, with $U^\nu = \gamma_u(c, \underline{u})$, where $\gamma_u = (1-\bar{u}^2)^{-1/2}$, and $\bar{u} = u/c$, and $\underline{u} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right)$.

Note we then have $d\tau = dt/\gamma_u$, and also

$$U^2 = U^\nu U_\nu = \left(\frac{ds}{d\tau}\right)^2 = c^2 \qquad (42)$$

4-acceleration: $$\underline{A} = A^\nu \underline{e}_\nu, \quad \text{where} \quad A^\nu = \frac{d^2x^\nu}{d\tau^2} = \frac{dU^\nu}{d\tau} \qquad (43)$$

so that $$U_\nu A^\nu = 0 \qquad (44)$$

4-current density: $$\underline{J}(x) = \rho_0(x)\underline{U}(x) = J^\nu(x)\underline{e}_\nu(x) \qquad (45)$$

where $\rho_0(x)$ is the proper charge density; then
$$\left.\begin{aligned} J^\nu &= \rho_0 \gamma_u(c, \underline{u}) \\ &\equiv (\rho c, \underline{j}) \end{aligned}\right\} \qquad (46)$$

where $\rho(x) = \gamma_u \rho_0(x)$, and $\underline{j}(x) = \rho(x)\underline{u}(x)$ is the 3-current density; from eqtn (42), we have

$$J^2 = J^\mu J_\mu = \rho^2 c^2 - j^2 \qquad (47)$$

is a Lorentz-invariant quantity.

Returning now to the Lorentz force eqtn, we notice that a contraction of (37) with a 4-vector U_μ gives $f^\mu U_\mu = qF^{\mu\nu}U_\mu U_\nu = 0$, because a 4-force, like a 4-acceleration, must be perpendicular to its associated velocity (cf eqtn. (44)); thus $F_{\mu\nu}$ must be antisymmetric, ie.,

$$F_{\mu\nu} = -F_{\nu\mu} \qquad (48)$$

We have introduced the tensor $F_{\mu\nu}(x)$ as a way of describing the Lorentz-force law in a relativistically invariant way — the presence of the velocity vector $\dot{\underline{r}} = \underline{u}$ in (24) naturally leads to a formulation in terms of the velocity 4-vector $U(x)$, and thence to an eqtn of the form in (37). If we now want to recover (24), we write $F_{\mu\nu}(x)$ in the form:

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & B_y & B_x & 0 \end{pmatrix} \tag{49}$$

Now, we also wish to find a <u>source equation</u> for the EM field - we may treat this as a general theoretical requirement, or again base it on experiment (EM fields act on charges, and are acted on by them as well). Rather than appeal to experiment, we can simply ask what is the natural relativistically invariant form for electric charge to generate the field. Since the charge $\rho(x)$ is simply one component of the 4-current $J(x)$, we look for a linear eqtn relating $F^{\mu\nu}(x)$ to $J^{\mu}(x)$. The obvious way to do this is write

$$\partial_{\mu} F^{\mu\nu}(x) = K J^{\nu}(x) \qquad K = \begin{Bmatrix} 4\pi/c & (\text{cgs}) \\ \mu_0 & (\text{MKS}) \end{Bmatrix} \tag{50}$$

(any other 4-vector A_{μ} contracted with $F^{\mu\nu}$ would do, but experiment reveals no other such quantities); the constant K is given by experiment.

At first it might seem that eqtns. (37) and (50) characterize the theory properly. But actually the components of $F_{\mu\nu}$ are not independent of each other - there are 6 apparently independent components in (49), but only 4 quantities in (50). One can either $F_{\mu\nu}(x)$ has to be therefore constructed from a 4-vector, or appeal to experiment for the relationship between the components of $F_{\mu\nu}(x)$. Both lines of argument lead to the conclusion that we can write $F_{\mu\nu}(x)$ in the form

$$F_{\mu\nu}(x) = \partial_{\mu} A_{\nu}(x) - \partial_{\nu} A_{\mu}(x) \tag{51}$$

where

$$A^{\mu}(x) = \begin{Bmatrix} (\phi(x)/c,\ \underline{A}(x)) & (\text{MKS}) \\ (\phi(x),\ \underline{A}(x)) & (\text{cgs}) \end{Bmatrix} \tag{52}$$

and $\phi(x)$, $A(x)$ are the electric and magnetic potentials. Another way to enforce the restriction to 4 independent components is the Bianchi identity, written as

$$\partial_{[\lambda} F_{\mu\nu]} \equiv \partial_{\lambda} F_{\mu\nu} + \partial_{\nu} F_{\lambda\mu} + \partial_{\mu} F_{\nu\lambda} = 0 \tag{53}$$

which actually follows directly from (51). There are thus 2 different ways we can write the FIELD EQTNS of classical electromagnetism:

$$\boxed{\begin{array}{c} \partial_{\mu} F^{\mu\nu}(x) = \mu_0 J^{\nu}(x) \\ F^{\mu\nu} = \partial^{\mu} A^{\nu} - \partial^{\nu} A^{\mu} \quad \text{OR} \quad \partial_{\lambda} F_{\mu\nu} + \partial_{\nu} F_{\lambda\mu} + \partial_{\mu} F_{\nu\lambda} = 0 \end{array}} \tag{54}$$

So far so good. Now observe that we have a rather peculiar theory, for the key OBSERVABLE variables are $E(x)$, $B(x)$, and of course $J(x)$, but the underlying field variable, which is not observable, for the EM field is $A^{\mu}(x)$.

It then follows that if we make any changes of variables, coordinate transformations, etc., the quantities $F_{\mu\nu}(x)$ and $J(x)$ ought to be invariant, but $A^\mu(x)$ does not have to be. This immediately leads to the possibility of GAUGE TRANSFORMATIONS. Suppose we make the transformation

$$A^\mu(x) \rightarrow \tilde{A}^\mu(x) = A^\mu(x) + \alpha^\mu(x) \qquad (55)$$

Then we have $$\tilde{F}_{\mu\nu}(x) = (\partial_\mu A_\nu - \partial_\nu A_\mu) + (\partial_\mu \alpha_\nu - \partial_\nu \alpha_\mu) \qquad (56)$$

and for $\tilde{F}_{\mu\nu} = F_{\mu\nu}$, we require $\alpha^\mu = \partial^\mu \chi$, so: $$\tilde{A}^\mu(x) = A^\mu + \partial^\mu \chi(x) \qquad (57)$$

Thus any one of an infinite set of functions $\tilde{A}^\mu(x)$ is just as valid as $A^\mu(x)$ for the treatment of classical EM fields.

We now wish to write an action functional for the classical EM theory, in preparation for the transition to QED. There will be 2 parts to this action, the matter part and the field part.

<u>Particle Action in EM theory</u>: For a set of particles of mass m_j, at spacetime coordinates x_j, it is well known that the action has the form

$$S_p = -\sum_j m_j c \int ds_j \equiv -c^2 \sum_j \int dt_j \,(1 - u_j^2/c^2)^{1/2} \qquad (58)$$

Suppose we ignore this term, concentrating only on the electromagnetic part that comes from the interaction of the current $J^\mu(x)$ with the gauge field $A_\mu(x)$, which from now on we take to be the fundamental EM field. The simplest scalar term for a Lagrangian density combining these two is just $J_\mu A^\mu$, and actually experiment tells us that we should have.

$$S_M = -\int d^4x \; J^\mu(x) A_\mu(x) \qquad (59)$$

At first glance this coupling seems problematic, because it is apparently not invariant under the gauge transformation of (57). However one can actually show that it is gauge invariant, as follows. Suppose we gauge transform (59):

$$\left.\begin{aligned} A^\mu(x) &\rightarrow A^\mu + \partial^\mu \psi(x) \quad \Rightarrow \quad S_M \rightarrow \tilde{S}_M \\ \tilde{S}_M &= -\int d^4x\, J^\mu (A_\mu + \partial_\mu \chi) = -\int d^4x \left[J^\mu A_\mu + \partial_\mu (J^\mu \chi) - (\partial_\mu J^\mu)\chi \right] \end{aligned}\right\} \quad (60)$$

Now the 2nd term, $-\int d^4x\, \partial_\mu (J^\mu \chi)$, can be rewritten as a surface integral of $J^\mu \psi$ at infinity, and we will assume no current sources at infinity. This leaves the 3rd term, and we see the coupling term (59) will be gauge invariant if the current conservation eqtn

$$\boxed{\partial_\mu J^\mu(x) = 0} \qquad (61)$$

is satisfied. But the truth of this is easily demonstrated, starting from the 1st eqtn of motion in (54); for if we take the derivative ∂_ν of (54), the left side of (54) must vanish because of the antisymmetric property (48) of the field $F_{\mu\nu}(x)$.

Field Action in EM theory : We want to find a scalar quantity which is gauge invariant for the EM field action. This time we cannot get away with using combinations of $A_\nu(x)$, because any combination like $A^\nu A_\nu$ will not be gauge invariant. Thus we must resort to $F_{\mu\nu}$, and the correct form of the field action is in fact

$$S_{EM} = -\frac{1}{4\mu_0}\int d^4x\; F^{\mu\nu}(x)\,F_{\mu\nu}(x) \qquad \text{(MKS units)} \tag{62}$$

where the prefactor is fixed by experiment.

Thus we find that the total EM action is

$$\boxed{S_{cl}[J^\mu, A^\mu] = -\int d^4x \left\{ \frac{1}{4\mu_0} F_{\mu\nu}(x) F^{\mu\nu}(x) + J_\mu(x) A^\mu(x) \right\}} \tag{63}$$

and, by varying this action, we can recover the eqtns of motion.

QUANTUM ELECTRODYNAMICS : Now let us start again, this time with the Dirac eqtn. for a fermionic field having the bare action

$$S_0 = \int d^4x\; \bar\psi(x)\,[i\gamma^\mu\partial_\mu - m]\,\psi(x) \tag{64}$$

What we now wish to show is that an action equivalent to the classical action above can be derived for the Dirac electron coupled to an EM field — but the arguments used are a little different (at least at first glance). They have the great advantage of being easily generalizable to a variety of matter fields (although one can quibble rather strongly in the case of quantum gravity).

Now one can make a very simple "gauge transformation" on (64), noting that the phase of $\psi(x)$ is arbitrary. Thus we make the GLOBAL GAUGE TRANSFORMATION

$$\psi(x) \rightarrow e^{-i\theta}\,\psi(x) \tag{65}$$

Under this transformation, S_0 is invariant; the system possesses a global $U(1)$ symmetry. However, suppose we allow θ to depend on the spacetime coordinates; this move is of course highly non-trivial, and is suggested by the following considerations:

(i) the original motivation of Weyl, in the 1920's; since one can envisage, in general relativity, a change of "coordinate measure" as one moves around in space, why not the same for phase measure (hence the name "gauge", meaning a measurement scale).

(ii) Quantum Mechanics naturally suggests, particularly in its path integral form, the role of phase as a fundamental variable in the theory. This reappears in the Aharonov-Bohm effect.

We shall see that the consequences of introducing such a LOCAL GAUGE TRANSFORMATION give further reason to look at it. So we now consider the transformation:

$$\psi(x) \longrightarrow e^{i\theta(x)}\psi(x) \tag{66}$$

We immediately see that S_0 is not invariant under this transformation; in fact we get

$$\left.\begin{aligned} S_0 \longrightarrow & \int d^4x\, \bar{\psi}(x)\left[i\gamma^\mu\left(\partial_\mu - i\partial_\mu\theta(x)\right) - m\right]\psi(x) \\ = & \; S_0 + \int d^4x\, \bar{\psi}(x)\gamma^\mu\partial_\mu\theta(x)\psi(x). \end{aligned}\right\} \tag{67}$$

However, we can have a gauge-invariant theory if we do the following

- Replace ∂_μ by a "gauge-covariant" derivative D_μ so that $D_\mu\psi(x) \rightarrow e^{-i\theta(x)}D_\mu\psi(x)$ is invariant. This will work if:
- We write $D_\mu = \partial_\mu + iqA_\mu$, and at the same time suppose that the field $A_\mu(x)$ transforms according to

$$\boxed{A_\mu(x) \rightarrow A_\mu(x) + \frac{1}{q}\partial_\mu\theta(x)} \tag{68}$$

thereby cancelling the extra term in (67).

Thus we are led to replace (64) by

$$\left.\begin{aligned} S_0 = & \int d^4x\, \bar{\psi}(x)\left[i\gamma^\mu D_\mu - m\right]\psi(x) \\ \equiv & \int d^4x\, \bar{\psi}(x)\left[i\gamma^\mu\left(\partial_\mu + iqA_\mu(x)\right) - m\right]\psi(x) \end{aligned}\right\} \tag{69}$$

and then, recognizing that the field $A_\mu(x)$ must have its own individual term in the Lagrangian (it is a new field), which must be gauge invariant, we are lead to the total action:

$$\boxed{S_{QED} = \int d^4x\left\{\bar{\psi}(x)\left[i\gamma^\mu\partial_\mu - m\right]\psi(x) - J_\mu(x)A^\mu(x) - \frac{1}{4\mu_0}F_{\mu\nu}(x)F^{\mu\nu}(x)\right\}} \tag{70}$$

where the Dirac current $J_\mu(x)$ is

$$\boxed{J_\mu(x) = q\bar{\psi}(x)\gamma_\mu\psi(x)} \tag{71}$$

Thus we have shown that the gauge invariance of the Dirac electron terms in the action, under a local PHASE transformation, leads naturally to the existence of a gauge field $A_\mu(x)$ of the same kind as appears in the original classical action for the EM field! We now see another reason to take all this seriously; the gauge invariance naturally prevents the existence of a term in $A^2(x)$, which would give the photon a mass, something excluded by experiment. It is also interesting to note that the field $F_{\mu\nu}(x)$

acquires a new significance in the quantum-mechanical theory (which in a path integral formulation can be related directly back to the classical theory, as we shall see). Consider the action of the CURVATURE operator

$$\hat{R}_{\mu\nu} = [\mathcal{D}_\mu, \mathcal{D}_\nu] = \mathcal{D}_\mu \mathcal{D}_\nu - \mathcal{D}_\nu \mathcal{D}_\mu \tag{72}$$

on the wave-function; we immediately find that

$$\hat{R}_{\mu\nu} \psi(x) = iq F_{\mu\nu}(x) \psi(x) \tag{73}$$

so that the different components of the field intensity $F_{\mu\nu}(x)$ (ie., the "physical" fields $\underline{E}(x)$ and $\underline{B}(x)$) are just components of this curvature. Thus we can argue that in reality we have

$$\left.\begin{array}{ll} \underline{A}(x) : & \text{underlying "EM field", or "EM vacuum", not directly accessible to us in classical EM theory;} \\ & \text{accessible as a PHASE variable in QM, modulo gauge transformations.} \\ F_{\mu\nu}(x) : & \text{The "curvature" or "polarization" of } \underline{A}(x)\text{, directly accessible to us via its effects on charge,} \\ & \text{with components } \underline{E}(x), \underline{B}(x)\text{, in both classical EM and QED.} \end{array}\right\} \tag{80}$$

We note of course that $F_{\mu\nu}(x)$ is gauge invariant, as we have already seen. Notice 2 key features of this $U(1)$ theory:

(i) The form of the coupling to the gauge field (indeed, the existence of the gauge field, in this argument) arises solely from the transformation of the MATTER field under local phase transformations; one ends up with a covariant or "minimal" coupling. The requirement of gauge invariance then uniquely determines the coupling (as well as the masslessness of the gauge field).

(ii) The photon does not couple to itself. There is no reason coming from gauge invariance why this does not happen (eg., involving higher powers of $F_{\mu\nu}$).

There is a 3rd key feature which we will come to below, concerning the connection between gauge transformations & conservation laws (Noether's theorem).

(ii) NON-ABELIAN GAUGE FIELDS; YANG-MILLS, etc.

In a pioneering paper published in 1954, Yang & Mills vastly extended the ideas of gauge fields beyond the $U(1)$ discussion given above. This paper, inspired by ideas from GR as well as by questions arising in particle physics, was almost entirely ignored for a decade - it was well ahead of its time. There was nothing in classical physics at that time which suggested such a quantum theory, apart from GR — although now we have very nice examples (eg., superfluid ^3He, where the quantum order parameter obeys a set of classical eqtns. of motion which look like a non-Abelian gauge theory on a background dynamic curved spacetime). The main reasons that the Yang-Mills theory received so little attention were

(i) Almost nobody was interested in or familiar with classical GR; and quantum gravity hardly existed even as an idea.

(ii) The Yang-Mills (YM) theory predicted massless particles. Adding a mass term broke the SU(N) gauge invariance. No mechanism at that time was known, at least in particle physics, that would break this invariance in a physically satisfactory way.

(iii) Nobody knew how to calculate with such theories.

As we will see, objections (i) and (ii) were slowly overcome; the key to solving (ii) was the Anderson-Higgs mechanism, implemented for YM theories by Salam & Weinberg. The key to (iii) was the use of path integral methods, without which even QED remained hard to understand. The main initiative here came from 't Hooft, and then 't Hooft & Veltman.

In what follows we will begin by going through the important special case of SU(2) gauge symmetry, which is a nice pedagogical example. I will then sketch how this is generalized to higher non-Abelian groups, and discuss the physical significance of all this.

SU(2) GAUGE THEORY

: Recall that the simple SU(2) group can be represented by the Pauli matrices $\hat{\tau}_j^{\alpha\beta}$, with

$$\left.\begin{aligned} \hat{\tau}_i \hat{\tau}_j &= \delta_{ij} + i\epsilon_{ijk}\hat{\tau}_k \\ [\hat{\tau}_i, \hat{\tau}_j] &= 2i\epsilon_{ijk}\hat{\tau}_k \end{aligned}\right\} \qquad (81)$$

and we may define the unitary operator $\hat{U}(\underline{\Omega}) \equiv \hat{U}(\hat{\underline{n}}\Omega)$, where $\hat{\underline{n}}$ is a unit vector on the Bloch sphere, as

$$\left.\begin{aligned} \hat{U}(\underline{\Omega}) = e^{-\frac{i}{2}\underline{\Omega}\cdot\underline{\tau}} &\equiv e^{-\frac{i}{2}\Omega_j \tau_j^{\alpha\beta}} \\ &= \lim_{N\to\infty}\left(1 - \tfrac{i}{2N}\underline{\Omega}\cdot\underline{\tau}\right)^N \end{aligned}\right\} \qquad (82)$$

where the last form gives us the infinitesimal operator. Note that $\hat{U}(\underline{\Omega}) \equiv \hat{U}_{\alpha\beta}(\underline{\Omega})$ is a matrix operator in "spin space", and the unit vector $\hat{\underline{n}}$ tells us the direction around which $\underline{\Omega}$ is effecting a rotation.

We may now go through much the same manoeuvres as we did for the Abelian gauge field. We introduce now a 2-component spinor Dirac field, $\psi_\alpha(x)$, and note that the Lagrangian

$$\mathcal{L}_0 = \bar{\psi}_\alpha(x)\left[i\gamma^\mu \partial_\mu \delta^\alpha_\beta - m\,\delta^\alpha_\beta\right]\psi^\beta(x)$$

is invariant under a global $SU(2)$ rotation. However let us now apply the local operator, where $\underline{\Omega} \to \underline{\Omega}(x)$; then

$$\left.\begin{aligned}\hat{U}(\underline{\Omega}(x))\,\psi(x) &\equiv U^{\alpha\beta}(\underline{\Omega}(x))\,\psi_\beta(x)\\ &= e^{-\frac{i}{2}\underline{\Omega}(x)\cdot\underline{\tau}^{\alpha\beta}}\,\psi_\beta(x) = \tilde{\psi}_\alpha(x)\end{aligned}\right\}\quad(83)$$

we find that $\mathcal{L}_0$ is not invariant, for the same reason as before, viz., we get an extra gradient term:

$$S_0 \longrightarrow S_0 + \int d^4x\,\bar{\psi}^\alpha(x)\left[U^{-1}_{\alpha\gamma}(\underline{\Omega}(x))\,\partial_\mu U^{\gamma\beta}(\underline{\Omega}(x))\right]\psi_\beta(x) \qquad (84)$$

We therefore introduce the gauge covariant derivative, with a charge g_0:

$$\left.\begin{aligned}D_\mu &= \left(\partial_\mu + i\frac{g_0}{2}\underline{\boldsymbol{\tau}}\cdot\underline{\boldsymbol{A}}_\mu\right)\\ \text{ie.}\quad D^{\alpha\beta}_\mu &= \left(\partial_\mu\delta^{\alpha\beta} + i\frac{g_0}{2}\underline{\tau}^{\alpha\beta}_j\,\underline{A}^j_\mu(x)\right)\end{aligned}\right\}\quad(85)$$

where we use a vector notation $\boldsymbol{\tau} = \underline{\tau}^{\alpha\beta}_j$ so as to suppress the clutter of spinor indices, with the boldface indicating a vector in real spacetime. Since $\bar{\psi}^\alpha(x)$ transforms to $\bar{\tilde{\psi}}_\alpha = U^{-1}_{\alpha\beta}\bar{\psi}_\beta$, we clearly want a transformation such that

$$D_\mu\psi \longrightarrow U(\underline{\Omega}(x))\,D_\mu\psi \qquad (86)$$

and, going through the algebra, analogous to that for $U(1)$ gauge fields, we find that

$$\tfrac{1}{2}\underline{\boldsymbol{\tau}}\cdot\underline{\boldsymbol{A}}_\mu \longrightarrow U(\underline{\boldsymbol{\Omega}})\,\tfrac{1}{2}\underline{\boldsymbol{\tau}}\cdot\underline{\boldsymbol{A}}_\mu\,U^{-1}(\underline{\boldsymbol{\Omega}}) + i g_0^{-1}\left(\partial_\mu U(\underline{\boldsymbol{\Omega}})\right)U^{-1}(\underline{\boldsymbol{\Omega}}) \qquad (87)$$

or, for an infinitesimal transformation $\delta\underline{\boldsymbol{\Omega}}(x)$ (cf. eqn (82)), we get

$$\left.\begin{aligned}\tfrac{1}{2}\underline{\boldsymbol{\tau}}\cdot\underline{\boldsymbol{A}}_\mu &\longrightarrow \tfrac{1}{2}\underline{\boldsymbol{\tau}}\cdot\underline{\boldsymbol{A}}_\mu + \tfrac{1}{2}\underline{\boldsymbol{\tau}}\cdot(\delta\underline{\boldsymbol{\Omega}}\times\boldsymbol{A}_\mu) + \frac{1}{2g_0}\boldsymbol{\tau}\cdot\partial_\mu\delta\underline{\boldsymbol{\Omega}}\\ \text{ie.}\quad \tfrac{1}{2}\tau_i A^i_\mu &\longrightarrow \tfrac{1}{2}\tau_i A^i_\mu + \tfrac{1}{2}\epsilon_{ijk}\tau^i\,\delta\Omega^j A^k_\mu + \frac{1}{2g_0}\tau_i\,\partial_\mu\delta\Omega^i\end{aligned}\right\}\quad(88)$$

where we write out all the components in the 2nd form. Thus we can now write everything in terms of a new non-Abelian gauge field, given by (85), with the transformation property

$$\underline{A}_\mu(x) \longrightarrow \underline{A}_\mu(x) + \delta\underline{A}_\mu(x)$$

with $$\delta\underline{A}_\mu(x) = \left(\delta\underline{\Omega}(x) \times \underline{A}_\mu(x)\right) + \frac{1}{g_0}\partial_\mu \delta\underline{\Omega}(x) \qquad (89)$$

and we can immediately generalize this from the infinitesimal transformation to the general transformation

$$\underline{A}_\mu(x) \longrightarrow \underline{A}_\mu(x) + \frac{1}{g_0}\partial_\mu \underline{\Omega}(x) + \underline{\Omega}(x) \times \underline{A}_\mu(x)$$

$$A^i_\mu(x) \longrightarrow A^i_\mu(x) + \frac{1}{g_0}\partial_\mu \Omega^i(x) + \epsilon^i{}_{jk}\,\Omega^j(x) A^k_\mu(x) \qquad (90)$$

which we should compare with the Abelian transformation in (68). We can also relate this to a "curvature tensor", or field intensity tensor; as with (72) above, we have an operator

$$R^{\alpha\beta}_{\mu\nu} = \mathcal{D}^{\alpha\gamma}_\mu \mathcal{D}^{\gamma\beta}_\nu - \mathcal{D}^{\alpha\gamma}_\nu \mathcal{D}^{\gamma\beta}_\mu \qquad (91)$$

and applying this to $\psi_\alpha(x)$, we find that

$$R^{\alpha\beta}_{\mu\nu}\psi_\beta(x) = \frac{i g_0}{2}\,\underline{\tau}^{\alpha\beta}\cdot\underline{F}_{\mu\nu}\,\psi_\beta(x) \qquad (92)$$

where we define

$$\underline{F}_{\mu\nu} = \left(\partial_\mu \underline{A}_\nu - \partial_\nu \underline{A}_\mu\right) + g\,\underline{A}_\mu \times \underline{A}_\nu$$

$$F^i_{\mu\nu} = \left(\partial_\mu A^i_\nu - \partial_\nu A^i_\mu\right) + g\,\epsilon^i{}_{jk} A^j_\mu A^k_\nu \qquad (93)$$

Unlike the Abelian $F_{\mu\nu}(x)$, this tensor is not gauge-invariant, as we see by making the transformation; we have

$$\underline{\tau}\cdot\underline{F}_{\mu\nu} \longrightarrow U(\underline{\Omega}(x))\,\underline{\tau}\cdot\underline{F}_{\mu\nu}(x)\,U^{-1}(\underline{\Omega}(x))$$

ie. $$\underline{F}_{\mu\nu}(x) \longrightarrow \underline{F}_{\mu\nu}(x) + \underline{\Omega}(x) \times \underline{F}_{\mu\nu}(x) \qquad (94)$$

However, the analogue of the Abelian field action term in (63) and (70) is gauge-invariant; ie., the term

$$\mathrm{Tr}\left\{(\underline{\tau}\cdot\underline{F}_{\mu\nu})(\underline{\tau}\cdot\underline{F}^{\mu\nu})\right\} = \tfrac{1}{2}F^i_{\mu\nu}F_i^{\mu\nu} \qquad (95)$$

is gauge-invariant. Thus we are led to the form of our spinor generalization of the Abelian gauge theory, taking the form of an action in which a vector gauge field $A_\mu(x)$ is coupled to a spinor fermion field $\psi(x)$, with the action

$$S[\bar{\psi},\psi;A_\mu] = \int d^4x\, \bar{\psi}(x)[i\gamma^\mu D_\mu - m]\psi(x) - \tfrac{1}{4}F^i_{\mu\nu}(x)F_i^{\mu\nu}(x) \qquad (96)$$

with D_μ given by (85). Now we would like to write this in a way analogous to (63), using a current operator. We can do this if we define

$$\boxed{\begin{aligned} \underline{J}^\mu(x) \equiv J_i^\mu(x) &= \frac{g}{2}\bar{\psi}(x)\gamma^\mu \underline{\tau}\,\psi(x) \\ &= \frac{g}{2}\bar{\psi}_\alpha(x)\gamma^\mu \tau_i^{\alpha\beta}\psi_\beta(x) \end{aligned}} \qquad (97)$$

A little later we will see how such a choice can be justified (and in the same way justify the choice (71) for Abelian QED). In any case, with this choice we can write the final form for the action:

$$S = \int d^4x \left\{ \bar{\psi}_\alpha(x)[i\gamma^\mu\partial_\mu - m]\psi^\alpha(x) - \underline{J}^\mu(x)\cdot\underline{A}_\mu(x) - \tfrac{1}{4}\underline{F}_{\mu\nu}(x)\underline{F}^{\mu\nu}(x)\right\} \qquad (98)$$

with the quantities defined as we have already seen.

Let us now move to a more general discussion.

<u>GENERAL NON-ABELIAN GAUGE THEORY</u> : We can formulate all of this for a general non-Abelian gauge group. Thus, one can imagine some simple Lie group G with generators $\{g_a\}$ satisfying the algebra

$$[g_a, g_b] = i f_{abc}\, g^c \qquad (99)$$

and we will represent this Lie algebra with matrices T which operate on n-dimensional fermion fields $\Psi(x) = \Psi_a(x)$. The matrices T then have the commutation relations

$$[T_a, T_b] = i f_{abc}\, T^c \qquad (100)$$

and the general local gauge transformation will act on $\Psi(x)$ according to

$$\hat{U}(\underline{\Lambda}(x))\,\underline{\Psi}(x) = e^{-ig_0 \underline{T}\cdot\underline{\Lambda}(x)}\,\underline{\Psi}(x) \qquad (101)$$

where we can think of $\underline{\Lambda}(x)$ as a "hyperangle" in the n-dimensional space. We introduce a generalized gauge field $\mathbf{A}^\mu$ with components $A^\mu_a(x)$, which transform according to

$$A^\mu_a \longrightarrow A^\mu_a + \partial^\mu\Lambda_a + g_0 f_{abc}\Lambda^b A^\mu_c \qquad (102)$$

and we define the gauge covariant derivative

$$D^\mu = (\partial^\mu + ig_0 \underline{T} \cdot \underline{A}^\mu) \tag{103}$$

Then the action of the commutator $R^{\mu\nu} = [D^\mu, D^\nu]$ on the fermion field is given by

$$R^{\mu\nu} = [D^\mu, D^\nu] = ig_0 \underline{T} \cdot \underline{F}^{\mu\nu} \tag{104}$$

with

$$\underline{F}^{\mu\nu}(x) = (\partial^\mu \underline{A}^\nu - \partial^\nu \underline{A}^\mu) + ig_0 [\underline{A}^\mu, \underline{A}^\nu] \tag{105}$$

which transforms according to

$$F^{\mu\nu}(x) \rightarrow \hat{U}(\underline{\Lambda}) F^{\mu\nu}(x) \hat{U}^{-1}(\Lambda) \tag{106}$$

Finally, we write the total action for this theory as

$$\boxed{\begin{aligned} S_{YM} &= \int d^4x \left\{ \bar{\Psi}(x) [i\gamma^\mu D_\mu - m] \Psi(x) - \tfrac{1}{4} \underline{F}_{\mu\nu}(x) \cdot \underline{F}^{\mu\nu}(x) \right\} \\ &= \int d^4x \left\{ \bar{\Psi}(x) [i\gamma^\mu \partial_\mu - m] \Psi(x) - \underline{J}^\mu(x) \cdot \underline{A}^\mu(x) - \tfrac{1}{4} \underline{F}_{\mu\nu}(x) \cdot \underline{F}^{\mu\nu}(x) \right\} \end{aligned}} \tag{107}$$

of which the SU(2) theory is obviously a special case.

(iii) PHYSICAL PROPERTIES of GAUGE FIELDS

There are many interesting things that one can say about gauge fields, particularly about non-Abelian gauge fields. In the following I will confine the remarks to some fairly basic points

EQUATIONS OF MOTION & CURRENT CONSERVATION : Up until now we've just looked at the action and the quantities in it. But an obvious physical question is – how do the fields effect each others' motion?

To answer this we find the eqtns of motion of the 2 fields, by functionally differentiating the action. We then have, as usual, that

$$\begin{aligned} \delta S_{YM} = \int d^4x \Big\{ & \Big[\partial_\mu \frac{\partial L}{\partial(\partial^\mu \underline{A}^\nu)} - \frac{\partial L}{\partial A^\nu}\Big] \delta A^\nu + \Big[\partial^\mu \frac{\partial L}{\partial(\partial^\mu \bar{\Psi})} - \frac{\partial L}{\partial \bar{\Psi}}\Big] \delta\bar{\Psi} \\ & + \Big[\partial^\mu \frac{\partial L}{\partial(\partial^\mu \Psi)} - \frac{\partial L}{\partial \Psi}\Big] \delta\Psi \Big\} \end{aligned} \tag{108}$$

giving us eqtns of motion for A^μ, $\bar{\Psi}$, and Ψ. The simplest of these eqtns is for the fermion field $\Psi(x)$; we have

$$(i\gamma^\mu D_\mu - m)\,\Psi(x) \equiv [(i\gamma^\mu \partial_\mu - m) + ig_0 \bar{\Psi}(x)\gamma^\mu \underline{T}\cdot A_\mu(x)]\,\Psi(x) = 0 \qquad (109)$$

so that the fermion field is acted upon by a "source" which combines the anti-field $\bar{\Psi}(x)$ and the gauge field term $\underline{T}\cdot A_\mu(x)$. A similar eqtn is obeyed by the anti-field $\bar{\Psi}(x)$.

The gauge field eqtn. of motion is taken from the 1st variation in (108), and it gives

$$\boxed{\begin{aligned} D_\mu \underline{F}^{\mu\nu}(x) &= g_0 \bar{\Psi}(x)\gamma^\nu \underline{T}\,\Psi(x) \\ &\equiv \underline{J}^\nu(x) \end{aligned}} \qquad (110)$$

which is just the generalization of the usual sourced eqtn of motion for the EM field.

It is very interesting and useful to look more closely at the connection between currents like $\mathbf{J}_\nu(x)$ in (110) and the symmetries that exist in the field theory of interest. Let us recall where an expression like (108) comes from; our theory has a Lagrangian $L(\mathbf{X}_p, \partial_\mu \mathbf{X}_p)$, where $\underline{\mathbf{X}}_p = (\phi(x), \psi_\alpha(x), \bar{\psi}(x), A^\nu(x), \ldots)$ is the set of all fields that L depends on, collected into one "superfield" $\underline{\mathbf{X}}_p$. We then have

$$\left.\begin{aligned} \delta S &= \int d^D x \left[\frac{\delta L}{\delta \underline{\mathbf{X}}_p}\,\delta \mathbf{X}_p + \frac{\delta L}{\delta(\partial_\mu \underline{\mathbf{X}}_p)}\,\delta(\partial_\mu \underline{\mathbf{X}}_p) \right] \\ &= \int d^D x \left[\frac{\delta L}{\delta \underline{\mathbf{X}}_p} + \frac{\delta L}{\delta(\partial_\mu \underline{\mathbf{X}}_p)}\,\partial_\mu \right] \delta \underline{\mathbf{X}}_p \end{aligned}\right\} \qquad (111)$$

Now from this we derive a form like (108) by throwing away boundary terms, arguing that at the boundary of our spacetime they vanish – this is done using integration by parts, to show that

$$\partial_\mu\left(\frac{\delta L}{\delta(\partial_\mu \underline{\mathbf{X}}_p)}\right) - \frac{\delta L}{\delta \underline{\mathbf{X}}_p} = 0 \qquad (112)$$

once we set $\delta S = 0$. However, substituting this into (111), we immediately find that δS in (111) can be written as a TOTAL derivative, ie.,

$$\delta S = \int d^D x\, \partial_\mu \left[\frac{\delta L}{\delta(\partial_\mu \underline{\mathbf{X}}_p)}\,\delta \underline{\mathbf{X}}_p \right] \equiv \int d^D x\, \delta L(\mathbf{X}_p, \partial_\mu \mathbf{X}_p) \qquad (113)$$

where we now identify the quantity in brackets as a "current". Before continuing with this argument, let's just consider what form the current $J^\mu_{(x)}$ might take. This clearly depends on what field $\underline{\mathbf{X}}_p$ we are dealing with. In what follows we will assume that all transformations of the Lagrangian we are interested in can be effected by unitary operators acting on the field $\underline{X}_p$, ie., that we can write

$$\tilde{\underline{X}}_p = e^{i\underline{G}^a_{pq}\underline{\omega}_a} \underline{X}_q \tag{114}$$

for the transformed field, where $\underline{G}^a_{pq}$ is the generator of the transformation and $\underline{\omega}^a$ is the "angle" by which it is effectuated. Then, eg., we have

$$\left.\begin{aligned} \underline{G}^a_{pq}\underline{\omega}_a &\to g_0 \underline{T}^a_{\alpha\beta}\cdot\underline{\Lambda}_a && (\text{general non-Abelian transformation}) \\ \underline{G}^a_{pq}\underline{\omega}_a &\to \frac{g}{2}\underline{\tau}^a_{\alpha\beta}\cdot\underline{\Omega}_a && (SU(2)\ \text{transformation}) \\ \underline{G}^a_{pq}\underline{\omega}_a &\to q\theta && (U(1)\ \text{transformation}) \end{aligned}\right\} \tag{115}$$

where in the case of a global transformation, $\underline{\omega}_a$ is independent of x, whereas for a local transformation, $\underline{\omega}_a \to \underline{\omega}_a(x)$.

From (114) we have

$$\left.\begin{aligned} \delta\underline{X}_p &= i\,\delta\underline{\omega}_a(x)\,\underline{G}^a_{pq}\,\underline{X}^q \\ &= \left(\frac{\partial \underline{X}_p}{\partial \underline{\omega}_a}\right)\delta\underline{\omega}_a \end{aligned}\right\} \tag{117}$$

and so now we can write δL in (113) as

$$\delta L = \partial_\mu\left[\frac{\partial L}{\partial(\partial_\mu X_p)}\ \frac{\partial \underline{X}_p}{\partial \underline{\omega}_a}\right]\delta\underline{\omega}_a \tag{118}$$

Now, the key point. Suppose we make a transformation of the fields, which will be assumed infinitesimal, parametrized by the infinitesimal $\delta\underline{\omega}_a$, and we find that L is unchanged. It then immediately follows that

$$\boxed{\begin{aligned} \partial_\mu \underline{J}^\mu_a(x) &= 0 \\ \underline{J}^\mu_a(x) &= \left[\frac{\partial L}{\partial(\partial_\mu \underline{X}_p)}\ \frac{\partial \underline{X}_p}{\partial \underline{\omega}_a}\right] \end{aligned}} \tag{119}$$

This is usually called "Noether's theorem", and evaluation of it for any of the fields we have looked at so far immediately gives us a conservation law for the currents we have defined. As an example of (119), consider the bare Lagrangian

$$L_0 = \bar{\Psi}^\alpha(x)\left[i\gamma^\mu\partial_\mu - m\right]\Psi_\alpha(x) \tag{120}$$

The infinitesimal field transformation is $\delta\Psi_\alpha(x) = \left[-ig_0\,\underline{T}^a_{\alpha\beta}\,\Psi_\beta(x)\right]\delta\underline{\Lambda}_a$ (121)

so that the current is

$$\left.\begin{aligned} \underline{J}^\mu_a(x) &= (i\bar{\Psi}^\alpha(x)\gamma^\mu)(-ig_0\,\underline{T}^a_{\alpha\beta}(x)\,\Psi^\beta(x)) \\ &= g_0\,\bar{\Psi}_\alpha(x)\gamma^\mu\,\underline{T}^{\alpha\beta}_a(x)\,\Psi_\beta(x) \end{aligned}\right\} \tag{122}$$

as previously derived in (110). From (119) we then see that for the Lagrangian in (120), $J^\mu_a(x)$ is conserved.

One can say a lot more about such conserved currents, but the basic message here is clear – symmetries lead to conservation laws, just as in classical physics and in ordinary Q.M.

PHYSICAL INTERPRETATION OF $F_{\mu\nu}(x)$

: It has already been stated that $F_{\mu\nu}(x)$ can be thought of as a kind of curvature, and here we amplify on this statement.

Consider the case where the transformation of the field we have been talking about now consists in looking at the change of the field as we move from one point to another. Thus we are interested in the transformation

$$\psi(x_2) = \hat{U}(x_2, x_1)\, \psi(x_1) \tag{123}$$

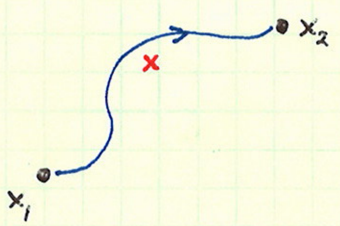
and more general in the correlator

$$G_2(x_2, x_1) = \langle 0 | \hat{T}\{\psi(x_2)\psi(x_1)\} | 0 \rangle \tag{124}$$

Now, lets consider the effect of adding a gauge field on the unitary transformation $\hat{U}(x_2, x_1)$. In path integral language, we can write an expression of the form.

$$\left.\begin{aligned} \hat{U}(x_2, x_1) &= e^{i/\hbar \int_{x_1}^{x_2} dx\, (L_0 + L_A)} \\ &= G^0_{21}[x]\, e^{i/\hbar \int_{x_1}^{x_2} dx\, L_A} \end{aligned}\right\} \tag{125}$$

where the "amplitude" $G^0_{12}[x]$ is the result of making the transformation in the absence of the gauge field, along a SPECIFIC PATH $\mathbf{x}$.



Let's now focus on the extra contribution here, which we call

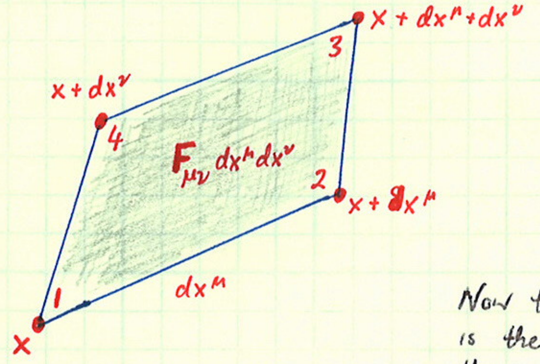
$$P_A(x_2, x_1 | \mathbf{x}) = e^{i/\hbar \int dx\, L_A(\bar{\psi}, \psi; A^\mu)} \tag{126}$$

for a general gauge field; thus, eg., for the non-Abelian Yang-Mills theory of (107), we have

$$P_A(x_2, x_1 | \mathbf{x}) = e^{i g_0 \int_{x_1}^{x_2} dx_\mu\, \underline{T} \cdot A^\mu(x)} \tag{127}$$

Now the interesting question here is – what happens if we take the system through a circuit, and bring it back to the same place? This question is a specific example of a "Berry phase" argument, which is more usually discussed for a simple wave-function, as in QM.

To extract the curvature it is sufficient to look at an infinitesimal circuit, which we assume to be oriented arbitrarily in spacetime. Let us imagine following this circuit along the counterclockwise path $1 \to 2 \to 3 \to 4$, ie., we wish to calculate the contribution



$$\left.\begin{aligned} P(x,x\,|\,4321) &= e^{ig_0 \oint dx^\mu \, \underline{T}\cdot \underline{A}_\mu(x)} \\ &= P(4,3)\,P(3,2)\,P(2,1)\,P(1,4) \end{aligned}\right\} \quad (128)$$

Now there are 2 obvious ways to do this. One is the simple and quick method of using Stokes's theorem, ie., we write

$$e^{ig_0 \oint_c dx^\mu \, \underline{T}\cdot \underline{A}_\mu(x)} = e^{ig_0 \oiint dx^\mu dx^\nu \left\{ (\partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)) + g_0 [A_\mu(x), A_\nu(x)] \right\}} \quad (129)$$

where $A_\mu(x) = \underline{T}\cdot \underline{A}_\mu(x)$. This result is obtained by noting the identity

$$e^{(\hat{A}+\hat{B})x} = e^{\hat{A}x} e^{\hat{B}x} e^{-[\hat{A},\hat{B}]x^2/2} + O(x^3) \quad (130)$$

for the non-commuting operators $A_\mu(x)$ and $A_\nu(x)$, and then using two of four integrations in the commutator term to remove the derivatives from the integrand $[\partial_\mu A^\nu, \partial_\nu A^\mu]$ in the exponent.

If this manoeuvre seems too much of a trick (and it looks much better if we phrase it in terms of differential forms), then we can do the path integral long-hand. We have

$$P(3,2)P(2,1) = e^{ig_0 A_\nu(x+dx^\mu)dx^\nu} \, e^{ig_0 A_\mu(x)\,dx^\mu}$$

$$= \exp\left\{ ig_0 \left[A_\mu dx^\mu + (A_\nu dx^\nu + \partial_\mu A_\nu \, dx^\mu dx^\nu) \right] - \frac{g_0^2}{2} [A_\mu, A_\nu] dx^\mu dx^\nu \right\}$$

and

$$P(1,4)P(4,3) = e^{-ig_0 A\,(x)\,dx} \; e^{-ig_0 A_\mu(x+\delta x^\nu)\,dx^\mu}$$

$$= \exp\left\{ ig_0 \left[A_\nu dx^\nu + (A_\mu dx^\mu + \partial_\nu A_\mu \, dx^\nu dx^\mu) \right] - \frac{g_0^2}{2} [A_\mu, A_\nu] dx^\mu dx^\nu \right\}$$

and so we get

$$P(x,x\,|\,4321) = \exp\left\{ ig_0 \left[(\partial_\mu A_\nu - \partial_\nu A_\mu) - ig_0 [A_\mu, A_\nu] \right] dx^\mu dx^\nu \right\} \quad (131)$$

From this result, which agrees with (129), we see that the net effect of moving around this path is to change the amplitude. Now we have already seen this effect in our discussion of the Aharonov-Bohm effect, for which the curvature at a point is just the magnetic field $\underline{B}(x)$ (assuming the field $A^\mu(x)$ is static). What we see in (131) is just the generalization of this to the non-Abelian case (cf. eqtn (105)).

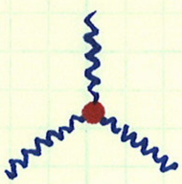
Why do we call this a <u>curvature</u>? Actually this is because of the analogy with GR, where we <u>measure</u> the curvature of spacetime by parallel transporting some 4-vector around a loop. Here we are actually transporting a field around the space of field configurations, in the Hilbert space of these configurations, by moving along a circuit in spacetime.

The effect of the commutator, and of the non-commutativity of the fields, has a profound effect on the field dynamics. Let us go back to the result we have for the total action, eqtn (107), then has terms of 3rd and 4th-order in the gauge field:

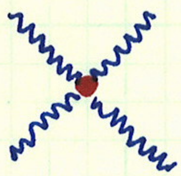
$$\left.\begin{aligned} -\tfrac{1}{4}\mathbf{F}^a_{\mu\nu}\cdot \mathbf{F}_a^{\mu\nu} &= -\tfrac{1}{2}\,\mathrm{Tr}\,\{(\underline{T}\cdot\underline{F}_{\mu\nu})(\underline{T}\cdot \mathbf{F}^{\mu\nu})\} \\ &= -g_0 f_a{}^{bc}\,\partial_\mu A^a_\nu A^\mu_b A^\nu_c - \frac{g_0^2}{4} f^{abc} f_{ade} A^\mu_b A^\nu_c A_\mu^d A_\nu^e \end{aligned}\right\} \quad (132)$$

where the f_{abc} are, as before, the elements characterizing the Lie group algebra (so for simple angular momentum-style operators, we would have $f_{abc} = \epsilon_{abc}$). Thus we immediately expect to see terms corresponding to diagram like those in (a) and (b) below left, which correspond to the 1st and 2nd self-interaction terms in (132) above. This is in addition to the free gauge line shown in (c).

(a)



(b)



(c)

From this we see a really crucial difference between $U(1)$ gauge fields like those in QED, and the non-Abelian Yang-Mills theory. If we refer back to p. 16, we see that after eqtn (80), a key feature of ordinary electrodynamics was highlighted, viz., that photons do not interact with themselves. However, Yang-Mills gauge fields have their own "self-charge", and act as a source for themselves (in addition to having matter fields as their source). The Yang-Mills gauge field is thus fundamentally NON-LINEAR. This has profound effects on the dynamics of YM fields, many of which have yet to be explored - we still do not have anything like a full understanding of them.

And of course all this takes no account of the coupling of YM fields to matter itself, which makes it all the more complicated! To properly deal with all of this would take us deep into the standard model.

Finally, notice that we have not yet <u>quantised</u> this theory! To do this we have to adapt path integrals to gauge theory, which we now do.

B.4 (c) PATH INTEGRALS FOR GAUGE FIELDS

It is perfectly possible to deal with QED using a conventional canonical approach, and the results of doing this are strewn across dozens of textbooks and thousands of papers. What is less often emphasized are certain difficulties in such an approach, which proved insurmountable in the case of non-Abelian gauge theories (at least at the time). For this reason the success of the path integral approached proved decisive for the subsequent development of particle physics.

The technical key which opened the door was the development by Fadeev & Popov in 1966-67 of their method of integrating over gauge-equivalent fields, following the discovery of "ghost" contributions by Feynman in the early 1960's. In what follows I will explain the simple picture behind the Fadeev-Popov idea, & then give its formal elaboration.

(i) FUNCTIONALLY INTEGRATING OVER REDUNDANT VARIABLES

Let's start by considering a simple problem in probability. Suppose we are given some probability $I[f]$ that a function $f(Q)$ of a variable $Q = (\{x_i\}; \{q_j\})$ will take the value I. To focus things, imagine the function I we are interested in is the intensity of sunlight in the upper atmosphere above some point on the earth's surface (itself with angular coordinates $\Omega = (\theta, \phi)$). Then $x_i \equiv (\underline{r}, t)$, where $\underline{r}$ is the radius vector of the point, measured from the centre of the earth, t is the time, and the $\{q_j\}$ are a set of atmospheric variables like pressure P, temperature T, humidity h, etc.

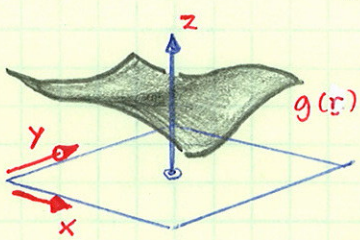
Now as physicists we understand clearly that the intensity depends only on the angle of the point on the earth with respect to the sun's direction - all other variables are irrelevant. If, however, we did not know this, we might have some trouble finding this relevant variable amongst all the others - a common problem in, eg., medical trials.

This suggests the following simple mathematical question. Suppose we are given a function

$$I = \int d^2r\, f(\underline{r}) \tag{133}$$

where $\underline{r} = (x, y)$ is a vector in the horizontal plane. We can also write this as

$$I = \int d^3R\, f(\underline{r})\, \delta(z), \quad \text{OR} \quad I = \int d^3R\, f(\underline{R})\, \delta(z - g(\underline{r})) \tag{134}$$



where in the first formula, we simply restrict the integral over the 3-d vector $\underline{R}$ to the xy-plane, whereas in the 2nd formula we extend the function $f(\underline{r})$ to the whole 3d space, assuming that

$$\left.\begin{aligned} f(\underline{R}) &= f(\underline{r}) \quad \forall\, z \\ \text{so that} \quad \frac{\partial^n f(\underline{R})}{\partial z^n} &= 0 \end{aligned}\right\} \tag{135}$$

ie., $f(\underline{R})$ is independent of z. Obviously in this case the variable z is quite irrelevant to our considerations.

Suppose however that it turns out that we don't know the explicit form of the eqtn for the surface in the form $z = g(z)$, but only know that the surface obeys the implicit eqtn

$$G(\underline{R}) = 0 \tag{136}$$

Then we would write*

$$I = \int d^3R \, f(R) \left| \frac{\partial G(\underline{R})}{\partial z} \right| \delta(G(\underline{R})) \tag{137}$$

to take care of the change of variables. All of this is a low-dimensional example of a more general problem; we have a function

$$\left.\begin{aligned} I &= \int d\underline{x}\, f(\underline{x}) & \underline{x} &= (x_1 \ldots x_n) \\ &= \int d\underline{Q}\, f(\underline{Q})\, \delta(\underline{q}) & q &= (q_1 \ldots q_m) \end{aligned}\right\} \tag{138}$$

where $\underline{Q} = (\underline{x}, \underline{q}) = (x_1 \ldots x_n ; q_1 \ldots q_m)$, and $f(\underline{Q}) = f(\underline{x}) \;\forall\, \underline{q}$. Then if we define a surface in $\underline{Q}$-space upon which $f(\underline{Q})$ varies, but it is independent of the other coordinates in $\underline{Q}$-space that are orthogonal to the surface variables, we can do the same as above. We define the surface as

$$G(\underline{Q}) = 0 \tag{139}$$

and then we have*

$$I = \int d\underline{Q}\, f(\underline{Q}) \det \left| \frac{\partial G(\underline{Q})}{\partial \underline{q}} \right| \delta(G(\underline{Q})) \tag{140}$$

What we now wish to do is apply this observation to the problem of functional integration over gauge fields, where the key problem is that instead of a function $f(\underline{Q})$ with redundant variables in it, we deal with a functional $Z[\bar{\psi}, \psi ; \underline{A}^\mu]$ of a gauge field $\underline{A}_\mu(x)$ which also has redundant variables in it, produced from some given $\underline{A}_\mu(x)$ by making a gauge transformation. The generating functional Z should be invariant under any gauge transformation, since no physical quantity should depend on which gauge we choose; and indeed we know it is invariant, because the action is invariant.

What this means is that the obvious form for the generating functional, viz.,

$$Z[\bar{\psi}, \psi ; \underline{A}_\mu] = \int D\bar{\psi} D\psi D\underline{A}_\mu \, e^{i/\hbar\, S[\bar{\psi}, \psi ; \underline{A}_\mu]} \quad ? \tag{141}$$

cannot be right, because it contains a "hidden infinity", coming from the integration over all gauge-transformed configurations of $\underline{A}_\mu$.

One might argue here that all one needs to do is fix a gauge, and then calculate from there. In the old canonical formulation, this is what was done with QED, but it led to severe technical problems, which we will note in passing below. But in the path integral formulation, all that one has to do is extract the determinant in (140) (or rather, its generalization to functionals).

* Note that for some function $f(x)$, with zeros at points $\{x_j^0\}$, $j = 1, 2, \ldots$, we have $\delta(f(x)) = |f'(x)|^{-1} \delta(x - x_j^0)$. For I in (137) we have, in the case where $G(\underline{R}) = z - g(r)$ (cf. (134)), that $\delta(G(\underline{R})) = \left|\frac{\partial G}{\partial z}\right|^{-1} \delta(z - g(r))$. The determinant in (140) is the multivariable version of this.

What happens if we do just apply (141) naively? We can do this most simply with QED, and so we are interested in the following integral:

$$I[J_\mu] = \int \mathcal{D}A_\mu \, e^{\frac{i}{\hbar}[S_0[A_\mu] + \int d^4x \, J_\mu(x) A^\mu(x)]} \tag{142}$$

where $S_0[A_\mu]$ is given by (62), which we rewrite as

$$\left.\begin{aligned} S_0[A_\mu] &= -\frac{1}{4\mu_0}\int d^4x \, F_{\mu\nu}(x) F^{\mu\nu}(x) \\ &= \frac{1}{2\mu_0}\int d^4x \, A_\mu(x)\left[\eta^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu\right] A_\nu(x) \end{aligned}\right\} \tag{143}$$

which we can also write in k-space as

$$S_0[A_\mu] = -\frac{1}{2\mu_0}\sum_q A_\mu(q)\left[q^2\eta^{\mu\nu} + q^\mu q^\nu\right] A_\nu(-q) \tag{144}$$

and you will notice the similarity of this result to that for the phonon system; we can in the same way divide this free field term into transverse and longitudinal parts, ie, write

$$A_\mu(q) = A_\nu^{\parallel}(q)\,\hat{q}_\mu \hat{q}_\nu + A_\nu^{\perp}(q)\left[\delta_{\mu\nu} - \hat{q}_\mu\hat{q}_\nu\right] \tag{145}$$

and we see that the operator $\hat{Q}_0^{\mu\nu}(q) = q^2\eta^{\mu\nu} - \hat{q}^\mu\hat{q}^\nu$, or equivalently the operator $Q_0^{\mu\nu}(x) = \eta^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu$, are entirely transverse (which is what we would expect for photons, which are indeed transverse excitations).

The next move is then clearly supposed to be to do the functional integral in (142), following the usual line:

$$\int \mathcal{D}A_\mu \, e^{\frac{i}{\hbar}\int d^4x \left[\frac{1}{2\mu_0}(A_\mu, Q_0^{\mu\nu} A_\nu) + (J_\mu A^\mu)\right]} \sim \frac{1}{|Q_0^{\mu\nu}|} e^{-\frac{i\mu_0}{2\hbar}\int d^4x \, (J Q_0^{-1} J)} \tag{146}$$

where the inverse operator $Q_0 = (\eta^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu)^{-1}$ (147)

However this operator inverse is formally infinite; writing

$$\left[\eta^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu\right]\left(Q_0(xx')\right)_{\nu\beta} = \delta^\mu_\beta \, \delta(x-x') \tag{148}$$

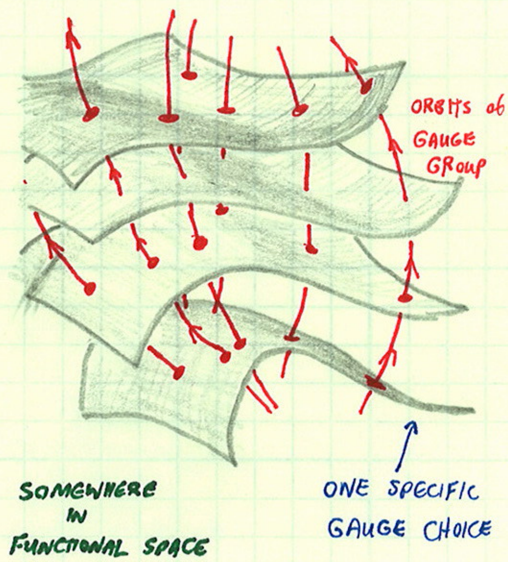
and multiplying to the left by ∂_μ, we get $\quad 0 \times Q_0 = \partial_\beta \, \delta(x-x')$ (149)

which is a contradiction unless $\hat{Q}_0$ is infinite. This infinity is a reflection of the gauge invariance, because $A^\mu(q)$ contains longitudinal (as well as transverse) degrees of freedom, which are untouched by $\hat{Q}_0$, ie., they have zero eigenvalue when acted upon by Q_0 (and hence infinite eigenvalue when operated on by Q_0^{-1}). Any gauge transformation of the form (57), adding a term $\partial^\mu\psi(x)$ to A^μ, will thus also have zero eigenvalue when acted upon Q_0, as we see:

$$Q_0^{\mu\nu}\,\partial_\mu\psi(x) = (\eta^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu)\,\partial_\mu\psi = 0 \tag{150}$$

So how do we deal with this? The naive answer is to fix the gauge before doing the calculation, but this has the disadvantage that we immediately lose Lorentz invariance in the calculations. In the early days of QFT this was a big problem, actually first solved by Tomonaga. Other methods introduced a photon mass into the calculation – this removed the problem of zero eigenvalues, but destroyed gauge invariance (the mass was set to zero at the end of the calculation). However all such techniques were extremely messy to implement, and quite impossibly complicated for non-Abelian gauge theories.

Let's describe the solution of Fadeev & Popov in geometric terms, and then go on to see how it works in detail.



Imagine the space of all possible configurations of some gauge field $A_\mu(x)$. This is of course a very large space, which we show here at left in a 3-d caricature.

Now suppose we are in some specific gauge, and we look at all possible configurations in this gauge. We show all such configurations in this gauge on a hypersheet, depicted as a simple 2-d sheet at left. We write all such configurations in this gauge as

$$A^\mu(x) \rightarrow \bar{A}^\mu(x) \quad \text{(fixed gauge)} \tag{151}$$

and then consider the set of all possible gauge transformation on $\bar{A}_\mu(x)$, to produce the full set of possible configurations, viz.,

$$\{A^\mu_a(x)\} = \{\bar{A}^\mu_a(x) + \alpha^\mu_a(x)\} \equiv \{\bar{A}^\mu_a(x) + (\partial^\mu \Lambda_a + g_0 f_{abc} \Lambda^b A^\mu_c)\} \tag{152}$$

where it is understood that this means that the set of all gauge configurations $\{A^\mu_a(x)\}$ is produced by starting with all configurations possible in some fixed gauge $\{\bar{A}^\mu_a(x)\}$, and then adding all possible gauge transformations $\{\alpha^\mu(x)\} \equiv \{\alpha^\mu_a(x)\}$, for arbitrary differentiable functions $\alpha^\mu(x)$. More precisely, we define the group G of all possible equivalence classes of a given $\bar{A}_\mu(x)$, called in mathematics the "orbit" of the gauge group; the full set of functions $A_\mu(x)$ is then produced by the set of all orbits of all possible configurations $\bar{A}_\mu(x)$ in a fixed gauge. That this geometrical picture is accurate is a job for mathematicians which we will not enter into here.

Consider now the functional integral in (142) again – we now write

$$\left.\begin{aligned} I[J_\mu] &= \int \mathcal{D}A_\mu \, e^{\frac{i}{\hbar}[S_0[A_\mu] + \int dx\, J_\mu(x) A^\mu(x)]} \\ &\equiv \int \mathcal{D}\alpha_\mu(x) \int \mathcal{D}\bar{A}_\mu(x)\, e^{\frac{i}{\hbar}(S_0[\bar{A}_\mu + \alpha_\mu(x)] + \int dx\, J_\mu(x)(\bar{A}^\mu(x) + \alpha^\mu(x)))} \\ &\equiv \int \mathcal{D}\alpha_\mu \int \mathcal{D}\bar{A}^\mu(x)\, e^{\frac{i}{\hbar}(S_0[A^\mu] + \int J_\mu A^\mu)} \end{aligned}\right\} \tag{153}$$

and we see that the problem in (153) is that the functional integrand, ie., $\exp\{\frac{i}{\hbar}(S_0[A^\mu] + \int J_\mu A^\mu)\}$, is invariant under changes in gauge, so that the functional integration $\int \mathcal{D}\alpha_\mu$ simply produces an infinite multiplication of the answer - what we would like to do is get rid of this.

To do so, let's recall the development in eqtns (138)-(140), and do the same now for gauge functionals, instead of ordinary functions. Notice that we can't just stick a factor like $\delta(\chi(x))$ into the functional integral; this is for 2 reasons. First, we want to keep the theory gauge invariant, for many different reasons (most of which are not yet clear). Second, in the functional integrals (142) and (153), we don't actually yet know how to properly "measure" the "volume", in the hyperspace of functionals, of the domain defined by a specific gauge choice. To see this we simply recall the determinant appearing in the finite-dimensional integral in (137), which acts as a Jacobian for the change of variables.

Thus it is not enough to just stick a factor $\int \mathcal{D}\alpha^\mu \, \delta(G(A^\mu(x)))$ into the functional integral (141). What we want is something like

$$Z[\bar\psi,\psi,A_\mu] = \int \mathcal{D}\alpha_\mu \int \mathcal{D}A_\mu \, \Delta_{FP} \, \delta(G(A_\mu)) \int \mathcal{D}\bar\psi \, \mathcal{D}\psi \, e^{\frac{i}{\hbar} S[\bar\psi,\psi,A_\mu]} \qquad (154)$$

where Δ_{FP} is the relevant determinant, now called the "Fadeev-Popov determinant". Our job is to find an expression for it, which is valid for any conceivable form for the gauge transformation.

Formally this is easy. In exact analogy with (140), we have

$$\left.\begin{aligned} \Delta_{FP} &= \det \left| \frac{\delta G(A_\mu(x))}{\delta \alpha_\mu(x')} \right| \equiv \Delta^{ab}_{FP}(x,x') \\ \text{for} \quad G(A_\mu(x)) &= 0 \qquad \text{(gauge constraint)} \end{aligned}\right\} \qquad (155)$$

However this formal expression is not terribly illuminating until we try to use it. Let's first rewrite (155) for a general non-Abelian gauge theory, and then say a little more about it for QED and for Yang-Mills theories.

If we assume that the gauge transformation can always be characterized by an "angle" in some space of gauge transformations (this angle being $\theta(x)$ for $U(1)$ transformation and $\Lambda(x) \equiv \Lambda_a(x)$ for YM theories), then we can rewrite the Fadeev-Popov factor in (154) as

$$\int \mathcal{D}\alpha_\mu \, \Delta_{FP} \, \delta(G(A_\mu(x))) \equiv \int \mathcal{D}\Lambda(x) \, \Delta_{FP} \, \delta(G(A_\mu(x))) \qquad (156)$$

where now

$$\boxed{\Delta_{FP} \equiv \Delta^{ab}_{FP}(x,x') = \det \left| \frac{\delta G(A^a_\mu(x))}{\delta \Lambda_b(x')} \right| = \det \left| M^{ab}(x,x') \right|} \qquad (157)$$

This is a much more transparent formula since the functional integral is over angles $\Lambda^a(x)$ in some angular space, something we know how to do. Note that the Fadeev-Popov determinant has a physical meaning that is evident from (157). If we make an infinitesimal change in the gauge by effecting a change $\delta\Lambda(x)$ in the gauge angle, then Δ_{FP} measures the "response" to this from the gauge-fixing function $G(A_\mu(x))$. In other words, if we write

$$\boxed{\delta G_a(A^\mu(x)) = \int d^4x' \, M^{ab}(x,x') \, \delta\Lambda_b(x')} \tag{158}$$

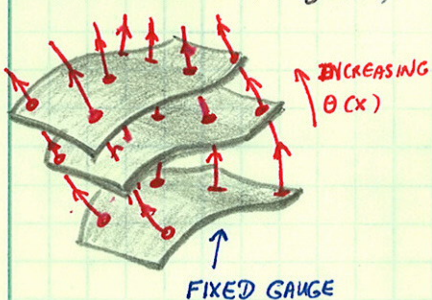
then $M^{ab}(x,x')$ measures the response of the a-th component of the gauge constraint function $G_a \equiv G(A^\mu_a(x))$ to the change $\delta\Lambda_b(x')$ in the gauge angle.

(ii) EXAMPLES of PATH INTEGRATION for GAUGE THEORIES

We will not go into too much detail here. The case of QED is relatively easy to understand, as we will see. Non-Abelian gauge theories are more messy because the number of degrees of freedom is large. It would take us too far afield to discuss the application of the results to either QED or to the electroweak theory (and the latter requires, in any case, an appeal to a spontaneous symmetry-breaking, ie., to the Anderson-Higgs mechanism).

QUANTUM ELECTRODYNAMICS :

This is of course the simplest theory. We have parametrized gauge transformations in this theory by the shift $A^\mu(x) \to A^\mu(x) + \partial^\mu \chi(x)$, but it is now time to formulate this a little more generally. We assume the gauge constraint in the form



$$G(A^\mu(x,\theta(x)) = 0 \qquad \left\{ \begin{array}{lll} G(A^\mu(x)) = \partial_\mu A^\mu(x) & & \text{(Lorentz)} \\ \quad '' \quad = \nabla \cdot A(x) & & \text{(Coulomb)} \\ \quad '' \quad = A_z(x) & & \text{(Axial)} \end{array} \right\} \tag{159}$$

where on the RHS of (159) we give 3 common examples of fixed gauge choices in QED. In what follows we will use a slightly on the Lorentz gauge, so as to use a trick devised by 't Hooft for YM theories. We write

$$\left.\begin{aligned} G(A^\mu) &= \partial^\mu A_\mu(x) - \chi(x) \\ &= (\partial^\mu \bar{A}_\mu(x) + \partial^2\theta(x)) - \chi(x) \end{aligned}\right\} \tag{160}$$

where $\bar{A}(x)$ is some fixed gauge satisfying $G(A^\mu) = 0$. Then we have*

$$\int \mathcal{D}\theta(x) \, \delta(G(A^\mu)) = (\det|\partial^2|)^{-1} \tag{161}$$

* Recall the footnote on p. 28. We are just using the functional generalization of the usual formula that $\int dx \, \delta(ax-b) = 1/a$ to $\int \mathcal{D}\theta(x) \, \delta(\partial^2\theta(x) - f(x)) = 1/\det(\partial^2)$.

so that the FP determinant is

$$\left.\begin{aligned}\Delta_{FP}(x,x') &= \det\left|(\partial^\mu\partial_\mu)_{x,x'}\right| = \det|\partial^2| \\ &= \Delta_{FP}\end{aligned}\right\} \quad (162)$$

which is a constant; thus we get

$$\left.\begin{aligned}I &= \int \mathcal{D}A^\mu(x)\, e^{i/\hbar S[A^\mu]} \\ &= \Delta_{FP}\int \mathcal{D}A^\mu(x)\,\delta(\partial^\mu A_\mu - \chi(x))\, e^{i/\hbar S[A]}\end{aligned}\right\} \quad (163)$$

with the constant outside the integration. However we still have to deal with the δ-functional $\delta(\partial^\mu A - \chi)$, and this is where the 't Hooft trick comes in handy. Suppose we functionally integrate now over $\chi(x)$, but now inserting some function $H_t(x)$ in the integrand, ie., we multipy up I with the factor $\int \mathcal{D}\chi(x)\, H_t(x)$. (where we if we choose $H_t = 1$, then we get rid of the χ-function as though it had never been there at all). We then have

$$I = \frac{1}{N}\int \mathcal{D}A^\mu(x)\, H_t(\partial^\mu A_\mu)\, e^{i/\hbar S[A]} \quad (164)$$

where N, the normalizing factor, is just a constant: $N = \dfrac{\int \mathcal{D}\chi\, H_t(\chi)}{\Delta_{FP}}$ (165)

The nice thing about this trick is that we can now make $H_t(\chi)$ the exponential of something — this way everything is now in the exponential, and we can read off the Feynman rules. The choice made by 't Hooft was

$$H_t(\chi) = e^{\frac{-i}{2\alpha}\int d^4x\, \frac{1}{\mu_0\hbar}\chi^2(x)} \quad (166)$$

where α is just a number; we then finally get (ignoring the factor $1/N$):

$$Z_{QED}[J_\mu] = \int \mathcal{D}\bar\psi\,\mathcal{D}\psi \int \mathcal{D}A^\mu\, e^{i/\hbar\left(S_0[A] - \frac{1}{2\alpha}\int d^4x\,\frac{1}{\mu_0}(\partial_\mu A^\mu)^2 + \int d^4x\, J_\mu(x) A^\mu(x)\right)} \quad (167)$$

with $S_0[A^\mu]$ given by (143) or (144). We may now do the usual functional integration that led to (146) and (147), but now we can write

$$Z_{QED}[J_\mu] = \int \mathcal{D}\bar\psi\,\mathcal{D}\psi\int \mathcal{D}A^\mu\, e^{-i/\hbar\int d^4x\left[\frac{1}{2\mu_0}A^\mu(x) D^0_{\mu\nu}(xx') A^\nu(x')\right] - \int d^4x\, J_\mu(x)A^\mu(x)} \quad (168)$$

where

$$D^0_{\mu\nu}(x,x') = \left(\eta_{\mu\nu}\partial^2 + (1/\alpha - 1)\partial_\mu\partial_\nu\right)^{-1} \quad (169)$$

and this operator does not have the pathology of $\Phi_0(x,x')$ in (147); it has an equivalent representation in momentum space as

$$\boxed{D^0_{\mu\nu}(q) = \frac{-1}{q^2+i\epsilon}\left[\eta_{\mu\nu} + (\alpha-1)\hat{q}_\mu \hat{q}_\nu\right]} \qquad (170)$$

where, as usual, $\hat{q}^\mu \equiv q^\mu/|q|$. The parameter α now acts as a "regularizer", getting rid of the "zero mode" problem we had before. Different values of α give different gauges that had been used in QED long before Fadeev & Popov; for example

$\alpha = 1$ Feynman gauge
$\alpha = 0$ Landau gauge.

We may now read off the Feynman rules for QED (compare the rules already derived for Dirac fermions in section B.3, eqtns. (44) - (48)). We have, for the Lagrangian

$$L = \bar{\psi}_\alpha(x)(i\gamma^\mu D_\mu - m)\psi_\alpha(x) + J_\mu(x)A^\mu(x) - \frac{1}{2\alpha\mu_0}(\partial_\mu A^\mu)^2 - \frac{1}{2\mu_0}F_{\mu\nu}F^{\mu\nu} \qquad (171)$$

the following rules:

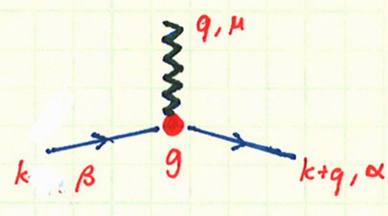
1. The fermion propagator is $G_0^{\alpha\beta}(k_j) = i\hbar S_F^{\alpha\beta}(k_j)$, for a fermion line carrying momentum k_j (172)

2. The photon propagator is given by the factor

$$i\hbar D^0_{\mu\nu}(q) = \frac{-i\hbar}{q^2+i\epsilon}\left[\eta_{\mu\nu} + (\alpha-1)\hat{q}_\mu\hat{q}_\nu\right] \qquad (173)$$

3. Each interaction vertex contributes the factor

$$\frac{-i}{\hbar}g = \frac{-i}{\hbar}q\gamma^\mu_{\alpha\beta}\;\delta\left(\sum_i k_i + \sum_j q_j\right) \qquad (174)$$



where the γ-matrix now contains spin indices α, β acting between $\bar{\psi}_\alpha$ and ψ_β.

Then, as usual, we sum over all spinor indices, integrate over all momenta k_i and q_j, and multiply by a symmetry factor; and the momentum integrations appear in the form

$$(-1)^L \int\frac{dk_1}{(2\pi)^D}\cdots\int\frac{dk_n}{(2\pi)^D}\;\int\frac{dq_1}{(2\pi)^D}\cdots\int\frac{dq_m}{(2\pi)^D} \qquad (175)$$

where L is the number of fermion loops.

Actually, the rules for QED are almost identical in form to those for the electron-phonon problem, covered in section B.3; and the topology of the diagrams is identical to that for the coupled field problem of section B.3.

YANG-MILLS SU(N) THEORY :

Let us go back to our key formula in (157), for the FP determinant. To make things clear, we will simply go through the same manoeuvres as we did for the U(1) gauge field, and we will pick the same generalized Lorentz gauge, viz.,

$$
\left.\begin{aligned}
G(A_a^\mu(x)) &= \partial_\mu A_a^\mu - \chi_a(x) \\
&\equiv \partial_\mu \bar{A}_a^\mu + \partial_\mu(\partial^\mu \Lambda_a + g_0 f_{abc} \Lambda^b \bar{A}_c^\mu) - \chi_a(x)
\end{aligned}\right\} \quad (176)
$$

Following through the steps as for the QED calculation, we then find that

$$
\boxed{\Delta_{FP}^{ab}(x,x') = \det \left| \delta^{ab} \partial^2 + g_0 f^{abc} \partial_\mu A_c^\mu \right| \delta(x-x')} \qquad (177)
$$

We may also carry out the integration using the 't Hooft trick, introducing the obvious generalization of (166) as

$$
H_t(\chi_a) = \exp \frac{-i}{2\alpha} \int d^4x \left(\frac{1}{\hbar} \chi_a(x) \chi^a(x) \right) \qquad (178)
$$

and so we can write the partition function/generating functional as

$$
\left.\begin{aligned}
Z_{YM}[J_\mu^a] &= \int \mathcal{D}\mathbf{A}^\mu(x)\, \mathbf{\Delta}_{FP}\, e^{i/\hbar \left(S_{YM}^0[\mathbf{A}] + \int d^4x [\mathbf{J}_\mu(x) \mathbf{A}^\mu(x) - \frac{1}{2\alpha}(\partial_\mu A_b^\mu)^2] \right)} \\
&\equiv \int \mathcal{D}A_a^\mu(x)\, \Delta_{FP}^{ab}(x)\, e^{\frac{i}{\hbar} \left(S_{YM}^0[A_b^\mu(x)] + \int d^4x [J_\mu^b(x) A_b^\mu(x) - \frac{1}{2\alpha}(\partial_\mu A_b^\mu)^2] \right)}
\end{aligned}\right\} \quad (179)
$$

with $S_{YM}^0[\mathbf{A}^\mu]$ given by:

$$
S_{YM}^0[\mathbf{A}^\mu] = -\frac{1}{4} \int d^4x\, \mathbf{F}_{\mu\nu}(x) \mathbf{F}^{\mu\nu}(x) \qquad (180)
$$

However there is now a big difference. In the case of QED, the Fadeev-Popov determinant Δ_{FP} in (162) was just a constant, independent of the field $A^\mu(x)$. This is no longer the case – the determinant in (177) is clearly dependent on $\mathbf{A}^\mu(x)$, and so we cannot take it outside the functional integral in (179), as we did for QED in (163). This makes the non-Abelian case much more difficult.

At this point the founders of modern QFT introduced a trick that had been invented by Feynman, during his earlier research into quantum gravity. The wrote the determinant Δ_{FP}^{ab} as a functional integration over a set of fake fermion fields – this takes us back to the result we found in section B.2 for integration over fermion fields, viz., that they give a determinant in the numerator (cf. eqtns (25) and (28) in section B.2). Thus we write

$$
\boxed{\begin{aligned}
\Delta_{FP}^{ab}(x,x') &= \det |M^{ab}(x,x')| = \int \mathcal{D}\bar{\mathbf{c}}(x) \int \mathcal{D}\mathbf{c}(x)\, e^{i S_{GH}[\bar{\mathbf{c}},\mathbf{c}]} \\
S_{GH}[\bar{c},c] &= \int d^4x \int d^4x'\, \bar{c}_a(x) M^{ab}(x,x') c_b(x)
\end{aligned}} \quad (181)
$$

where the fermion fields $\mathbf{c}(x) \equiv c^a(x)$ and $\bar{\mathbf{c}}(x)$ obey all the Grassman rules discussed before. The fields were called "ghost fields" by Feynman, and they created much confusion at the time (Feynman introduced them to prevent the theory from losing unitarity – it was only later that their determinental role was understood).

Thus we can finally write that

$$\boxed{Z_{YM}[J^a_\mu] = \int \mathcal{D}\bar{\mathbf{c}}\, \mathcal{D}\mathbf{c} \int \mathcal{D}A^\mu\, e^{\frac{i}{\hbar}\left(S^0_{YM}[A^\mu] + S_{GH}[\bar{\mathbf{c}},\mathbf{c}] + \int d^4x \left[J_\mu(x) A^\mu(x) - \frac{1}{2\alpha}(\partial_\mu A^\mu)^2\right]\right)}} \tag{182}$$

The Feynman rules for this functional are extracted by writing the action in terms of a non-interacting "free field" part

$$S_0[\bar{\mathbf{c}},\mathbf{c};A^\mu] = \int d^4x \left[\bar{\mathbf{c}}(x)\partial^2\mathbf{c}(x) - \tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\alpha}(\partial_\mu A^\mu)^2\right] \tag{183}$$

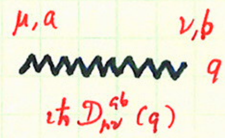
and an interacting part

$$\left.\begin{aligned} S_{int}[\bar{\mathbf{c}},\mathbf{c};A^\mu] &= -\int d^4x \Big\{ i g_0 (\bar{\mathbf{c}}\,\mathbf{f}\,\mathbf{c}\,\partial_\mu A^\mu) + \tfrac{1}{2} g_0 F_{\mu\nu}\,\mathbf{f}\, A^\mu A^\nu \\ &\qquad - \tfrac{1}{4} g_0^2\, \mathbf{f}\mathbf{f} (A_\mu A^\mu)^2 \Big\} \\ &\equiv -\int d^4x \Big\{ i g_0 (\bar{c}_a f^{abc} c_b\, \partial_\mu A^\mu_c) + \tfrac{1}{2} g_0 F^a_{\mu\nu} f_a{}^{bc} A^\mu_b A^\nu_c \\ &\qquad - \tfrac{1}{4} g_0^2 f^a{}_{bc} f_a{}^{de} A^b_\mu A^c_\nu A^\mu_d A^\nu_e \Big\} \end{aligned}\right\} \tag{184}$$

and this allows us to read off the Feynman rules for the Yang-Mills gauge field (which, I emphasize, is still not coupled to the real world fermions that one might expect to exist in a real theory like the electroweak theory). We then have.

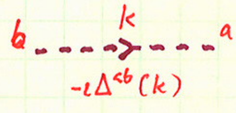
1. A vector boson propagator given by the YM generalization of (173), viz.,

$$i\hbar D^{ab}_{\mu\nu}(q) = \frac{-i\hbar}{q^2 + i\epsilon}\,\delta^{ab}\left[\eta_{\mu\nu} + (\alpha - 1)\hat{q}^\mu \hat{q}^\nu\right] \tag{185}$$



2. A ghost propagator given by

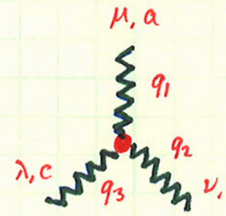
$$-i\Delta^{ab}(k) = -i\delta^{ab}\,\frac{1}{k^2 + i\epsilon} \tag{186}$$



where there is no $\hbar$ because of the def^n (181); we notice that the ghost propagator is massless, as is obvious from (184), and is a scalar field as well

3. Finally, from the interaction in (184) we get a whole variety of vertices, as follows:

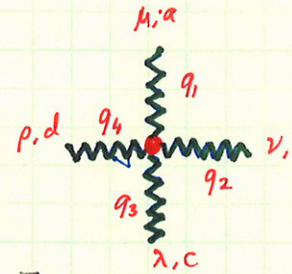
(a) 3-boson vertex, given by the 2nd term in (184), as.



$$\frac{i}{\hbar}\lambda_3 = \frac{i}{\hbar} g_0 f^{abc}\left[\eta_{\mu\nu}(q_1-q_2)_\lambda + \eta_{\nu\lambda}(q_2-q_3)_\mu + \eta_{\lambda\mu}(q_3-q_1)_\nu\right]\delta(q_1+q_2+q_3)$$

in which 3 vector bosons $A^\mu(q)$ interact at a point in spacetime; scattering off "curvature";

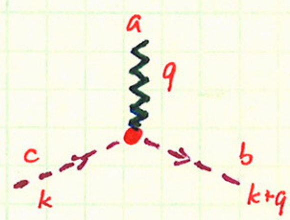
(b) A 4-boson vertex, given by the 3rd term in (184), as



$$\frac{i}{\hbar}\lambda_4 = \frac{-i}{\hbar} g_0^2\left[f^{ab}{}_e f^{cd}{}_e\,(\eta_{\mu\lambda}\eta_{\nu\rho} - \eta_{\mu\rho}\eta_{\nu\lambda}) + f^{ac}{}_e f^{bd}{}_e\,(\eta_{\mu\nu}\eta_{\lambda\rho} - \eta_{\mu\rho}\eta_{\lambda\nu}) + f^{ad}{}_e f^{cb}{}_e\,(\eta_{\mu\lambda}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\lambda\rho})\right]\delta(q_1+q_2+q_3+q_4)$$

with 4 bosons coupled at a point; and

(c) A boson-ghost fermion interaction, coming from the 1st term in (184), given by



$$\frac{i}{\hbar} g_3 = g_0 f^{abc} q_\mu$$

which is produced by the absorption of a gauge boson by the ghost fermion.

Finally, we integrate over momenta as before. - however, there are no external ghost fermion lines, and we still have a loop factor $(-1)^L$ for L ghost fermion loops.

This concludes this introduction to gauge fields in QFT. One of the most interesting topics that is usefully examined in this formalism, outside of particle physics, is the variety of non-relativistic fermionic & bosonic superfluids in Nature, as well as spin fluids. And of course, if we go to spin-2 gauge fields, we can also discuss quantum gravity.