NBI-HE-76-14

OF GHOULIES AND GHOSTIES

AN INTRODUCTION TO QCD

B. Lautrup

The Niels Bohr Institute, University of Copenhagen

DK-2100 Copenhagen Ø, Denmark

From Ghoulies and Ghosties and
longleggedy Beasties and Things
that go bump in the Night
good Lord preserve us

Old English prayer

Lectures given at the International School of Elementary Particle Physics, Baŝko Polje, 1976

Contents

1. Introduction

In these lectures I have tried to explain the structure of gauge theories from simple unitarity arguments starting with quantum electrodynamics. There is nothing new in these lectures except perhaps certain pedagogical means of expression. They are intended for a group of young experimentalists, but can also serve as the first introduction to these concepts for young theorists. They assume nothing more than a superficial acquaintance with QED and willingness to generalize. I have completely avoided the use of fields and kept the language entirely within the terminology of particle physics. On the other hand I have also tried to be as precise as possible in the calculations that are actually carried out.

It is not very difficult to introduce fields and I have at several points been strongly tempted to do so. But from a pedagogical point of view the fields tend to obscure the arguments by being far away from the natural language of particle physics. In particular the unitary arguments that are so crucial for the understanding of gauge theories do not seem natural in the language of field theory. For example, the Feynman-Fadeev-Popov ghosts are introduced in section 6 after a study of unitarity in quark-antiquark annihilation into two gluons. Formally they arise in field theory from the logarithm of a certain functional determinant and their connection with unitarity is obscure. There are, of course, things that cannot be done easily in such a simple language as I have adopted here. The proof of generalized Ward identities and the breakdown of symmetry are best done in the language of the functional formalism and path integrals. Likewise the more exotic aspects of gauge fields, magnetic monopoles, solitons etc. are also bound to the field formalism. For this reason these subjects have all been avoided here.

The lectures are organized in the following way. In section 2 QED in the Coulomb gauge is reviewed and the photon propagator is calculated. In section 3 this theory is cast into covariant form and it is recognized that this can only be done if two unphysical photons are introduced, one of which is produced with negative probability. It is also seen that in order that negative cross sections do not arise a certain linear superposition of unphysical photons (the gaugeon) must decouple from the S-matrix. In section 4 we generalize QED to quantum chromodynamics and demand that the gaugeons decouple. This leads in quark-gluon scattering to the result that the couplings belong to a Lie algebra. In section 5 the general structure of compact Lie algebras is analyzed and the section finishes with a discussion of the Cartan classification. In section 6 we study quark-antiquark annihilation into two gluons and discover that although the gaugeons do decouple there is still an extra piece in the covariant probability. In order to remove this piece it is necessary to introduce still more unphysical scalar particles with negative pair production cross section. These are the Fermion ghosts. In section 7 we look at gluon-gluon scattering and recognize the need for a four-gluon vertex, and in section 8 we show how the gaugeons decouple in a simple loop calculation. In section 9 we show that the theory so constructed is renormalizable by power counting. It is then shown - in section 10 - that a photon mass can be introduced with impunity in QED, but in QCD things go badly wrong as shown in section 11, which also has a short discussion of spontaneous breakdown of symmetry. In section 12 we return to the discussion of renormalization and in section 13 we discuss the Callan-Symanzik equations. The last two sections are mainly intended for theorists.

At no place in these lectures do I construct specific models. There is ample literature on this subject and the lectures by M. Gaillard at this school will cover this point.

Also I do not go into a deeper discussion of spontaneously broken symmetries, which will be covered by Zinn-Justin.

In writing these lectures I have profited much from discussions with Predrag Čvitanović to whom I direct my thanks. I have lifted unashamedly any material I needed from the vast literature on the subject. There are only a few references at the end of these notes and I apologize for the many sins of omission that I have wittingly - and unwittingly - committed.

2. QED in the Coulomb gauge

The basic object in quantum electrodynamics is the photon. Experimentalists treat the photons at equal footing with other particles. At all the big accelerators secondary photon beams are available. High energy experimentalists rarely think of the wave nature of the photons they use. Conversely opticians rarely think of their particle nature. There is, incidentally, one experiment in which optics meets with particle physics. If you focus the light from a ruby laser on a stored high energy electron beam, tremendous head-on collisions will occur between the photons and the electrons, and the photons may get their energy violently increased.

For the theoreticians the dual nature of particles and waves is always apparent. In the language of Feynman all processes are broken down into basic processes, absorption and emission, for which certain simple amplitudes are given. The amplitude that a photon comes in with momentum $\bar k$ to get absorbed at the point $\bar x$ at time t is represented by

$$\underset{\bar x, t}{\bullet}\!\sim\!\sim\!\sim_{\bar\varepsilon,\bar k} = \varepsilon_i\, e^{i\bar k\cdot\bar x - i\omega t} \tag{2.1}$$

where ω is its energy ($\hbar = c = 1$ everywhere). Since photons are massless we have $\omega = |\bar k|$. This is nothing but a classical plane wave with polarization vector $\bar\varepsilon$. It is well-known that plane electromagnetic waves have polarization orthogonal to their motion (in the Coulomb gauge), i.e.

$$\bar k\cdot\bar\varepsilon = 0 \tag{2.2}$$

From the correspondence principle we conclude that this must also be true quantum mechanically.

Similarly, the complex conjugate of (1)

$$\bar k,\bar\varepsilon\sim\!\sim\!\sim\underset{\bar x, t}{\bullet} = \varepsilon_i^{*}\, e^{-i\bar k\cdot\bar x + i\omega t} \tag{2.3}$$

represents the amplitude that a photon comes out with momentum $\bar k$ after having been emitted at the point $\bar x, t$.

The most general solution to (2) is

$$\bar\varepsilon = c_1\,\bar\varepsilon^1 + c_2\,\bar\varepsilon^2 \tag{2.4}$$

where c_1 and c_2 are complex numbers, and $\bar\varepsilon^1$ and $\bar\varepsilon^2$ are real vectors which may be chosen to satisfy

$$\bar\varepsilon^1\cdot\bar k = \bar\varepsilon^2\cdot\bar k = \bar\varepsilon^1\cdot\bar\varepsilon^2 = 0 \tag{2.5}$$

$$|\bar\varepsilon^1| = |\bar\varepsilon^2| = 1 \tag{2.6}$$

The raised index $\lambda = 1,2$ may be thought of as the quantum number of polarization.

Imagine now a collision process in which - among other things- a single photon is emitted. The amplitude must be of the form

$$\begin{aligned}\bar k,\bar\varepsilon\sim\!\sim\!\sim\!\oslash &= \int d^3\bar x\, dt\; \bar\varepsilon^{*} e^{i\omega t - i\bar k\cdot\bar x}\,(-i)\,\mathfrak{J}(\bar x,t) \\ &= \bar\varepsilon^{*}\cdot(-i)\,\mathfrak{J}(\bar k,\omega)\end{aligned} \tag{2.7}$$

where the blob represents the collision. Mathematically, the amplitude that the emission occurs at $\bar x, t$ is a certain function $(-i)\mathfrak{J}(\bar x,t)$ and (7) arises when we integrate the product of (3) and $(-i)\mathfrak{J}(\bar x,t)$ over all $\bar x, t$. $\mathfrak{J}(\bar k,\omega)$ is the four-dimensional Fourier transform of $\mathfrak{J}(\bar x,t)$. The factor $(-i)$ is for later convenience.

In accordance with the general principles of quantum mechanics, the probability is the square of the amplitude, $|\bar\varepsilon^{*}\cdot\mathfrak{J}(\bar k,\omega)|^2$. If we do not measure the polarization of the emerging photon we must sum over the quantum number of polarization to get

$$P_T = |\bar{\varepsilon}^1 \cdot \bar{J}|^2 + |\bar{\varepsilon}^2 \cdot \bar{J}|^2 = \bar{J}^2 - \left(\frac{\bar{k}\cdot\bar{J}}{\omega}\right)^2 \tag{2.8}$$

where T stands for transverse. The last equation follows from the completeness of the basis formed by $\bar{\varepsilon}^1, \bar{\varepsilon}^2$ and $\bar{k}/\omega$. This may also be expressed in the form

$$\varepsilon^1_i\,\varepsilon^1_j + \varepsilon^2_i\,\varepsilon^2_j = \delta_{ij} - \frac{k_i k_j}{\omega^2} \equiv \tau_{ij} \tag{2.9}$$

where we have used tensor notation. The tensor τ_{ij} is called the transverse projection tensor.

From classical radiation theory we know that a changing electromagnetic current emits and absorbs radiation. In the correspondence limit a current must then also emit photons and by more detailed arguments one can show that $\bar{J}(\bar{x}, t)$ is really a kind of current. Imagine now that the photon emitted by one current is absorbed by another. The amplitude for this process must be of the form

$$A\ \underset{i\quad T\quad j}{\overset{x\qquad\ y}{\rightsquigarrow}}\ B = \int d^4x \int d^4y\,(-i)J^A_i(x)\,D_{ij}(x-y)\,(-i)J^B_j(y) \tag{2.10}$$

where we use four-dimensional notation for the integrals over space-time. The kernel $D_{ij}(x-y)$ represents the propagation of the photon between A and B and is called the propagator for transverse photons. It must be given by the expression

$$\underset{i\quad T\quad j}{\overset{x\qquad\ y}{\rightsquigarrow}} \equiv D_{ij}(x-y) = \Theta(x_0-y_0)\sum_{\bar{k}\lambda} \underset{i\quad \bar{k}\lambda\quad j}{\overset{x\qquad\ y}{\rightsquigarrow\rightsquigarrow}} + \Theta(y_0-x_0)\sum_{\bar{k}\lambda} \underset{j\quad \bar{k}\lambda\quad i}{\overset{y\qquad\ x}{\rightsquigarrow\rightsquigarrow}} \tag{2.11}$$

which says that if $x_0 > y_0$ then the photon is emitted from y and absorbed at x passing through every intermediate state $\bar{k}\lambda$. When $x_0 < y_0$ the roles of emitter and absorber are reversed. The sum over λ can be carried out using (9) and the sum over $\bar{k}$ is replaced by the integral $\int \frac{d^3\bar{k}}{(2\pi)^3 2\omega}$. Then we get

$$\underset{i\quad T\quad j}{\overset{x\qquad\ y}{\rightsquigarrow}} = \Theta(x_0-y_0)\int\frac{d^3\bar{k}}{(2\pi)^3 2\omega}\,\tau_{ij}\,e^{-i\omega(x_0-y_0)+i\bar{k}\cdot(\bar{x}-\bar{y})} + \Theta(y_0-x_0)\int\frac{d^3\bar{k}}{(2\pi)^3 2\omega}\,\tau_{ij}\,e^{i\omega(x_0-y_0)-i\bar{k}\cdot(\bar{x}-\bar{y})} \tag{2.12}$$

Using the Fourier representation of the Θ-function

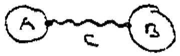
$$\Theta(t) = \int_{-\infty}^{\infty}\frac{du}{2\pi}\,\frac{i}{u+i0}\,e^{-iut}$$

we may write this in the form

$$\underset{i\quad T\quad j}{\overset{x\qquad\ y}{\rightsquigarrow}} = \int\frac{d^4k}{(2\pi)^4}\,\frac{i\tau_{ij}}{k^2+i0}\,e^{-ik\cdot(x-y)} \tag{2.13}$$

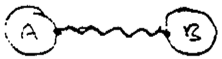
where we have used four-dimensional notation throughout.[1] The integral over k_0 may be viewed as the sum over all the transients created by the abrupt switch of roles for emitter and absorber at $x_0 = y_0$. This equation shows that a photon can propagate over finite distances without fulfilling the mass-shell condition $k^2 = 0$. This may also be interpreted as a consequence of the uncertainty relations. The small imaginary part specifies the way the singularity at $k^2 = 0$ is passed in the k_0 - integration.

Charges also interact via their instantaneous Coulomb interaction. From elementary quantum mechanics we know that in the first approximation the Coulomb transition amplitude is time integral of the interaction energy



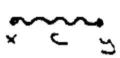
$$= -i\int_{-\infty}^{\infty} dt \int d^3x \int d^3y\ \rho^A(\vec{x},t)\,\frac{1}{4\pi|\vec{x}-\vec{y}|}\,\rho^B(\vec{y},t) \qquad (2.14)$$

where $\rho^A(\vec{x},t)$ and $\rho^B(\vec{y},t)$ are the "charge densities" of A and B. Since this amplitude should be added to (10) and since the charge density is the zero'th component of the 4 -current density it is convenient to write the sum of the two amplitudes in four-dimensional form



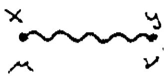
$$= \int d^4x \int d^4y\ (-i)J^A_\mu(x)\, D^{\mu\nu}(x-y)\,(-i)J^B_\nu(y) \qquad (2.15)$$

where D_{ij} is given by (13), and D_{oo} arises from (14)



$$D_{oo}(x-y) \equiv \quad = i\,\frac{\delta(x_0-y_0)}{4\pi|\vec{x}-\vec{y}|} = \int\frac{d^4k}{(2\pi)^4}\,\frac{i}{\vec{k}^2}\,e^{-ik\cdot(x-y)} \qquad (2.16)$$

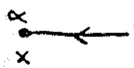
Collectively we have



$$\equiv D_{\mu\nu}(x-y) = \int\frac{d^4k}{(2\pi)^4}\,\frac{i\,T_{\mu\nu}}{k^2+i\epsilon}\,e^{-ik\cdot(x-y)} \qquad (2.17)$$

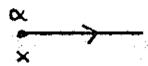
where $T_{oo} = k^2/\vec{k}^2$ and T_{ij} is as before.

What is meant by QED is normally the interaction between electrons, positrons and photons. Incoming plane wave electrons are described by solutions to the Dirac equation



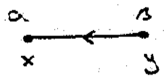
$$\vec{p}, u = u_\alpha\, e^{-ip\cdot x} \quad ; \quad \left[p_0 = \sqrt{\vec{p}^2+m^2}\right] \qquad (2.18)$$

where u is a positive energy spinor $((\not{p}-m)u = 0)$ while incoming positrons are described by



$$\vec{q}, v = \bar{v}_\alpha\, e^{-iq\cdot x} \quad ; \quad \left[q_0 = \sqrt{\vec{q}^2+m^2}\right] \qquad (2.19)$$

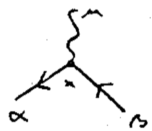
where v is a negative energy spinor $((\not{q}+m)v = 0)$ Outgoing particles are the adjoints of these. The propagator is



$$= \int\frac{d^4p}{(2\pi)^4}\left(\frac{i}{\not{p}-m+i\epsilon}\right)_{\alpha\beta} e^{-ip\cdot(x-y)} \qquad (2.20)$$

and it propagates electrons as well as positrons.

Finally there is the basic vertex



$$= -ie\,(\gamma_\mu)_{\alpha\beta} \qquad (2.21)$$

which accounts for the detailed interaction of the various spin and polarization components. The constant e is the electron charge. This is the "current" J_μ which we introduced above.

By means of the Feynman rules arbitrary S-matrix elements can now be calculated. There are of course divergence problems which we ignore here. I shall assume that you know something about quantum electrodynamics. This exposition has certainly not been adequate for obtaining an understanding of the subject. I merely wanted to show clearly how the propagator (17) arises.

3. Covariant quantum electrodynamics

The theory of relativity was discovered from the properties of light and it is therefore disconcerting that quantum electrodynamics, as formulated in the previous section, takes an - at least superficially - non-covariant form. Photon transversality is not a Lorentz invariant concept, but that seems also to be our only trouble. The rest of the theory, the electron propagator and the vertex, is nicely covariant.

From the basic vertex (2.21) it seems as if we can calculate the emission amplitude for photons that are not transverse. That has a priori no meaning. Physical photons are always transverse. Ask any experimentalist. Let us anyway try to allow arbitrary four vectors as polarization vectors replacing eq. (2.1) by

$$\overset{\mu}{\underset{x}{\bullet}}\!\sim\!\sim\!\sim\, \bar{k},\varepsilon = \varepsilon_\mu \, e^{-ik\cdot x} \; ; \quad (k_0 = \omega) \qquad (3.1)$$

This is a fundamental and discontinuous change of the theory. Instead of admitting two basic polarization states we now have four. Let us introduce a basis which keeps contact with the previous section. Besides the two transverse physical photons (T)

$$\varepsilon^\lambda_\mu = (0, \bar{\varepsilon}^\lambda) \qquad \lambda = 1,2 \qquad (3.2)$$

we define a longitudinal photon (L)

$$\varepsilon^3_\mu = (0, \bar{k}/\omega) \qquad (3.3)$$

and a scalar photon (S)

$$\varepsilon^0_\mu = (1, \bar{0}) \qquad (3.4)$$

The last two are unphysical. Nobody has ever seen them. We shall se why in a moment.

The new basis satisfies the orthonormality relation

$$\varepsilon^\lambda \cdot \varepsilon^{\lambda'} = g^{\lambda\lambda'} \qquad (3.5)$$

and every four-vector can be expanded along this basis

$$\varepsilon_\mu = c_0 \varepsilon^0_\mu + c_1 \varepsilon^1_\mu + c_2 \varepsilon^2_\mu + c_3 \varepsilon^3_\mu \qquad (3.6)$$

Using (5) we get $c_\lambda = g_{\lambda\lambda'} \, \varepsilon^{\lambda'} \cdot \varepsilon$ and hence from (6) we have

$$g_{\lambda\lambda'} \varepsilon^\lambda_\mu \varepsilon^{\lambda'}_\nu = \varepsilon^0_\mu \varepsilon^0_\nu - \varepsilon^1_\mu \varepsilon^1_\nu - \varepsilon^2_\mu \varepsilon^2_\nu - \varepsilon^3_\mu \varepsilon^3_\nu$$
$$= g_{\mu\nu} \qquad (3.7)$$

There is no way of getting around the change of signs. It is tied to the geometry of Minkowsky space.

The amplitude for emission of a photon in an arbitrary collision takes now the form

$$\bar{k},\varepsilon \sim\!\sim\!\sim \underset{x}{\bigcirc} = \int d^4x \, \varepsilon^*_\mu \, e^{ik\cdot x} (-i) J^\mu(x) = \varepsilon^*_\mu (-i) J^\mu(k) \qquad (3.8)$$

It is easy to see that the amplitude does not vanish in general when ε_μ is unphysical! It seems as if the unphysical photons can be produced in collisions between physical particles. The longitudinal photon is produced with probability $P_L = |\varepsilon^3_\mu J^\mu|^2$ while the scalar is produced with probability $P_S = |\varepsilon^0_\mu J^\mu|^2$. What is worse, the total probability for production of a photon with arbitrary polarizations $P = P_T + P_L + P_S$ is not covariant! So it seems as if we have completely failed in our purpose of constructing a covariant theory.

It is actually not that bad. Since the unphysical particles are introduced artificially and do not represent known states we may change our definition of probability. Using (7) we find that

$$P_C \equiv P_T + P_L - P_S = -g^{\mu\nu} J_\mu J_\nu^* \qquad (3.9)$$

If J_μ is a four-vector then P_c is certainly covariant. This is called the covariant probability and is not a probability at all, because of the minus sign. On the other hand, the minus sign cannot be avoided if we want a manifestly covariant theory. It is tied to the metric of Minkowski space. Because of the minus sign the scalar photon is called a ghost (what else could it be!).

The minus sign rescues in fact the situation. Since nobody has ever observed unphysical photons we must always add the probability for producing an arbitrary number of them in all possible processes. Since probability has to be positive we should then at least require $P_L - P_S \geq 0$. But if $P_L - P_S > 0$ then the theory predicts a finite probability that something happens when nothing should (even if we cannot observe unphysical states we can always find that the sum of the probabilities for producing physical states is less than one). The only way out is to require that the total probability for creation of unphysical states is zero. In our case this means

$$P_L - P_S = 0 \qquad (3.10)$$

Then the covariant probability is equal to the physical probability and everything seems alright.

There is a linear combination of unphysical photons which is particularly interesting, namely

$$\varepsilon_\mu^0 + \varepsilon_\mu^3 = k_\mu / \omega \qquad (3.11)$$

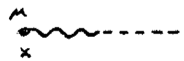
Since it will occur frequently in the following we shall give it a special name. We shall call it the gaugeon, because it is deeply connected with gauge invariance.

In QED the gaugeon decouples which means that

$$k^\mu J_\mu(k) = 0 \qquad (3.12)$$

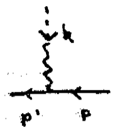
in a collision between physical particles. In coordinate space it means that $J_\mu(x)$ is a conserved current. Combining (11) and (12) we can derive (10).

In order to prove (12) and for future use it is convenient to introduce a graphical notation. Let us define an incoming gaugeon to have the amplitude

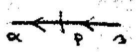

$$= k_\mu e^{-ik\cdot x} \qquad (3.13)$$

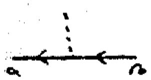
We shall allow gaugeons to be off-shell, i.e. we do not require $k^2 = 0$.

When a gaugeon enters the vertex (2.21) it produces the amplitude (in momentum space where all $e^{-ik\cdot x}$ factors are absent)

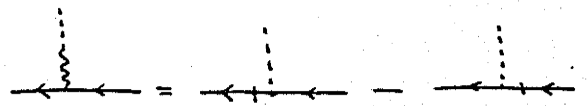

$$= -ie\,\not{k} = -ie(\not{p}' - \not{p}) = -ie(\not{p}' - m) + ie(\not{p} - m) \qquad (3.14)$$

On the right hand side we recognize the inverse electron propagator in momentum space. Let us introduce the graphical notation


$$= (\not{p} - m)_{\alpha\beta} \qquad (3.15)$$


$$= -ie\,\delta_{\alpha\beta} \qquad (3.16)$$

Then eq. (4) takes the graphical form

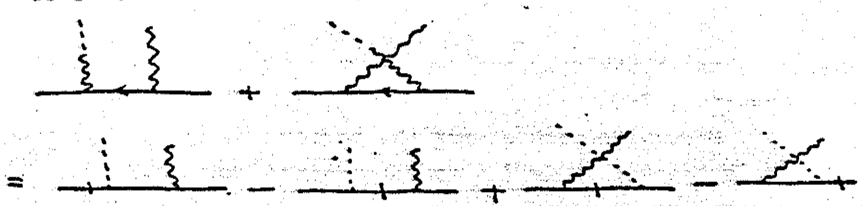
 (3.17)

This is what is called the basic Ward identity in QED.

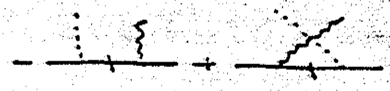
If the electrons are on-shell in (17), i.e. if we multiply with appropriate spinors, then the right hand side vanishes because of the Dirac equation $(\not p - m)u = (\not p' - m)u = 0$ Thus the gaugeon decouples from the on-shell vertex. This is quite general. Consider for example Compton scattering which in the lowest order of approximation is described by the diagrams

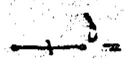
 (3.18)

Applying (17) to the final photon we get

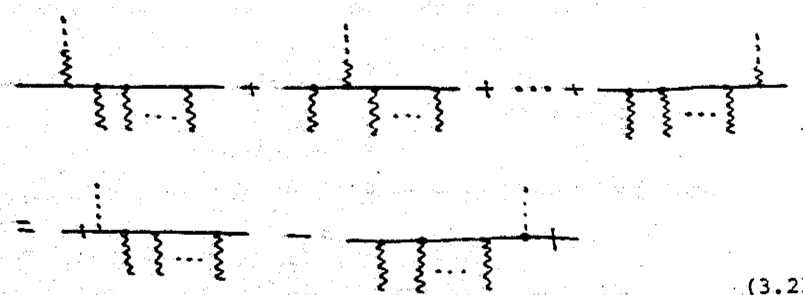
 (3.19)

The first and the last diagrams vanish on-shell and the two middle ones cancel

 $= ie^2\gamma_\mu - ie^2\gamma_\mu = 0$ (3.20)

because  $= i$. This relation (which is a Ward identity) is also valid off-shell. So the gaugeon decouples in the final state of Compton scattering.

In general we find from repeated application of (17) and (20)

 (3.21)

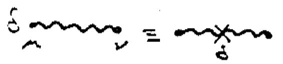
Since photons always are symmetrized with respect to each other this shows that the gaugeons always decouples from a line running through a graph. If we bend the electron line into a closed loop the right hand side also vanishes and this shows that the gaugeon decouples completely. The interaction between photons and electrons can never produce a gaugeon. This is what has become of gauge invariance in QED.

This proof also shows that an arbitrary number of gaugeons decouple. The photons themselves played no real role in the proof. Hence we have demonstrated in full generality that the unphysical states cancel in the final state of all processes with physical initial state.

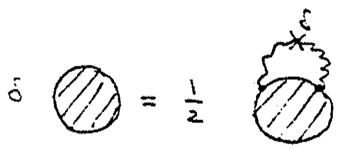
Let us return to eq. (12) and write it in the form

 (3.22)

where the blob represents the rest of the process. The only condition is that the electrons are on-shell. Let us try to change the photon propagator everywhere by the infinitesimal amount

 $= -i(k_\mu \Lambda_\nu(k) + k_\nu \Lambda_\mu(k))$ (3.23)

where $\Lambda_\mu(k)$ is an arbitrary infinitesimal function of k. Then it is pretty clear that the change to an arbitrary amplitude takes the form

 (3.24)

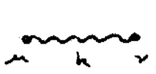
where the factor $\frac{1}{2}$ is due to statistics. But since the variation (23) contains a gaugeon in each term, (24) vanishes when the electrons are on-shell. By adding up infinitesimal elements of the form (23) we conclude that the S-matrix is invariant under arbitrary finite changes of this form.

If we introduce the "time-pointer" $n_\mu = (1, \bar{0})$ then $\tau_{\mu\nu}$ in the propagator (2.17) can be written

$$\tau_{\mu\nu} = -g_{\mu\nu} + \frac{n\cdot k\,(k_\mu n_\nu + k_\nu n_\mu)}{(n\cdot k)^2 - k^2} - \frac{k_\mu k_\nu}{(n\cdot k)^2 - k^2} \quad (3.25)$$

and the proof above shows that the last two terms can be dropped. With these terms the last non-covariant elements of QED have disappeared. Everything is now manifestly covariant, but the physical S-matrix elements are the same as before. The physical predictions are unchanged.

The general proof also shows that if we choose a photon propagator

 $= -i\,\frac{g_{\mu\nu}}{k^2} + i(1-a)\,\frac{k_\mu k_\nu}{(k^2)^2}$ (3.26)

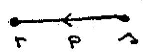
where a is an arbitrary parameter, the S-matrix is unchanged, i.e. independent of a. For $a = 1$ we obtain the Fermi gauge, and for $a = 0$ the Landau gauge. The limit of $a \rightarrow \infty$ is called the unitary gauge for reasons that will be given in section 10.

Summing up we can say that manifest covariance in QED requires the introduction of two new unphysical photon polarization states, one of which is produced with negative probability. The probability for producing unphysical states is however zero and this restores unitarity (the sum of all probabilities being 1) to the physical states.

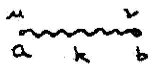
4. Quantum Chromodynamics (QCD)

A whole new world is opened up when we generalize the above ideas to include more than one electron and more than one photon. Instead of calling them electrons and photons we now call them quarks and gluons. Instead of electric charge the quarks and gluons carry colour. Like charge, colour cannot be destroyed.[2)]

The quarks are indexed by $r = 1, 2, \cdots, R$ and the gluons by $a = 1, 2, \cdots, A$. These indices enumerate the various colours and colour combinations. The basic free propagators are supposed to leave the colours unchanged



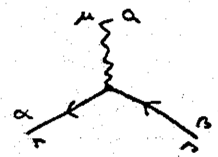
$$= \frac{i}{\not{p} - m} \delta_{rs} \qquad (4.1)$$



$$= -i \left[\frac{g_{\mu\nu}}{k^2} + (a-1) \frac{k_\mu k_\nu}{(k^2)^2} \right] \delta_{ab} \qquad (4.2)$$

We have chosen to work in the covariant gauge (3.26). The analysis in the preceding section shows that the non-physical gluons should then cancel against each other.

The basic process is the absorption and emission of a gluon by a quark



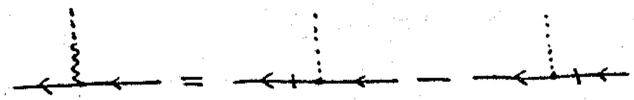
$$= -i\, T_{ras} (\gamma_\mu)_{\alpha\beta} \qquad (4.3)$$

We have allowed for a very general coupling scheme. The numbers T_{ras} are arbitrary complex numbers. They define a set of matrices

$$(T_a)_{rs} = T_{ras} \qquad (4.4)$$

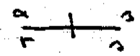
acting in quark colour space. These matrices must be linearly independent, for if they were not, there would be a linear combination of gluons that did not couple to the quarks and consequently could be left our of the theory. The factorization of the colour couplings and the spin couplings in (1), (2) and (3) just expresses that the colour of a quark or gluon does not depend on its state of polarization.

The basic Ward identity obeyed by the vertex (3) is as before (eq. (3.17))

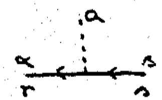


(4.5)

where now, instead of (3.15) and (3.16),



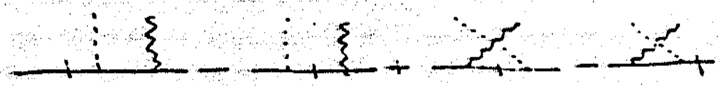
$$= \delta_{rs} (\not{p} - m)_{\alpha\beta} \qquad (4.6)$$



$$= -i\, T_{ras}\, \delta_{\alpha\beta} \qquad (4.7)$$

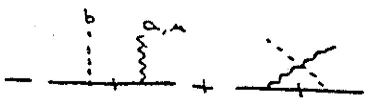
As before the gaugeons decouple (in the vertex) from on-shell quarks.

Let us now see what happens in the equivalent of Compton scattering (3.18), namely quark-gluon scattering. After a calculation entirely analogous to the previous one we get that the amplitude for emission of a gaugeon is



(4.8)

The first and the last vanish on-shell. The two middle ones become



$$= -i\,[T_a, T_b]\,\gamma_\mu \qquad (4.9)$$

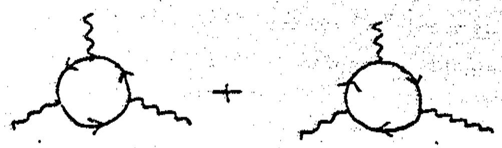
where we have used matrix notation for the colour couplings.

If the commutator in (9) vanishes

$$[T_a, T_b] = 0 \qquad (4.10)$$

then as before the theory is gauge-invariant, unitary in the physical subspace etc. etc. Since, as we shall argue below, the T_a are Hermitian matrices (in the same way as e was a real number in QED) they may be diagonalized simultaneously by unitary transformation of the quarks among themselves. A theory of this kind is just a set of juxtaposed QED's and contains nothing really new.

If the commutator does not vanish we are in trouble. Gaugeons can be emitted in quark-gluon scattering under the assumptions that we have made up to now. However, since a quark of one colour in this case can become a quark of a different colour (even allowing for unitary equivalences) by emitting a gluon, the gluon must itself carry colour. But then a gluon of one colour should be able to convert into a gluon of a different colour under the emission of a third gluon. Such a transition is for example furnished in third order in T by the diagrams



(4.11)

where the quarks form a closed loop. Because the T_a do not commute, the Furry theorem does not hold as in QED. It is easy to see that the sum is proportional to

$$\mathrm{Tr}\{T_a T_b T_c\} - \mathrm{Tr}\{T_c T_b T_a\} = \mathrm{Tr}\{[T_a, T_b] T_c\}$$

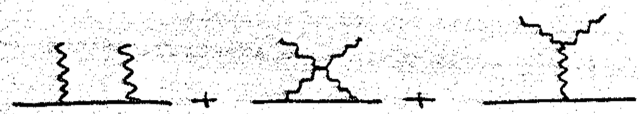
where Tr means the trace. Notice that the colour couplings and the polarization couplings factorize in (11) and that $\mathrm{Tr}\{[T_a, T_b] T_c\} = \mathrm{Tr}\{T_a [T_b, T_c]\}$ is completely antisymmetric in a, b and c.

Let us consequently try to add a triple gluon vertex to the theory and thereby make quark-gluon scattering gauge invariant. We choose the vertex to be

$$= C_{a_1 a_2 a_3} \{ g_{\mu_1\mu_2} (k_2 - k_1)_{\mu_3} + g_{\mu_2\mu_3} (k_3 - k_2)_{\mu_1} + g_{\mu_3\mu_1} (k_1 - k_3)_{\mu_2} \} \qquad (4.12)$$

where $C_{a_1 a_2 a_3}$ is totally antisymmetric. Actually there is not much freedom in this choice. The gluons are Bosons because they couple to the Fermi-quarks as in (3), and the triple vertex should consequently be totally symmetric in the momenta (all running towards the vertex) and the indices. If we factorize the colour couplings and the polarization couplings there are only two possibilities. They are either both completely symmetric or both completely antisymmetric. The simplest way of coupling three vectors is linear in the momenta. The totally symmetric case gives the coupling $g_{\mu_1\mu_2} k_{3\mu_3} + g_{\mu_2\mu_3} k_{1\mu_1} + g_{\mu_3\mu_1} k_{2\mu_2}$. But this is unacceptable because it does not couple purely transverse gluons with each other. Hence the only case that is left is the totally antisymmetric colour and polarization couplings, i.e. (12).

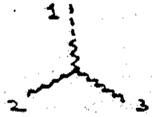
With this vertex in the theory, quark gluon scattering gets an extra contribution in the lowest order



(4.13)

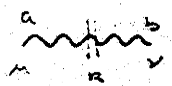
What is then the relationship which should be obeyed by the T_{ra} and the C_{abc} in order to eliminate the gaugeon amplitude?

First we find



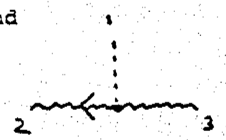
$$= C_{a_1a_2a_3}(g_{\mu_2\mu_3}k_2^2 - k_{2\mu_2}k_{2\mu_3}) + C_{a_1a_3a_2}(g_{\mu_3\mu_2}k_3^2 - k_{3\mu_3}k_{3\mu_2}) \qquad (4.14)$$

where we have made use of momentum conservation $k_1 = -k_2 - k_3$. So if we define



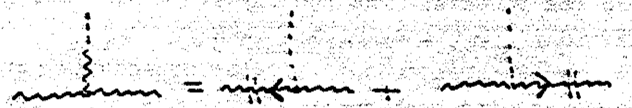
$$= (g_{\mu\nu}k^2 - k_\mu k_\nu)\delta_{ab} \qquad (4.15)$$

and



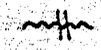
$$= C_{a_1a_2a_3}\, g_{\mu_2\mu_3} \qquad (4.16)$$

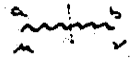
then eq. (14) takes the form



(4.17)

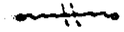
The vertex in (16) is <u>not</u> symmetric in the Bose variables 2 and 3. It violates the spin-statistics connection. It is therefore necessary with a sign convention. We choose a small arrow to distinguish one of the gluons.

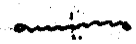
Eq. (17) is the basic gluon Ward identity. Since  vanishes for a gaugeon or an on-shell transverse gluon there is a good chance that gaugeons might decouple. The expression (15) is not the inverse of the propagator (2), which is



$$= (g_{\mu\nu}k^2 + (\tfrac{1}{\alpha} - 1)k_\mu k_\nu)\,\delta_{ab} \qquad (4.18)$$

except in the limit of $\alpha \to \infty$.

In the following we shall need a graphical notation for the relation between the double-barred propagator  and the single barred propagator  $= -i g_{\mu\nu}\delta_{ab}$. In keeping with the previous notation (3.13) we have



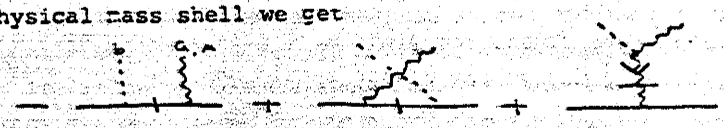
$$= -i(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2})\delta_{ab} =$$

(4.19)

where the dots represent a scalar propagator[3)]

$$= -\frac{i}{k^2}\delta_{ab} \qquad (4.20)$$

After these preparations we can now evaluate the gaugeon amplitude (4.13). Dropping all terms that vanish on the physical mass shell we get



$$= -i([T_a, T_b] - i C_{abc}T_c) \qquad (4.21)$$

instead of (9). In the third graph we have used (19) and (5) and the mass-shell conditions. So we conclude that the gaugeon amplitude vanishes in physical gluon-quark collisions provided that

$$[T_a, T_b] = i\,C_{abc}\,T_c \qquad (4.22)$$

When this condition is fulfilled the gaugeon amplitude is zero and the negative probability for creation of a scalar gluon will exactly be cancelled by the positive probability for the creation of a longitudinal gluon.

The condition (22) is nothing but the defining relations for a Lie algebra with generators T_a and structure constants C_{abc}. We shall discuss Lie algebras in the following section.

Lie algebras usually arise in the theory of continuous groups. The elements in the neighborhood of the identity

$$U = 1 - i \sum_a T_a\,\omega_a \qquad (4.23)$$

where ω_a are real infinitesimal parameters, may be shown to satisfy the Lie condition (22). The matrices represent infinitesimal rotations in colour space and must be unitary in order to preserve probability. But that implies that the T_a are Hermitian. If they are not Hermitian, ghosts may arise as in the case of the Lorentz group. In the standard introductions to gauge theories the group aspect is emphasized much more than here. We have shown that the Lie algebra (22) arises solely as a consequence of the demand that the gaugeons decouple, i.e. from unitarity. Note, however, that we have only shown that one gaugeon decouples. The general proof which we carried out in the previous section for QED cannot be taken over unmodified because the gluons themselves interact with each other.

5. The structure of Lie Algebras

We have seen that a necessary condition for the decoupling of the gaugeons is that the quark-gluon-quark coupling constants T_{rab} form a Lie algebra

$$[T_a, T_b] = i\, C_{abc} T_c \qquad (5.1)$$

The generators T_a are linearly independent and Hermitian

$$T_a^+ = T_a \qquad (5.2)$$

From this it follows that the antisymmetric structure constants, i.e. the triple gluon couplings,

$$C_{abc} = -C_{bac} = -C_{acb} = -C_{cba} \qquad (5.3)$$

are real. The algebra itself is formed by all linear combinations $T_a \omega_a$ (implicit sum over a) where the ω_a are arbitrary real parameters.

In the following we shall go through some of the more basic concepts of Lie algebras. This analysis may be found in the standard textbooks on the subject[4] but we extract it here in order to isolate the physical contents. The main result will be that we shall understand how many independent, and freely variable, coupling constants that the theory possesses. It is clear that the relations (1), (2) and (3) impose restrictions on the general couplings but it is not immediately clear how many there really are.

An element of the algebra $T_a \omega_a$ is called <u>Abelian</u> if it commutes with all other elements, i.e.

$$[T_a \omega_a, T_b] = 0 \qquad (5.4)$$

for all b. The set of all Abelian elements form a subalgebra called the center of the original algebra. We can always choose the generators in such a way that the first few

span the center. A Lie algebra without a (non-trivial) center is called <u>semi-simple</u>. A Lie algebra which is equal to its center is called <u>Abelian</u>. We discussed briefly Abelian algebras in the previous section (eq. (4.10)).

If a generator T_a is Abelian we get, because of the linear independence of the T_c in eq. (1), that $C_{abc} = 0$ for all b and c. If T_b and T_c are non-Abelian generators then $[T_b, T_c] = i\, C_{bca} T_a$ and the antisymmetry of $C_{abc} = C_{bca}$ now shows that an Abelian generator can never arise as the result of commuting two arbitrary generators. Thus the center is completely independent of the rest of the algebra. We may group the generators into two sets, one consisting of all the Abelian generators, and one consisting of all the non-Abelian generators. The former spans the center while the latter spans a semi-simple Lie algebra in its own right. We are therefore in the following free to assume that our Lie algebra is semi-simple. We can always get back to the general case by adding a center.

The Jacoby identities for the generators say that

$$[[T_a, T_b], T_c] + [[T_b, T_c], T_a] + [[T_c, T_a], T_b] = 0 \qquad (5.5)$$

Expressed in terms of the structure constants we have

$$C_{abu} C_{ucv} + C_{bcu} C_{uav} + C_{cau} C_{ubv} = 0 \qquad (5.6)$$

Let us define the $A \times A$ matrices

$$(\Theta_a)_{bc} = i\, C_{bac} \qquad (5.7)$$

Then eq. (6) takes the more convenient form

$$[\Theta_a, \Theta_b] = i\, C_{abc} \Theta_c \qquad (5.8)$$

The matrices Θ_a are Hermitian because of the antisymmetry and reality of the structure constants. They have the same commutation relations as the original generators and are said to generate the adjoint or regular representation of the algebra.

If a generator T_a is Abelian then $\Theta_a = 0$. Conversely, if no generators are Abelian, i.e. if the algebra is semi-simple, then the regular generators are linearly independent. For if $\omega_a \Theta_a = 0$ then $[\omega_a T_a, T_b] = 0$ for all b, and since the algebra is semi-simple no Abelian element exists except $\omega_a = 0$. From this it follows that the <u>Cartan-Killing form</u>[5)]

$$g_{ab} = C_{acd}\, C_{bcd} = (\Theta_c \Theta_c)_{ab} \qquad (5.9)$$

which by definition is real, symmetric and positive definite, is non-singular for a semi-simple Lie algebra. To verify this we notice that $g_{ab}\,\omega_a \omega_b = \sum_{bc} |\omega_a C_{abc}|^2 \geq 0$ can only vanish for $\omega_a C_{abc} = 0$. For a semi-simple algebra this implies $\omega_a = 0$.

A central question in the analysis of semi-simple Lie algebras concerns the existence of matrices $Z = \{Z_{ab}\}$ that commute with the generators of the regular representation

$$[Z, \Theta_a] = 0 \qquad (5.10)$$

Such matrices are said to be invariant. As the structure constants are real both the real and imaginary part of Z must satisfy this equation. Hence we can safely assume Z to be real. We shall now also show that Z is symmetric. We first establish the trivial relations

$$\mathrm{Tr}\,[\Theta_a Z \Theta_b] = (\Theta_c Z \Theta_c)_{ab} = (Z g)_{ab} \qquad (5.11)$$

The first follows by manipulating the indices using the antisymmetry while the last follows from (9) and (10). The trace is, however, invariant under transposition. Hence $\mathrm{Tr}\,[\Theta_a Z \Theta_b] = \mathrm{Tr}\,[\Theta_b Z^T \Theta_a]$ where Z^T is the transposed matrix $(\Theta_a^T = -\Theta_a)$. But because the trace of a product is independent of the order of the factors $(\mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA))$, we get $Z g = Z^T g$ and since g is non-singular for a semisimple Lie algebra $Z = Z^T$, i.e. Z is symmetric.

It now follows that if two matrices Z_1 and Z_2 are invariant, then the product $Z_1 Z_2$ is also invariant. The same is true about the commutator $[Z_1, Z_2]$. But the commutator is antisymmetric and must consequently vanish. In other words, <u>all invariant matrices must commute with each other</u>. A set of commuting, real symmetric matrices can be diagonalized simultaneously by means of an orthogonal transformation. Such a transformation only introduces new linear combinations of gluons and changes nothing in the physics. We are therefore free to assume that all invariant matrices are diagonal

$$Z_{ab} = \lambda_a\, \delta_{ab} \qquad \text{(no sum)} \qquad (5.12)$$

where the eigenvalues λ_a are real numbers.

But the eigenvalues are not independent. Inserting (12) into (10) we get

$$(\lambda_a - \lambda_b)\, C_{abc} = 0 \qquad \text{(no sum)} \qquad (5.13)$$

This is a terribly strong condition. For if $\lambda_a \neq \lambda_b$ then $C_{abc} = 0$, and conversely if $C_{abc} \neq 0$ then $\lambda_a = \lambda_b = \lambda_c$. So for each non-vanishing structure constant we get a triplet of identical eigenvalues. The smallest semi-simple Lie algebra must at least have three generators. It is the well-known algebra of angular momenta $[J_x, J_y] = iJ_z$ etc. In the general case we have more than one triplet of eigenvalues.

If some of the triplets overlap we get higher multiplets of identical eigenvalues. In this way the index set A of the generators is divided into the union of mutually disjoint sets A_σ , $\sigma = 1, 2, \dots, S$, each corresponding to a multiplet of identical eigenvalues such that different sets correspond to different eigenvalues.

If T_a and T_b belong to different sets then $\lambda_a \neq \lambda_b$ and $C_{abc} = 0$. Hence $[T_a, T_b] = 0$. Furthermore, if T_a and T_b belong to the same set, A_σ , and if $[T_a, T_b] = i C_{abc} T_c$ with non-vanishing C_{abc} , then T_c belongs to that same set. Thus the generators T_a for $a \in A_\sigma$ form by themselves a closed Lie algebra, and commute with all the generators not in this set. The subset of generators has furthermore the property that all invariant matrices belonging to this subset are proportional to the unit matrix because all the eigenvalues are identical. Such an algebra is called <u>simple</u>.

Recapitulating we have proved that every semisimple Lie algebra is the direct sum of simple Lie algebras. In general <u>an arbitrary Lie algebra*) is the direct sum of an Abelian center and a number of simple components</u>.

We now quote without proof a very remarkable and striking result: <u>The simple Lie algebras can be completely classified and exhaustively enumerated (Cartan).</u> The Cartan analysis is not very difficult but a proof would lead us too far away and I refer you to the mathematical literature.[4)]

---

*) Strictly speaking this result has only been proven for <u>compact</u> Lie algebras where the regular generators can be chosen Hermitian. The theorem is modified slightly for non-compact algebras.[4)]

The simple Lie algebras belong - apart from isomorphisms - to four infinite set and a finite exceptional set. They are

A_n $\quad n = 1, 2, \dots$

These algebras are generated by the special unitary groups $SU(n+1)$ of main importance for physics.

B_n $\quad n = 2, 3, \dots$

These algebras are generated by the odd orthogonal groups $O(2n+1)$. The familiar rotation group $O(3)$ is isomorphic (in its algebra) to $SU(2)$.

C_n $\quad n = 3, 4, \dots$

These algebras are generated by the even symplectic groups, $Sp(2n)$. They are best characterized for physics as those linear transformations on the p's and q's that leave the Poisson brackets invariant in classical canonical mechanics. $Sp(2)$ is (algebraically) isomorphic to $SU(2)$ and $Sp(4)$ to $O(5)$.

D_n $\quad n = 4, 5, \dots$

They are generated by the even orthogonal group $O(2n)$ $O(2)$ is Abelian and isomorphic to $U(1)$, $O(4)$ is isomorphic to the direct product $SU(2) \times SU(2)$ (and not simple) while $O(6)$ is isomorphic (algebraically) to $SU(4)$.

E_6, E_7, E_8, F_4, G_2

These are the exceptional algebras and have not (yet) received much attention from physicists.

The Cartan classification leads to a canonical choice for the structure constants of simple Lie algebras in which the Cartan-Killing form is proportional to the unit matrix (it is an invariant as may easily be verified). This fact is not changed by an arbitrary orthogonal transformation (of no physical importance) of the gluons among each other followed by an overall change of scale. If we call this scale g then

$$\sum_{abc} |C_{abc}|^2 = \nu g^2 \qquad (5.14)$$

where ν is some positive pure number, say 2 or $5\frac{1}{3}$. The scale g measures the strength of the triple gluon coupling. If the Lie algebra consists of more than one simple component, each component has its own coupling constant, g_σ. Because of (1) the coupling constant g_σ will also be the overall scale for those generators T_a that belong to the same simple component, σ. In general we can write

$$g_{ab} = \mathrm{Tr}\,[G_a G_b] = C_a g_a^2 \delta_{ab} \qquad (5.15)$$

$$\mathrm{Tr}\,[T_a T_b] = C_a^T g_a^2 \delta_{ab} \qquad (5.16)$$

Both the traces are easily seen to be invariant matrices and hence diagonal (in the semisimple case). The factors C_a and C_a^T are purely numerical. All quantities g_a, C_a, C_a^T are independent of a within a simple component. The ratio C_a^T / C_a is called the index of the representation T_a.

Finally some words about the representations of Lie algebras. In the analysis above we only considered the regular gluon representation, G_a. The quark representation T_a does not possess the same simple structure. The generators act on R-dimensional column vectors $u = \{u_r\}$, and we now ask the question that happens when they act successively on the same vector u.

We get the sequence of vectors

$$u,\ T_a u,\ T_b T_a u,\ \ldots \qquad (5.17)$$

From a certain stage and onwards the new vectors must be linearly dependent on the preceding ones, because the space is finite dimensional. This means that one can associate a maximal number of linearly independent vectors with each vector u. They span a subspace $\| u$ constructed from all possible linear combinations of the vectors (17). This subspace is invariant in the sense that if we act on an arbitrary vector in it, say v, then $T_a v$ also belongs to $\| u$. Let now $\perp u$ be the subspace orthogonal to $\| u$. It consists of all the vectors w satisfying that $w^+ v = 0$ when v is in $\| u$. The orthogonal space $\perp u$ is also invariant. For if w is in $\perp u$ then, because of the Hermiticity of T_a, we have $(T_a w)^+ v = w^+ T_a v = 0$, since $T_a v$ is in $\| u$.

We have now shown that for each vector u there are two orthogonal invariant subspaces $\| u$ and $\perp u$, in the space of quark colours. If $\perp u$ is not empty, i.e. if $\| u$ is not the whole space, then we may choose basic vectors (i.e. quark colours) lying entierely in $\| u$ and $\perp u$. Since $w^+ T_a v = 0$ we see that T_a takes the block form

$$T_a = \left\{ \begin{array}{c|c} T_a^{\|} & 0 \\ \hline 0 & T_a^{\perp} \end{array} \right\} \qquad (5.18)$$

If follows immediately that $T_a^{\|}$ and $T_a^{\perp}$ both must satisfy the basic relations (1). The matrix $T_a^{\|}$ only acts in the subspace $\parallel u$ and $T_a^{\perp}$ only acts in $\perp u$. We may therefore continue the reduction of T_a inside each of these subspaces. Eventually T_a takes the block form

$$T_a = \left\{ \begin{array}{c|c|c} T_a^1 & C & \\ \hline O & T_a^2 & \\ \hline & & \ddots \\ \hline & & T_a^I \end{array} \right\} \tag{5.19}$$

where each matrix T_a^i, $i = 1, 2, \cdots, I$ obeys the same commutation relations as T_a. The matrices T_a^i are supposed to be <u>irreducible</u> which means that the corresponding subspace S_i cannot be reduced further. If u is an arbitrary vector in S_i, then the sequence $u, T_a^i u, \ldots$ must span S_i completely. We can get from any vector to another by repeated application of the generators (and forming linear combinations). The trivial representation $T_a^i = 0$ cannot occur in (19) because that would mean that a subset of quarks was completely uncoupled and should have been left out of the theory from the beginning.

Irreducible representations are related to the concept of simplicity of a Lie algebra. In fact, the analysis which carried out above was nothing but a decomposition of the regular representation into irreducible components. We proved the important result that any matrix, which commuted with the regular generations of a simple algegra, was proportional to the unit matrix. A similar result holds in general for irreducible representations (Schuurs Lemma). Let

$$[Z, T_a] = 0 \tag{5.20}$$

for all generators of an irreducible representation, and let u be an eigenvector of Z with eigenvalue λ

$$Z u = \lambda u \tag{5.21}$$

All matrices have at least one eigenvector. Applying T_a repeatedly to both sides of this equation, using (20), we see that the sequence $u, T_a u, \cdots$ and all its linear combinations also satisfy (21). Hence because of the irreducibility, all vectors u in the space satisfy (21). Hence $Z_{rs} = \lambda \delta_{rs}$.

The smallest Lie algebra is A_1 corresponding to $SU(2)$. It has three generators I_1, I_2, I_3 with the commutation relations

$$\begin{aligned} [I_1, I_2] &= i I_3 \\ [I_2, I_3] &= i I_1 \\ [I_3, I_1] &= i I_2 \end{aligned} \tag{5.22}$$

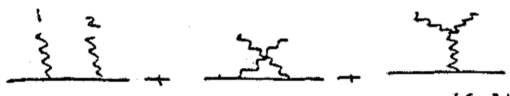
The regular representation is $(\Theta_i)_{kl} = -i\,\varepsilon_{ikl}$ where ε_{ikl} is the Levi-Civita symbol. The lowest dimensional representation is $I_i = \frac{1}{2}\tau_i$ where τ_i are the 2×2 Pauli matrices. In the Weinberg model an Abelian generator Y, called hypercharge, is added. The group is now $SU(2) \times U(1)$. The lowest non-trivial representation is now 3-dimensional because a two-dimensional Y would have to commute with the Pauli-matrices and only the unit matrix does that.*) In $SU(3)$, I_1, I_2, I_3 and Y are chained together by the addition of four more generators.

*) The Pauli matrices form (necessarily) an irreducible representation of the simple algebra A_1.

6. Quark annihilation

By requiring that a single gaugeon decouples in quark-gluon scattering we have deduced that the quark-gluon and triple gluon couplings must form a Lie algebra, and we have thoroughly analyzed the consequences of this condition. It is, however, not at all evident that the gaugeons will decouple in other processes.

The annihilation of a quark and an antiquark into two gluons is described by the same diagrams (4.13) as quark-gluon scattering. Making explicit the dependence on the gluon polarizations we denote the amplitude

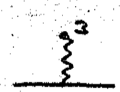
$$\varepsilon_1^{\mu_1}\varepsilon_2^{\mu_2} J^{a_1 a_2}_{\mu_1\mu_2} =$$  (6.1)

where $J^{a_1 a_2}_{\mu_1\mu_2}$ is the rest of the amplitude including quark spinors. Using the Ward identity (4.21) we find in the case that one of the gluons is a gaugeon

$$k_1^{\mu_1}\varepsilon_2^{\mu_2} J^{a_1 a_2}_{\mu_1\mu_2} =$$ 

$=$ 

$$= C_{a_1 a_2 a_3}\left(\varepsilon_2^{\mu_3} k_2^2 - \varepsilon_2\cdot k_2\, k_2^{\mu_3}\right) J^{a_3}_{\mu_3} \qquad (6.2)$$

where $J^{a_3}_{\mu_3}$ is the amplitude of 

If gluon 2 is physical then $\varepsilon_2\cdot k_2 = 0$ and $k_2^2 = 0$, so that the amplitude vanishes. This also happens if the second gluon is a gaugeon. The amplitude for producing two gaugeons is zero, as well as the amplitude for producing one transverse gluon and one gaugeon. But the amplitude for producing one gaugeon and one longitudinal (or scalar) gluon does not vanish. This is different from QED where all amplitudes with just one gaugeon vanished independently of the polarization of the other photons. We can consequently not be sure that the probability for producing unphysical gluons is zero.

The covariant probability, i.e. the probability for producing arbitrarily polarized gluons is

$$P_{CC} = \frac{1}{2} g^{\mu_1\nu_1} g^{\mu_2\nu_2} J^{a_1 a_2}_{\mu_1\mu_2} J^{a_1 a_2}_{\nu_1\nu_2} \times \qquad (6.3)$$

The factor $\frac{1}{2}$ is due to the identity of the gluons in the final state. In the same way as (3.9) we have

$$\begin{aligned} P_{CC} = {} & P_{TT} + P_{TL} - P_{TS} + P_{LT} - P_{ST} \\ & + P_{LL} - P_{LS} - P_{SL} + P_{SS} \end{aligned} \qquad (6.4)$$

where the indices correspond to the polarizations of the final gluons. Since the gaugeon decouples when the second gluon is transverse we have

$$P_{TL} - P_{TS} = P_{LT} - P_{ST} = 0 \qquad (6.5)$$

But as we shall see $P_{LL} - P_{LS} - P_{SL} + P_{SS}$ does not vanish.

The easiest way to find the value of the unphysical probability is by calculating

$$P_{TT} = \frac{1}{2} \sum_{\substack{\text{transverse}\\ \text{gluons}}} \left| \varepsilon_1^{\mu_1} \varepsilon_2^{\mu_2} J^{a_1 a_2}_{\mu_1 \mu_2} \right|^2$$

$$= \frac{1}{2} \tau_1^{\mu_1 \nu_1} \tau_2^{\mu_2 \nu_2} J^{a_1 a_2}_{\mu_1 \mu_2} J^{a_1 a_2 \times}_{\nu_1 \nu_2} \qquad (6.6)$$

where $\tau_1^{\mu_1\nu_1}$ and $\tau_2^{\mu_2\nu_2}$ are of the form (3.25) with $k_1^2 = k_2^2 = 0$. Since one gaugeon decouples when the other is transverse we can replace $\tau_2^{\mu_2\nu_2}$ by $-g^{\mu_2\nu_2}$ to get

$$P_{TT} = \frac{1}{2} \left(g^{\mu_1\nu_1} - \frac{n^{\mu_1} k_1^{\nu_1} + n^{\nu_1} k_1^{\mu_1}}{n \cdot k_1} + \frac{k_1^{\mu_1} k_1^{\nu_1}}{(n \cdot k_1)^2} \right)$$

$$\cdot\, g^{\mu_2\nu_2} J^{a_1 a_2}_{\mu_1 \mu_2} J^{a_1 a_2 \times}_{\nu_1 \nu_2} \qquad (6.7)$$

By making repeated use of eq. (2) and the analogous equation obtained by exchanging 1 and 2 we get

$$P_{TT} = P_{CC} -$$

$$- \mathrm{Re} \left(C_{a_2 a_1 a_3} k_1^{\mu_3} J^{a_3}_{\mu_3} C_{a_1 a_2 b_3} k_2^{\nu_3} J^{b_3 \times}_{\nu_3} \right) \qquad (6.8)$$

The last term in (7) gives no contribution while the two middle terms give contributions that are complex conjugates of each other. Using the antisymmetry of C_{abc} and the fact that $(k_1+k_2)^{\mu_3} J^{a_3}_{\mu_3} = -k_3^{\mu_3} J^{a_3}_{\mu_3} = 0$ for on-shell quarks, the last term in (8) becomes a perfect square and we can write

$$P_{CC} \overset{!}{=} P_{TT} + \left| C_{a_1 a_2 a_3} k_2^{\mu_3} J^{a_3}_{\mu_3} \right|^2 \qquad (6.9)$$

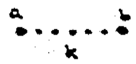
By comparison with (4) we see that the unphysical probabilities do in fact not cancel, but we have

$$P_{LL} - P_{LS} - P_{SL} + P_{SS}$$

$$= \left| C_{a_1 a_2 a_3} k_2^{\mu_3} J^{a_3}_{\mu_3} \right|^2 \qquad (6.10)$$

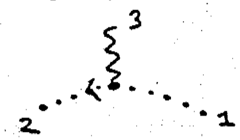
This is a disastrous result. Even if we cannot find any way of observing the unphysical gluons directly, we can observe them indirectly because the sum of all probabilities must be equal to 1. When quarks and antiquarks annihilate, the above result predicts that in a certain fraction of the annihilation processes nothing observable comes out.

The way out of this dilemma was originally found by Feynman[6] and given a more theoretical justification by Fadeev and Popov.[7] For each gluon a new unphysical particle is included in the theory with propagator



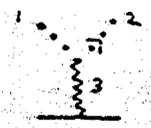
$$= -\frac{i}{k^2} \delta_{ab} \qquad (6.11)$$

and coupling



$$= C_{a_1 a_2 a_3} k_2^{\mu_3} \qquad (6.12)$$

This particle can never be present initially and cannot be observed in the final state. Since quarks can annihilate into a pair of these particles via the diagram



$$= C_{a_1 a_2 a_3} k_2^{\mu_3} J^{a_3}_{\mu_3} \qquad (6.13)$$

we must include the pair production probability

$$P_{F\bar{F}} = \left| C_{a_1 a_2 a_3} k_2^{\mu_3} J^{a_3}_{\mu_3} \right|^2 \qquad (6.14)$$

in the total probability for producing unphysical states. However, in order to cancel (10) we must use a negative sign

$$P_{LL} - P_{LS} - P_{SL} + P_{SS} - P_{F\bar{F}} = 0 \qquad (6.15)$$

The new particles are not Bosons because their coupling (12) is asymmetric in the variables 1 and 2. It is for this reason that an arrow has been put on leg number 2 to distinguish it from leg number 1.*) If we, on the other hand, assume them to be Fermions this would also explain the minus sign in front of $P_{F\bar{F}}$. The spin-statistics theorem is related to positive definite pair production probabilities and if we violate this theorem by making scalar Fermions we get naturally a minus sign in the pair production "probability". The Feynman-Fadeev-Popov (FFP) ghost is different from its antiparticle and hence (12) needs not be antisymmetric in 1 and 2. The symmetry consideration do in fact not determine the statistics of the FFP ghost. only the minus sign in (15) does that.[8)]

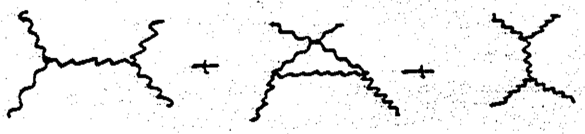
Having introduced the new unphysical particles they will show up other places in the theory as well, as for example in closed loops. It will be shown in section 8 that the minus sign in front of all ghost loops precisely allows for the cancellation of unwanted terms that otherwise would prevent the decoupling of the gaugeons from the gluon loops.

---

*) The arrow does not indicate the flow of momentum which is always towards the vertex.

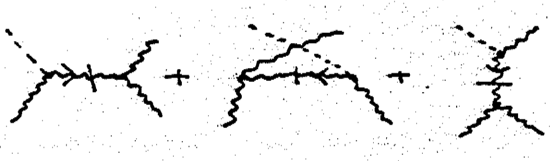
7. Gluon-gluon scattering

Continuing our investigation of the theory we now turn to gluon-gluon scattering which in the lowest order of approximation is given by the diagrams (in the theory so far constructed)



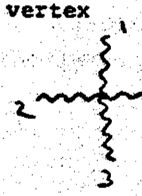
(7.1)

Taking one of the final gluons to be a gaugeon and the others to be transverse (physical) we get after dropping terms that vanish on-shell in analogy with (4.21)



(7.2)

Most of this cancels because the gluon couplings belong to the regular representation satisfying (5.8), but a quick calculation shows that not all of it vanishes due to the momentum dependence of the basic vertex. The structure is, however, so simple that the non-vanishing remainder way be eliminated by means of a quadruple gluon vertex



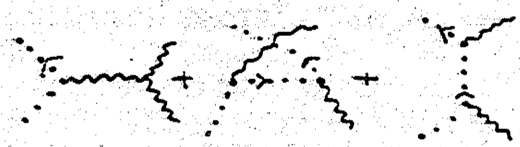
$$= -i\{C_{a_1a_2b}C_{ba_3a_4}(g_{\mu_1\mu_3}g_{\mu_2\mu_4} - g_{\mu_1\mu_4}g_{\mu_2\mu_3}) + C_{a_1a_3b}C_{ba_4a_2}(g_{\mu_1\mu_4}g_{\mu_2\mu_3} - g_{\mu_1\mu_2}g_{\mu_3\mu_4}) + C_{a_1a_4b}C_{ba_2a_3}(g_{\mu_1\mu_2}g_{\mu_3\mu_4} - g_{\mu_1\mu_3}g_{\mu_2\mu_4})\} \tag{7.3}$$

A little algebra now shows that adding this diagram to (1) the gaugeon decouples

(7.4)

This Ward identity is valid for all momenta, even off-shell ones.

In calculating the physical cross section for the process the Feynman-Fadeev-Popov (FFP) ghosts will also participate (in the covariant calculation). In this case there is, however, more than one diagram



(7.5)

When the square of this is subtracted from the probability for production of arbitrary, physical and non-physical gluons, the non-physical contribution all cancel and only the transverse part remains, exactly as in the previous section.

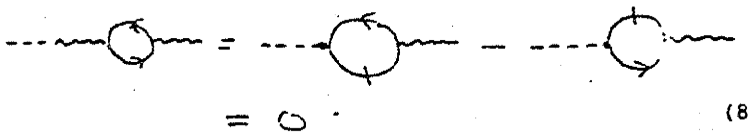
8. Closed loops

The specific quantum effects of field theory show up in closed loops. We have so far only considered tree diagrams without closed loops and have no guarantee that the closed loops do not present special problems. In QED the closed loops could only be formed by electrons and the gaugeons decoupled nicely due to the Ward identity (3.16). In QCD the gluons and FFP ghosts may also form closed loops and a more complicated analysis is necessary.

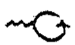
We shall only consider the simplest closed loops, namely those contributing to gluon vacuum polarization. The quarks contribute only one diagram in lowest order

(8.1)

If the left hand gluon is a gaugeon we get from (4.5)



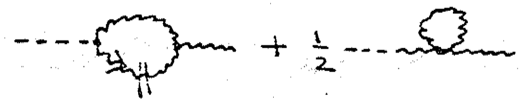
(8.2)

The last equation follows by direct calculation, or simpler by making use of the Ward identity (4.21), bending the quark lines into a loop. The last diagram in (4.21) gives rise to a tadpole  which vanishes because it is a Lorentz invariant vector and only the null vector is invariant.

The contribution to vacuum polarization from gluons and FFP ghosts is

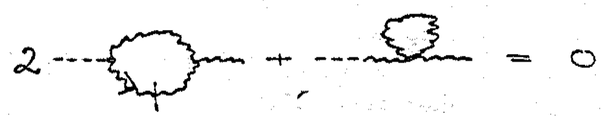
(8.3)

where we have written statistical factors and the Fermion minus sign explicitly. When the left hand gluon is a gaugeon we get from the two first diagrams using (4.17)



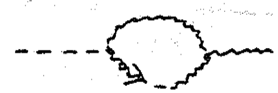
(8.4)

The two diagrams arising from (4.17) are topologically identical and that removes the factor ½. From the Ward identity (7.4) we get by tying the two lower gluons together



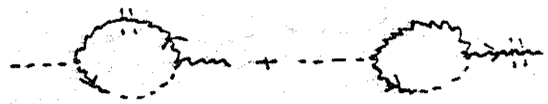
(8.5)

Using this identity and (4.19) we get from (4) the remainder

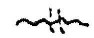


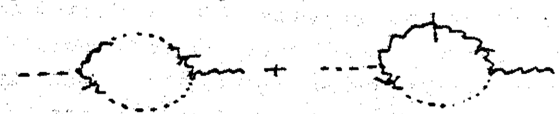
(8.6)

In the right hand vertex we have again an incoming gaugeon and can apply (4.17) to get



(8.7)

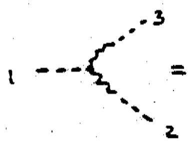
The last diagram must vanish because the loop has only one Lorentz index and must accordingly be proportional to the gaugeon momentum k_μ which is orthogonal to  $= g_{\mu\nu}k^2 - k_\mu k_\nu$. So only the first remains. Here we use again (4.19) to get



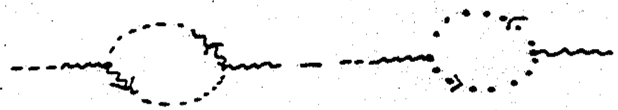
(8.8)

The last vanishes because it is proportional to $\int d^4k \, \frac{k_\mu}{k^2}$

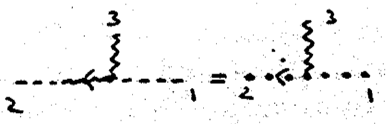
which - although badly divergent - must vanish for symmetry reasons. The left hand vertex in the first diagram is according to (4.16)

 $= C_{a_1 a_2 a_3} k_2 \cdot k_3 = C_{a_3 a_2 a_1} k_2 \cdot k_1 - C_{a_1 a_2 a_3} k_2^2$ (8.9)

where we have made use of momentum conservation. The last term vanishes for symmetry reasons $\int d^4 l \, l_\mu / l^2 = 0$ when inserted into the loop and the first term corresponds to moving the wiggle on leg 3 to leg 1. Hence the remainder is finally

 (8.10)

where we have also added the ghost contribution. Since

 $= C_{a_1 a_2 a_3} k_{2\mu_3}$ (8.11)

follows from (4.16) and (6.13) the two diagrams cancel.

So the gaugeons also decouple in vacuum polarization and this is entirely due to the inclusion of the FFP ghosts.[9)]

Closed loops are, however, not so simple as this discussion has led you to believe. They are badly divergent and this section only makes sense, if the divergences are brought under control by means of a regularization procedure. This procedure should furthermore be such that the naive combinatorics that we have performed is not changed. The dimensional regularization procedure invented by 't Hooft and Veltman[10)] is just such a scheme.

9. Power counting

The physical interpretation of the theory requires that its basic parameters are renormalized. This means that the original parameters with which the theory was originally constructed (masses and couplings) in fact do not have the interpretation that we gave them then, but change due to the interaction. The proper physical theory should be expressed in terms of the masses and couplings that you measure in the laboratory.

In a certain class of theories this renormalization which in practice is a bit more complicated than just described, removes all the divergences. In other words, the only true divergences in the theory are associated with the shifts in masses and couplings due to the interaction. The remainder of the theory, f.ex. the energy dependence of a cross section, is finite. This class of theories is called renormalizable.

The way to recognize a renormalizable theory is by power counting. Let us consider an arbitrary graph consisting entirely of loops, for example

(9.1)

Since the divergences reside in the loops it is sufficient to study only such graphs. They are called "one-particle irreducible".

Let the graph have Q external quark lines, G external gluon lines and F external FFP ghost lines. The ghosts are present in the theory and should be treated as normal particles from the point of view of renormalization. Let $\mathcal{N}_1$ be the number of quark-quark-gluon vertices, $\mathcal{N}_2$ the number of ghost-ghost-gluon vertices and $\mathcal{N}_3$ and $\mathcal{N}_4$ the numbers of triple and quadruple gluon vertices. Let furthermore the number of internal lines be denoted Q', G' and F'.

If we cut all internal lines such that the graph decomposes into a collection of vertices we get the relations

$$Q + 2Q' = 2\mathcal{N}_1 \tag{9.2}$$

$$G + 2G' = \mathcal{N}_1 + \mathcal{N}_2 + 3\mathcal{N}_3 + 4\mathcal{N}_4 \tag{9.3}$$

$$F + 2F' = 2\mathcal{N}_2 \tag{9.4}$$

from which we may determine the numbers of internal lines.

The number of loops is the number of independent internal momenta

$$L = Q' + G' + F' - (\mathcal{N}_1 + \mathcal{N}_2 + \mathcal{N}_3 + \mathcal{N}_4 - 1) \tag{9.5}$$

Momentum conservation in each vertex removes one independent momentum, except for overall momentum conservation.

Finally we can count the powers. Each loop integral gives four powers of momentum in the nummerator while each triple gluon vertex and each ghost vertex gives one power. In the denominator the quark propagators give each one power while gluon and ghost propagators give two. Hence the superficial degree of divergence is

$$D = 4L + \mathcal{N}_2 + \mathcal{N}_3 - Q' - 2G' - 2F' \tag{9.6}$$

By means of eqs. (2) - (5) internal line numbers may be eliminated with the result

$$D = 4 - \frac{3}{2}Q - G - F \tag{9.7}$$

The $\mathcal{N}$-dependence fell out! This shows that the theory is renormalizable.

The superficial degree of divergence does not specify how the graph diverges. However, a convergent graph must have $D < 0$, not only for itself, but for all its subgraphs. This theorem was originally proven by Weinberg. Conversely, if just one subgraph or the graph itself has $D \geq 0$ then the graph diverges. If all subgraphs have $D < 0$ but the graph itself has $D \geq 0$ then it is called primitively divergent. The primitively divergent graphs are important because they form the basis from which all divergences may be constructed.

In the figure below all the diagrams that have $D \geq 0$ are exemplified. Except for the vacuum graphs, they are all primitively divergent

vacuum diagrams (D=4)

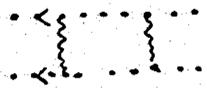
quark self-energy (D=1)

quark vertex (D=0)

ghost self-energy (D=2)

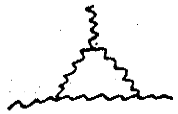
ghost vertex (D=1)

ghost-gluon scattering (D=0)

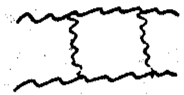


ghost-ghost scattering (D=0)

gluon self-energy (D=2)



gluon vertex (D=1)



gluon-gluon scattering (D=0)

Ghost-gluon and ghost-ghost scattering appear formally in this list but are actually convergent due to the asymmetric nature of the ghost vertex. One of the vertices (where the arrow sticks out) does not involve the loop momentum and should not be counted. Similarly the ghost self-energy and ghost vertex diverge one power less than indicated.

From the above list we conclude that it is necessary to renormalize the quark mass and wave function, the quark-gluon coupling, the ghost wave function and coupling, and the gluon wave function and couplings. We expect that the ghost and gluon mass remain zero because of gauge invariance. The various non-vanishing renormalization constants are also related to each other by gauge invariance via the Ward identities, or as they are called in non-Abelian gauge theories, the Slavnov-Taylor identities.[11]

10. Massive QED

As far as we know the photon is massless. Strictly speaking it could have a very small mass of the order of 10^{-48} g.[12] It is therefore satisfying that QED can be formulated for massive photons, and that the limit of the photon mass M going towards zero is smooth and continuous.

Structurally massive QED is quite different from massless QED. A massive photon can have three different states of polarization all satisfying $k \cdot \varepsilon = 0$. Besides the two transverse ones ε^1_μ and ε^2_μ there is a third solution[13] to this equation:

$$\varepsilon^3_\mu = \left(\frac{|\vec{k}|}{M}, \frac{\omega}{M} \frac{\vec{k}}{|\vec{k}|} \right) \tag{10.1}$$

where $\omega = \sqrt{\vec{k}^2 + M^2}$. They correspond to photons with helicity ± 1 and 0.

Let us begin by considering the theory in which only physical photons are present (the unitary theory). Here the sum over polarizations is

$$\sum_{\lambda=1,2,3} \varepsilon^\lambda_\mu \varepsilon^\lambda_\nu = -g_{\mu\nu} + \frac{k_\mu k_\nu}{M^2} \tag{10.2}$$

because the right hand side is the projection matrix for the solutions to $k \cdot \varepsilon = 0$, normalized so its trace is -3 $(k^2 = M^2)$. Thus the probability for emission of a physical massive photon is

$$\sum_{\lambda=1,2,3} |\varepsilon^\lambda_\mu J^\mu|^2 = -\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{M^2}\right) J^\mu J^{\nu *} \tag{10.3}$$

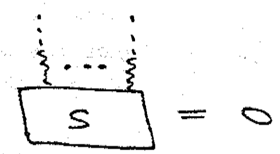
Similarly the propagator is

$$\mu \sim\sim\sim \nu = -i \frac{g_{\mu\nu} - \frac{k_\mu k_\nu}{M^2}}{k^2 - M^2} \tag{10.4}$$

QED with massive photons is manifestly covariant at the outset and there is no immediate need for introduction of unphysical states.

The power counting on the other hand goes wrong with the propagator (4) because it does not vanish as $1/k^2$ for $k \to \infty$. This means that the unitary theory formally is non-renormalizable.

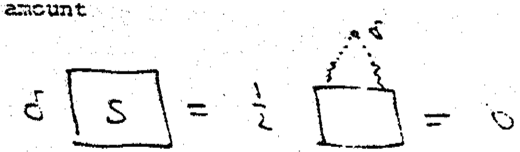
The special structure of QED, however, allows the construction of a renormalizable theory which is unitary in the physical subspace. Since the Ward identities for QED only derive from the electron vertex, via eq, (3.12), and not involves the photon propagator, we may introduce a single non-physical state with polarization k_μ, the gaugeon, which will decouple from the physical states. Apart from unimportant details the gaugeons have zero amplitude in the S-matrix



(10.5)

The gaugeons need not even be on-shell. The only condition is that all the other particles - not pictured here, - are on-shell and physical. This is true for any photon propagator in particular the massive one (4).

If we change the photon propagator by adding an infinitesimal $k_\mu k_\nu$-piece, it will change the S-matrix by the amount



(10.6)

which vanishes because of (5). Hence we may drop the $k_\mu k_\nu$ terms in (4) without changing the S-matrix. The theory has now become manifestly renormalizable and the S-matrix is the same as before in the physical subspace.

It is often convenient to use the propagator

$$\underset{\mu\qquad\nu}{\sim\!\sim\!\sim\!\sim} = -i\,\frac{g_{\mu\nu}-k_\mu k_\nu/M^2}{k^2-M^2} - i\,\frac{k_\mu k_\nu/M^2}{k^2-aM^2}$$

$$= -i\,\frac{g_{\mu\nu}-(1-a)\,k_\mu k_\nu/k^2-aM^2}{k^2-M^2} \qquad (10.7)$$

which, as the last line shows, has acceptable behaviour. It corresponds to giving the gaugeon a mass $\sqrt{a}\,M$. Here a is a real (positive) parameter. For $a\to\infty$ the unitary propagator (4) is obtained, as well as for $k^2\approx M^2$. For $M=0$ it reduces to the propagator (3.26). The parameter a is called the gauge parameter and all the gauges $0\leq a<\infty$ are manifestly renormalizable. For $a\to\infty$ we arrive at the unitary gauge. Since the S-matrix elements are independent of a, the theory is both renormalizable and unitary.

Finally we shall explicitly demonstrate that the transition to the massless limit is continuous. Writing

$$\varepsilon^3_\mu = \frac{|\bar{k}|}{M\omega}k_\mu + \frac{M}{|\bar{k}|\omega}(0,\bar{k}) \qquad (10.8)$$

we see that the first term is a gaugeon which decouples while the second vanishes for $M\to 0$. Hence the amplitude for emission of a physical longitudinal photon with mass zero is zero. In the limit $M\to 0$ there are only two physical states.

Concluding we can say that in massless QED the introduction of non-physical particles was necessitated by the demand for manifest covariance. The decoupling of the

gaugeons was absolutely necessary for this demand to be fulfilled. In massive QED manifest covariance was present from the beginning, but the theory was not manifestly renormalizable. The decoupling of the gaugeons permitted the theory to be cast in a manifestly renormalizable form without change to its physical contents.

11. Why massive QCD fails

We now turn to non-Abelian gauge theory and ask whether the same procedure can be applied in this case; in other words whether a gluon mass can be introduced "by hand". We shall choose a manifestly renormalizable gauge (10.7) with gluon propagator

$$\overset{a}{\underset{\mu}{}}\sim\sim\sim\overset{b}{\underset{\nu}{}} = -i\delta_{ab}\left[\frac{g_{\mu\nu}-k_\mu k_\nu/M^2}{k^2-M^2}+\frac{k_\mu k_\nu/M^2}{k^2-\alpha M^2}\right] \qquad (11.1)$$

This corresponds to giving the three physical gluon states a mass M and the single unphysical gluon state, the gaugeon with polarization k_μ, a mass $\sqrt{\alpha}\,M$. We shall use the same vertices for quarks, gluons and ghosts as in the massless case.

In order for the theory to be unitary it is necessary that the total production probability for unphysical particles, the gaugeons and the ghosts, vanishes. When this is the case we expect that the production probability for physical particles will be independent of the gauge parameter, α. In principle we can then let $\alpha \to \infty$ and obtain a manifestly unitary theory. But as we shall see the theory fails much earlier because the unphysical probabilities do not cancel.

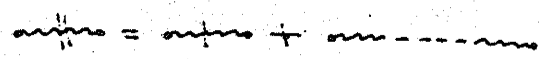
Let us first investigate whether a single gaugeon decouples in quark-gluon scattering. The basic Ward identity for quarks (4.5) remains unchanged. It is also clear that eq. (4.14) is unchanged. Adding and subtracting $-M^2 C_{a_1a_2a_3}$ to the two terms on the right hand side we obtain again (4.17), but with

$$\sim\!\overset{|}{\underset{||}{}}\!\sim\; = i\,\delta_{ab}\left(g_{\mu\nu}(k^2-M^2)-k_\mu k_\nu\right) \qquad (11.2)$$

This is the inverse of the unitary propagator and vanishes when acting on on-shell physical gluons because they have $k^2=M^2$ and $k\cdot\varepsilon=0$. Similarly the inverse of (1) is

$$\sim\!\overset{|}{}\!\sim\; = \delta_{ab}\left(g_{\mu\nu}(k^2-M^2)+\left(\tfrac{1}{\alpha}-1\right)k_\mu k_\nu\right) \qquad (11.3)$$

as may easily be verified. The relation (4.19) now again takes the form

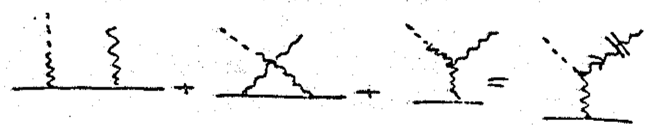


(11.4)

where now

$$\overset{a}{\bullet}\text{-----}\overset{b}{\bullet} = -\frac{i}{k^2-\alpha M^2}\,\delta_{ab} \qquad (11.5)$$

With these preparations it is easy to see that for on-shell quarks we have the one-gaugeon amplitude



(11.6)

because the coupling constants satisfy the Lie condition (4.22). This shows that if the incoming gluon is physical and on-shell, single gaugeon amplitude vanishes in quark gluon scattering

$$P_G = 0 \qquad (11.7)$$

Hence no problem arises here. Notice that this result is independent of the mass of the gaugeon and thus in particular valid for a gaugeon mass $\sqrt{\alpha}\,M$. Similarly one may derive - as in section 7 - that the 4-gluon vertex guarantees that a single gaugeon is not produced in gluon-gluon scattering.

Turning to quark-quark annihilation, problems arise. Since the physical gluons have a mass which is different from the gaugeon mass (when $a \neq 1$), we can physically distinguish between gluons and gaugeons by measuring their mass. According to (10.3) the probability for observing physical gluons is

$$P_{PP} = \frac{1}{2} \sum_{\text{Physical}} \left| \varepsilon_1^{\mu_1} \varepsilon_2^{\mu_2} J^{a_1 a_2}_{\mu_1 \mu_2} \right|^2$$

$$= \frac{1}{2} \left(g^{\mu_1 \nu_1} - \frac{k_1^{\mu_1} k_1^{\nu_1}}{M^2}\right)\left(g^{\mu_2 \nu_2} - \frac{k_2^{\mu_2} k_2^{\nu_2}}{M^2}\right)$$

$$\cdot\ J^{a_1 a_2}_{\mu_1 \mu_2} J^{a_1 a_2}_{\nu_1 \nu_2}{}^{\times} \qquad (11.8)$$

Since the physical gluons and the quarks are on-shell a change of the gauge parameter, a, will not affect this probability, because that only changes the $k_\mu k_\nu$-terms in the gluon propagator.

From (6) we conclude that the probability for producing a single gaugeon together with a physical gluon in quark annihilation is zero

$$P_{PG} = P_{GP} = 0 \qquad (11.9)$$

Writing eq. (6) in the form (6.2) (using eq. (2)) we get

$$k_1^{\mu_1} \varepsilon_2^{\mu_2} J^{a_1 a_2}_{\mu_1 \mu_2} = C_{a_1 a_2 a_3} \left(\varepsilon_2^{\mu_3} (k_2^2 - M^2) - \right.$$
$$\left. - k_2^{\mu_3} \varepsilon_2 k_2 \right) J^{a_3}_{\mu_3} \qquad (11.10)$$

Replacing ε_2 by k_2 we obtain the two-gaugeon amplitude

$$k_1^{\mu_1} k_2^{\mu_2} J^{a_1 a_2}_{\mu_1 \mu_2} = -M^2 C_{a_1 a_2 a_3} k_2^{\mu_3} J^{a_3}_{\mu_3} \qquad (11.11)$$

This amplitude does not vanish for $M \neq 0$ and we now expect trouble. For $M \to 0$ the right hand side vanishes at it should.

In calculating the probability P_{GG} for double gaugeon emission we must normalize the one-gaugeon states that have up to now been left unnormalized. Without going into the details, it follows from the propagator (1) that the properly normalized gaugeon-gaugeon probability is

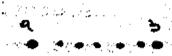
$$P_{GG} = \frac{1}{2} \left| \frac{k_1^{\mu_1}}{M} \frac{k_2^{\mu_2}}{M} J^{a_1 a_2}_{\mu_1 \mu_2} \right|^2 = \frac{1}{2} \left| C_{a_1 a_2 a_3} k_2^{\mu_3} J^{a_3}_{\mu_3} \right|^2 \qquad (11.12)$$

The factor $\frac{1}{2}$ is due to the statistical identity of the final gaugeons.

The gaugeons are unphysical but so are the FFP-ghosts and we should subtract the probability for production of a ghost pair (6.14)

$$P_{F\bar{F}} = \left| C_{a_1 a_2 a_3} k_2^{\mu_3} J^{a_3}_{\mu_3} \right|^2 \qquad (11.13)$$

Since we have assumed that the vertices are unchanged this must be the correct expression. If cancellation should be possible, the FFP ghosts must have the same mass as the gaugeons, i.e.



$$= -i \frac{\delta_{ab}}{k^2 - aM^2} \qquad (11.14)$$

Otherwise we could distinguish between gaugeons and ghosts by measuring their mass. The total unphysical probability is accordingly

$$P_{GG} - P_{F\bar{F}} = -\frac{1}{2} \left| C_{a_1 a_2 a_3} k_2^{\mu_3} J^{a_3}_{\mu_3} \right|^2 \qquad (11.15)$$

and this does not vanish![14]

The result (15) is cathastropic: The unphysical probability is not zero. Even if we don't know of any way of observing gaugeons and ghosts, the fact that all probabilities in some sense must sum up to 1 (unitarity), will nevertheless allow us to ascertain that something in being created when nothing should. Experimentalists would look for the missing mass.

We may understand this result in a fairly simple way. In the massless case there were two unphysical gluons, the scalar and longitudinal ones, and that required a certain FFP ghost contribution. In the massive case there is only one unphysical gluon, the gaugeon, and that needs only half the ghost contribution. This accounts for the left-over ghost contribution on the right hand side of eq. (15).

In the limit of the mass $M \to 0$ the amplitudes of the present theory go continuously over into the amplitudes of the massless theory, but the way we interpret the theory physically, by means of the probabilities, changes abruptly. The physical probability (8) can be written

$$P_{PP} = P_{TT} + P_{3T} + P_{T3} + P_{33} \tag{11.16}$$

where 3 refers to the physically longitudinal polarization (10.8). In the limit of $M \to 0$

$$P_{3T} \doteq P_{T3} \to 0 \tag{11.17}$$

because the gaugeon decouples from the transverse states. In the same limit we see from (10.8) that

$$P_{33} - P_{GG} \to 0 \tag{11.18}$$

Apart from terms that vanish for $M \to 0$, P_{33} is identical to P_{GG}. This shows explicitly that a part of the physical probability in the massive case becomes unphysical in the massless case.

Recapitulating, we have shown, that massive renormalizable QCD fails because of the discontinuous change in the number of physical and unphysical gluon states when going from the massless to the massive case. Since this change is deeply related to the structure of the representations of the Poincaré group it is impossible in a trivial way to introduce a mass in non-Abelian gauge theories. This leads to serious difficulties because it is then impossible to cut off the infrared divergences by means of a small mass as one does in QED. The infrared behaviour of QCD is still largely unknown and presumably related to the question of why quarks and gluons have not yet been observed, even if everything points to quarks as being the constituents of elementary particles.

Although there is no easy way of introducing a mass for the gluons there is a sophisticated approach which allows for the construction of theories with massive vector particles.

It is called the Higgs-Kibble mechanism and involves a spontaneous symmetry breakdown analogous to that which is responsible for the spontaneous magnetization of a ferromagnet. The theory of spontaneous symmetry breaking and in particular the Higgs-Kibble mechanism will be discussed by Zinn-Justin in his lectures. The basic idea - expressed in our language - is to introduce still more unphysical scalar particles (Goldstone Bosons) that will cancel the superfluons probability (15). They will for reasons of gauge invariance always be accompanied by physical scalar Bosons (Higgsons) and this changes the theory considerably. We shall not go further into these questions here.

12. Renormalization

In this section we return to the massless theory defined in sections 3-9. We shall study the divergences at the one-loop level and show that they can be removed by suitable counter terms. Throughout this section we assume that a regularization procedure exists which makes all integrals finite without disturbing the naive combinatiorical Ward identities.[15)]

a) The quark self-energy

According to the table in section 9 the quark self-energy is linearly divergent and will become convergent if we differentiate twice after the external momentum. Hence the divergent part must be a polynomial of first order in the momentum, so that we can write

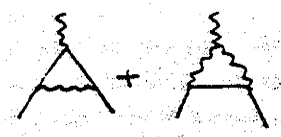
$$= -i(A + \not{p}B)\delta_{rs} + \text{finite} \qquad (12.1)$$

where A and B are divergent constants. We shall assume that the quarks belong to an irreducible representation and the gluons to a simple component of the algebra. From the discussion in section 5 we know that we may consider each irreducible representation and each simple component by itself. This is the reason for the δ_{rs} in eq. (1) because according to Schuurs Lemma, that is the only invariant matrix. Explicitly the algebraic factor in (1) is $T_a T_a$ which evidently is invariant.

The separation between the finite and infinite part in eq. (1) is arbitrary. We may, for example, shift the finite parts of A and B around to obtain, that the finite remainder in (1) and its derivative vanishes at some momentum. We shall return to this question in the next section.

b) The proper quark vertex

According to the table in section 9 we see that the proper quark vertex corrections are logarithmically divergent and hence of the form

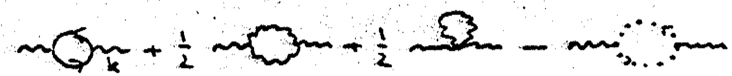


$$= -iL\, T_a \gamma_\mu + \text{finite} \qquad (12.2)$$

Again the Lie algebra allows the factor T_a and the Lorentz invariance only γ_μ.

c) The gluon self-energy

It is quadratically divergent and of the form



$$= iC(g_{\mu\nu}k^2 - k_\mu k_\nu)\delta_{ab} + \text{finite} \qquad (12.3)$$

The reason that the second degree divergent polynomial has this special form is the decoupling of the gaugeon which was demonstrated in section 8. The divergent part must itself decouple and hence (3) is the only possible form.

The constant C is in fact only logarithmically divergent.

d) The ghost self-energy

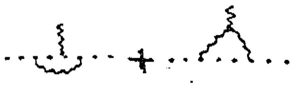
A priori it is quadratically divergent, but because one momentum "sticks out" at the arrow, the actual integral is only linearly divergent. Lorentz invariance then forces the form

$$= i\tilde{B}\, k^2 \delta_{ab} + \text{finite} \qquad (12.8)$$

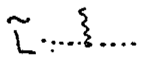
The constant $\tilde{B}$ is only logarithmically divergent.

e) The ghost vertex

Here the form of the divergence is



$$= \tilde{L}$$

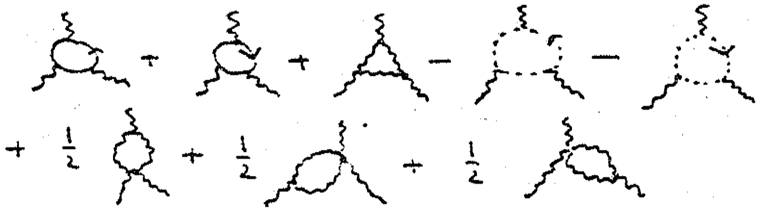


$$+ \text{finite} \qquad (12.9)$$

where $\tilde{L}$ is logarithmically divergent.

f) The triple gluon vertex

There are 8 diagrams contributing in this case



$$= D$$

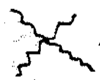
$$+ \text{finite} \qquad (12.10)$$

This follows from fairly simple analysis but we shall refrain from doing it here. The constant D is logarithmically divergent.

g) The quadruple gluon vertex

There are 21 diagrams that contribute here. Again a closer analysis shows that

$$= E$$



$$+ \text{finite} \qquad (12.11)$$

where E is logarithmically divergent.

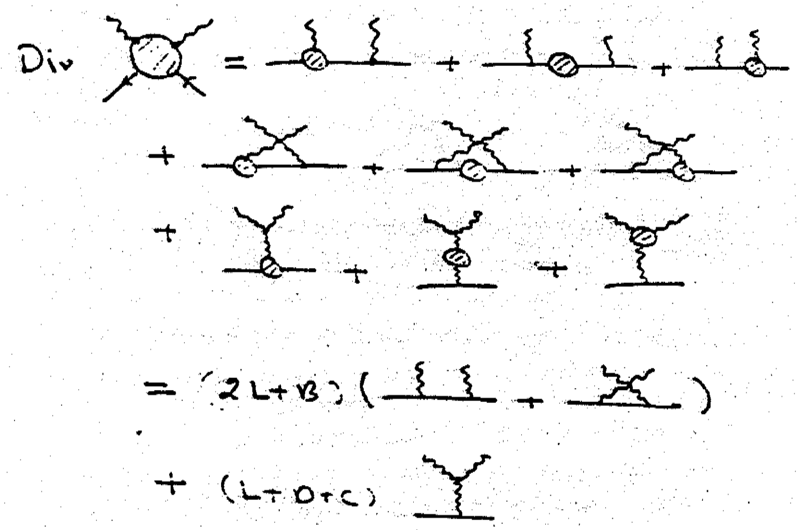
What is the effect of these divergences. Let us look at eq. (1). The corrections to the quark propagator are

$$= \frac{i}{\not p - m}\left\{1 + \frac{A + \not p B}{\not p - m}\right\} + \text{finite}$$

$$\approx \frac{i}{\not p - m - A - mB} + \frac{iB}{\not p - m} + \text{finite} \qquad (12.12)$$

The last equation is correct to first order in A and B. This shows that the quark mass is changed from m to $m + A + Bm$. So if we assume that the bare quark mass is divergent in such a way as to make $m + A + mB$ finite, then the first term will be finite. This is the essence of mass renormalization. We shall get back to the interpretation of the other divergences in a moment.

The divergences are not independent, but are related by the Slavnov-Taylor identities. They may be derived directly from the Ward identities by means of naive combinations, but we shall derive them here in a slightly indirect way. Let us for example consider the divergent part of quark-gluon scattering leaving out corrections to the external legs



(12.13)

Here we have used the explicit form of the divergences established above and we have left out the mass renormalization term. It will simply correct the mass of the quark propagator in the lowest order diagrams. Also we have dropped longitudinal terms in the gluon propagator. They do not contribute for on-shell gluons.

We have not exactly proven before that the loop corrections to quark gluon scattering guarantee the decoupling of the gaugeons, but this may be proven by general means. Hence the divergent part must by itself decouple the gaugeons. The only way this can happen is if

$$2L + B = L + D + C \tag{12.14}$$

because then the lowest order diagrams are reproduced in (13). This is the first Slavnov-Taylor identity. Notice, that the corrections to the external lines do not change this result.

In an exactly similar way we obtain from ghost-gluon scattering that

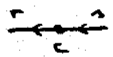
$$2\tilde{L} + \tilde{B} = \tilde{L} + D + C \tag{12.15}$$

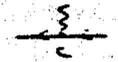
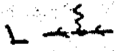
Finally by considering the divergent part of gluon-gluon scattering we obtain

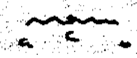
$$2D + C = E \tag{12.16}$$

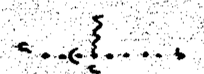
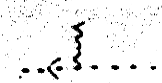
Eqs. (14) - (16) are all the one-loop Slavnov-Taylor identities.

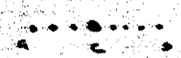
Instead of trying to interpret the divergences physically as corrections to various parameters it is more convenient to follow our practice of defining new vertices to repair the theory. Let us introduce the basic vertices

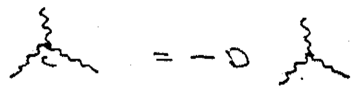
 $= i\,(A + \not p\, B)\,\delta_{rs}$ (12.17)

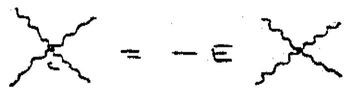
 $= i\,L\,T_a\,\gamma_\mu = -L$  (12.18)

 $= -i\,C\,(g_{\mu\nu}k^2 - k_\mu k_\nu)\,\delta_{ab}$ (12.19)

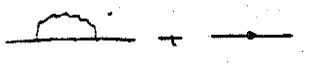
 $= -\tilde{L}$  (12.20)

 $= -i\,\tilde{B}\,k^2\,\delta_{ab}$ (12.21)


$= -D$ (12.22)

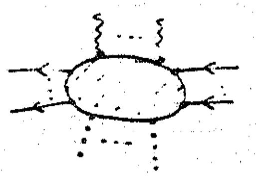

$= -E$ (12.23)

The little c's on the left hand side indicates that these vertices are compensating the divergences. Since A and B are of second order in the coupling the self-energy calculated to second order is now


$=$ finite (12.24)

By this procedure the radiative corrections have all been made finite to lowest non-trivial order. All this seems rather artificial but the interesting thing is that this procedure can be carried through to arbitrary high orders of perturbation theory. The only change is that the constants in the counterterms (17) - (23) get higher order corrections. The proof of this assertion is straight-forward but laborious and is standard renormalization theory. We shall not go through it here.

Let us now assume that the counter terms have been determined such that the theory is finite order by order in perturbation theory, and let us ask about the relationship between the theory with and without counter terms. It turns out that this relationship is quite simple. Consider the amplitude for a general process

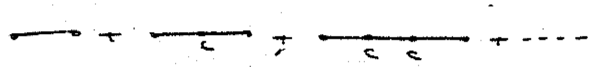
$\Gamma_c =$

(12.25)

Here we only consider diagrams that cannot be split into two parts connected by a single propagator (one-particle irreducible graphs). The general graphs can evidently be built up from these. The index c indicates that counterterms are included so that Γ_c is finite order by order in the coupling. In keeping with the notation of section 9 there are

Q external quark lines, G external gluon lines, and

F external FFP ghost lines.

If we reorder the perturbation series it is seen that each quark propagator is accompanied by a series of corrections of the type (17)



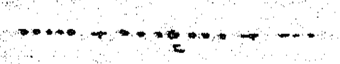
$$= \frac{i}{(1+B)\not{p} - m + A} \qquad (12.26)$$

The last result follows by summing up the geometric series. This means that the only effect of the vertex (17) is to replace all free quark propagators by $\frac{i}{Z_2(\not{p} - m_0)}$ where

$$Z_2 = 1 + B \qquad (2.27)$$

$$m_0 = \frac{m - A}{1 + B} \qquad (2.28)$$

Similarly the ghost propagator becomes

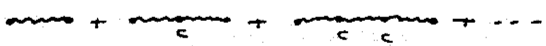


$$= -\frac{i}{k^2}\frac{1}{1+\tilde{B}} = -\frac{i}{\tilde{Z}_2 k^2} \qquad (2.29)$$

where we have put

$$\tilde{Z}_2 = 1 + \tilde{B} \qquad (2.30)$$

The gluon propagator is a bit more complicated. It becomes



$$= -i\,\frac{g_{\mu\nu}-\frac{k_\mu k_\nu}{k^2}}{(1+C)k^2} - i a \frac{k_\mu k_\nu}{(k^2)^2} \tag{12.31}$$

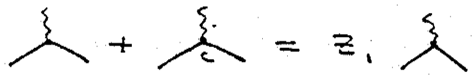
$$= \frac{1}{Z_3}\left(-i\,\frac{g_{\mu\nu}-\frac{k_\mu k_\nu}{k^2}}{k^2} - i a_0 \frac{k_\mu k_\nu}{(k^2)^2}\right)$$

where we have put

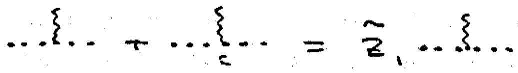
$$Z_3 = 1 + C \tag{12.32}$$

$$a_0 = a Z_3 \tag{12.33}$$

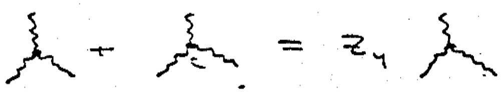
The vertices are very simple. We have



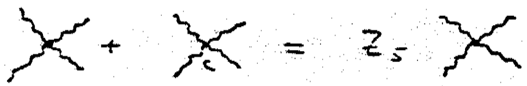
(12.34)



(12.35)



(12.36)



(12.37)

where we have put

$$Z_1 = 1 - L \tag{12.38}$$

$$\tilde{Z}_1 = 1 - \tilde{L} \tag{12.39}$$

$$Z_4 = 1 - D \tag{12.40}$$

$$Z_5 = 1 - E \tag{12.41}$$

From this analysis we see that the amplitude (25) with counterterms included can be calculated from the amplitude Γ_0 without counterterms, but using the propagators (26), (29) and (31) and the vertices (34) - (37). This may be formulated a little simpler in the following way. The compensated amplitude is a function $\Gamma_c(p, a, m, g, g, g^2, g)$ of the various momenta, collectively denoted p, the gauge parameter a, the mass parameter m, and the various coupling constants. Because we consider only a simple Lie algebra there is just one free factor in T_a which we call g. The triple gluon coupling C_{abc} is also proportional to g. The quadruple gluon coupling is g^2 apart from numerical factors and the ghost-gluon coupling is g apart from numerical factors.[16)]

With this notation we can write down the relationship between the compensated and uncompensated amplitudes in the precise form

$$\Gamma_c(p, a, m, g, g, g^2, g)$$
$$= (\sqrt{Z_2})^{Q} (\sqrt{Z_3})^{G} (\sqrt{\tilde{Z}_1})^{F}$$
$$\Gamma_0(p, a_0, m_0, g_0, g_1, g_2^2, g_3) \tag{12.42}$$

The quantity Γ_0 on the right hand side is calculated without compensating terms using a quark propagator $i/(\not p - m_0)$ a gluon propagator $-i(g_{\mu\nu}-k_\mu k_\nu/k^2)/k^2 - i a_0 k_\mu k_\nu/(k^2)^2$

where m_0 and a_0 are given by (28) and (33), and the coupling constants

$$g_0 = \frac{Z_1}{Z_2} \frac{1}{\sqrt{Z_3}} g \tag{12.43}$$

$$g_1 = \frac{Z_4}{(\sqrt{Z_3})^3} g \tag{12.44}$$

$$g_2^2 = \frac{Z_5}{(\sqrt{Z_3})^4} g^2 \tag{12.45}$$

$$g_3 = \frac{\tilde{Z}_1}{\tilde{Z}_2} \frac{1}{\sqrt{Z_3}} g \tag{12.46}$$

What we have done here is to take the factors $1/Z_2, 1/\tilde{Z}_2$ and $1/Z_3$ in the propagators (26), (29) and (31) and associating $1/\sqrt{Z_2}, 1/\sqrt{\tilde{Z}_2}$ and $1/\sqrt{Z_3}$ with each vertex at the end of the propagator. This has the effect of changing all the couplings as shown in the list above. The factors in front of Γ_0 in eq. (42) arise because the external lines have no propagators associated with them to give $1/\sqrt{Z}$ factors.

The effect of the counterterms is thus completely accounted for by eq. (42). One problem remains. Since the counterterms compensate the divergences, and since the gaugeons decouple from the divergences, they must decouple from the finite, compensated, amplitudes. We have at least demonstrated this in the lowest order. But the decoupling of the gaugeons only affects the momenta on the left hand side, so the gaugeons should also decouple from Γ_0 on the right hand side. This is only possible if the coupling constants are identical,

$$g_0 = g_1 = g_2 = g_3 \tag{12.47}$$

Without this condition the Lie algebra would not even close. The conditions (47) are equivalent to

$$\frac{Z_1}{Z_2} = \frac{\tilde{Z}_1}{\tilde{Z}_2} = \frac{Z_4}{Z_3} = \sqrt{\frac{Z_5}{Z_3}} \tag{12.47}$$

These are the generalized Slavnov-Taylor identities.[17] Expanding to first order in the counterterms using the definitions of the Z's they are seen to be equivalent to eqs. (14) - (16).

We can finally rewrite eq. (42) as

$$\Gamma_c(p, a, m, g) = Z_2^{\frac{1}{2}Q} \tilde{Z}_2^{\frac{1}{2}F} Z_3^{\frac{1}{2}G} \Gamma_0(p, a_0, m_0, g_0) \tag{12.48}$$

making explicit that there is only one coupling constant on both sides. Γ_0 is now simply the original amplitude without counterterms buth with different parameters a_0, m_0 and g_0. The $\sqrt{Z}$ factors outside can be viewed as a renormalization of the wave functions of the external particles.

Although the left hand side of (48) is finite m is not the "physical" mass of the quarks. But the relation between the physical mass, the pole in the propagator, and m is a finite relationship, $m_{ph} = f(m, g)$. Although (48) is renormalized it is not what is called the on-shell renormalized amplitude. Finite renormalizations may still have to be performed.

13. The Callan-Symanzik equations

The unrenormalized amplitudes $\Gamma_0(p, a_0, m_0, g_0, \Lambda)$ actually depend on some parameters, Λ, associated with the regularization procedure, which cuts off the divergences. It may, for example, be a large mass characterizing the cut-off in momentum or it may be $1/n-4$, where n is the dimension of space-time. For $\Lambda \to \infty$ the divergences reappear in the theory. The renormalization constants Z_i also depend on Λ but the renormalized amplitudes (12.48) are finite in the limit of $\Lambda \to \infty$. That this happens requires a great conspiracy between the renormalization constants and the unrenormalized amplitudes, and we may ask the question of whether all vestiges of this conspiracy have been lost in the renormalized amplitudes. The answer is no, it has not. The asymptotic behaviour for $p \to \infty$ of the renormalized amplitudes is restricted in a way which is deeply connected with the renormalizability of the theory.

Instead of studying the asymptotic behaviour for $p \to \infty$ we shall look at the behaviour of Γ_0 for $m_0 \to 0$. This must be very close to being the same thing. If Γ_0 diverges for $m_0 \to 0$, we say that a mass singularity occurs at $m_0 = 0$. The theory of mass singularities is well understood.[18] The result is, essentially that one can use naive power counting for mass singularities (looking for the simultaneous divergence of propagators for small loop momenta). So, provided the external momenta are in the Euclidean region, where they themselves and all partial sums of momenta are space-like, then mass singularities do not occur. We can even say more. If we expand the bare electron propagator in powers of m_0 we get

$$\frac{i}{\not p - m_0} = \frac{i}{\not p} + \frac{i m_0}{\not p^2} + \frac{i m_0^2}{\not p^3} + \frac{i m_0^3}{\not p^3(\not p - m_0)} \qquad (13.1)$$

In the Euclidian region we would naively expect that only the last term could give rise to a mass singularity $\int \frac{d^4p}{p^4}$ in the limit of $m_0 \to 0$. We would therefore (erroneously) be lead to believe that Γ_0 can be expanded to $O(m_0^2)$ without the appearance of $\log m_0$. This is, however, wrong. When the gluon propagator correction  is expanded to order m_0^2, it will for dimensional reasons behave like m_0^2/k^4, which gives rise to an infrared divergence in the integration over the gluon momentum. Hence we have only

$$\Gamma_0(p, a_0, m_0, g_0, \Lambda) = A_0(p, a_0, g_0, \Lambda) + m_0 B_0(p, a_0, g_0, \Lambda) + O(m_0^2 \log m_0) \qquad (13.2)$$

where A_0 and B_0 do not depend on m_0.

The renormalization constants are all only logarithmically divergent. Since they are dimensionless and only depend on the parameter m (because all subtraction points are scaled by m) they can only contain $\log m$. Any term of $O(m)$ would be of $O(m/\Lambda)$ and hence vanishing for $\Lambda \to \infty$. This means that we have

$$\Gamma_c(p, a, m, g) = A(p, a, m, g) + m B(p, a, m, g) + O(m^2 \log m) \qquad (13.3)$$

where A and B only depend on $\log m$ and are related to A_0 and B_0 by

$$A(p,a,m,g) = Z\,A_0(p,a_0,g_0,\Lambda) \quad (13.4)$$

$$B(p,a,m,g) = Z Z_m\,B_0(p,a_0,g_0,\Lambda) \quad (13.5)$$

$$Z = Z_2^{\frac{1}{2}G}\,\tilde{Z}_2^{\frac{1}{2}F}\,Z_3^{\frac{1}{2}G} \quad (13.6)$$

$$Z_m = m_0/m \quad (13.7)$$

Each of the functions A and B must be finite because they are the leading logarithmic and first subdominant terms in the expansion of the renormalized amplitude for small m. Another way of saying this is that the only way A and B can depend on $\log m$ is via the renormalization constants. This is the restriction which was mentioned above.

In differential form this can be stated in the following way. If we change the mass m infinitesimally the only way A and B change is via the dependence on the renormalization constants. There is no intrinsic dependence on m. Precisely

$$0 = Z\frac{\partial}{\partial m_0}A_0(p,a_0,g_0,\Lambda)$$

$$= \frac{\partial}{\partial m_0}(Z A_0) - \frac{\partial Z}{\partial m_0}A_0$$

$$= \left(\frac{\partial m}{\partial m_0}\frac{\partial}{\partial m} + \frac{\partial g}{\partial m_0}\frac{\partial}{\partial g} + \frac{\partial a}{\partial m_0}\frac{\partial}{\partial a} - \frac{\partial \log Z}{\partial m_0}\right)A$$

In differentiating m, g, a and Z after m_0 we keep all the unrenormalized parameters fixed. Dividing by $\partial m/\partial m_0$ and multiplying by m we get[19])

$$\left(m\frac{\partial}{\partial m} + \beta g\frac{\partial}{\partial g} + \beta' a\frac{\partial}{\partial a} + \gamma\right)A(p,a,m,g) = 0 \quad (13.8)$$

$$\left(m\frac{\partial}{\partial m} + \beta g\frac{\partial}{\partial g} + \beta' a\frac{\partial}{\partial a} + \gamma - \delta\right)B(p,a,m,g) = 0 \quad (13.9)$$

where the last equation is derived in a similar way from eq. (5). The coefficients are

$$\beta = K\frac{\partial \log g}{\partial \log m_0} \quad (13.10)$$

$$\beta' = K\frac{\partial \log a}{\partial \log m_0} \quad (13.11)$$

$$\gamma = -K\frac{\partial \log Z}{\partial \log m_0} \quad (13.12)$$

$$\delta = K\frac{\partial \log Z_m}{\partial \log m_0} \quad (13.13)$$

where

$$K = \frac{m/m_0}{\partial m/\partial m_0} = \frac{1}{\frac{\partial \log m}{\partial \log m_0}} \quad (13.14)$$

and we have expressed everything in terms of logarithmic derivatives. Using $Z_m = m_0/m$ in (13) we get from (14)

$$K = 1 + \delta \quad (13.15)$$

From (6) and (12) we get

$$\gamma = \frac{1}{2}G\gamma_2 + \frac{1}{2}F\tilde{\gamma}_2 + \frac{1}{2}G\gamma_3 \quad (13.16)$$

where the γ's are the logarithmic derivatives of the wave function renormalization constants. Using that $a = a_0/Z_3$ (eq. (12.33) and eq. (11) we get

$$\beta' = \gamma_3 \tag{13.17}$$

By varying G, F and G it follows from (8) and (9) that $\beta, \gamma_2, \tilde{\gamma}_2, \gamma_3$ and δ are all finite because they can be expressed as derivatives of A and B after the renormalized parameters.

Defining

$$Z_g = g_0/g \tag{13.18}$$

we find from (10)

$$\beta = -k \frac{\partial \log Z_g}{\partial \log m_0} \tag{13.19}$$

It is intuitively fairly clear that if the theory should make physical sense, Z_m and Z_g must be independent of the gauge parameter, a. Hence from (14) and (19) we get that β and δ are independent of the gauge parameter and since they are finite and dimensionless, they can only depend on g, the coupling constant. The other coefficients, $\gamma_2, \tilde{\gamma}_2,$ and γ_3 are all gauge dependent.

Following the standard arguments we define the effective coupling $\bar{g}(g,\lambda)$ by

$$\lambda \frac{d\bar{g}}{d\lambda} = \bar{g}\, \beta(\bar{g}) \tag{13.20}$$

$$\bar{g}(g, 1) = g \tag{13.21}$$

Using that $A(\lambda p, m, g) = \lambda^D A(p, m/\lambda, g)$ we may cast the Callan-Symanzik equations (8) and (9) into differential equations after λ and the solutions take the form (for $a = 0$)

$$A(\lambda p, m, g) = \lambda^D A(p, m, \bar{g}(g,\lambda))\, e^{\int_g^{\bar{g}(g,\lambda)} dx \frac{\gamma(x)}{x\beta(x)}} \tag{13.22}$$

$$B(\lambda p, m, g) = \lambda^{D-1} B(p, m, \bar{g}(g,\lambda))\, e^{\int_g^{\bar{g}(g,\lambda)} dx \frac{\gamma(x) - d(x)}{x\beta(x)}} \tag{13.23}$$

where D is the superficial degree of divergence (9.7). So, apart from the overall scaling factors λ^D and λ^{D-1} the whole dependence on λ lies in the effective coupling constant, $\bar{g}(g,\lambda)$. We have used the Landau gauge $a = 0$ in order to avoid problems with the gauge dependence.

The β-function can be calculated in perturbation theory

$$\beta(g) = \beta_1 g^2 + \beta_2 g^4 + \cdots \tag{13.24}$$

The leading coefficient is for a simple Lie algebra[20)]

$$\beta_1 = \frac{1}{16\pi^2}\left(\frac{4}{3} C^T - \frac{11}{3} C\right) \tag{13.25}$$

where C^T and C are the coefficients defined in eqs. (5.15) and (5.16). So the sign of β_1 depends on the index of the representation C^T/C. For $SU(N)$ we have $C = N$. The value of C^T depends on which representations the quarks belong to. If we have m fundamental N-dimensional multiplets (in $SU(3)$ triplets of quarks) and n adjoint representations of dimension $N^2 - 1$ (octets in $SU(3)$) then $C^T = m\frac{1}{2} + n \cdot N$

This shows that for sufficiently many quarks β_1 will always become positive. In $SU(2)$, β_1 is negative for $m + 4n < 11$, i.e. we can have 10 quarks doublets and no triplets, doublets and one triplet or two doublets and two triplets. In $SU(3)$, β_1 is negative for

$2m+12n < 33$ so that we can have 16 quark triplets and no octets, 10 quark triplets and one octet, or four triplets and two octets.

Using only the lowest order approximation to β we get from (20) and (21)

$$\bar{g}^2 = \frac{g^2}{1-2\beta_1 g^2 \log\lambda} \tag{13.26}$$

If β_1 is positive and λ grows from 1 towards infinity we see that for $\log\lambda = 1/2\beta_1 g^2$ this expression diverges. This means that the formulas (22) and (23) are of no use because we do not know what happens to $A(p,m,\bar{g})$ when $\bar{g}$ becomes infinite. On the other hand, if β_1 is negative, then no singularity is met and $\bar{g}^2$ decreases steadily towards zero for $\lambda \to \infty$. As seen from (22) and (23) the asymptotic limit becomes calculable because perturbation theory is valid for small coupling constants. This justifies in turn the use of the lowest order approximation for β. The exponential factors in (22) and (23) only vary slowly for $\lambda \to \infty$.

Since the effective coupling constant vanishes in the asymptotic region the theory is called asymptotically free. It is an odd fact that among all the "normal" field theories the only way asymptotic freedom can be obtained is by including sufficiently many non-Abelian gauge fields. There seems to be no simple argument which will explay why β_1 is negative in gauge theories. That is just the way it is.

References and Footnotes

1. A four-vector is written $x = (x_0, \vec{x})$ and the scalar product is $x \cdot y = x_0 y_0 - \vec{x}\cdot\vec{y}$. Generally the notation of Bjorken and Drell (Relativistic Quantum Mechanics) is used.
2. Conventional QCD has three quark colours and eight gluon colours for each quark flavor. We shall not put such a limitation on the following discussion.
3. With the convention that k_μ in ---~~~ always runs from the "dashes to the waves", the last term in (4.19) is decomposed in the form
   $$\text{---}\sim\sim\sim = (-k_\mu)\,\frac{(-i)}{k^2}\,(k_\nu)$$
4. For example, J.E. Humphreys, Introduction to Lie Algebras and Representation Theory (Springer, 1972).
5. This is the negative of what is usually called the Cartan-Killing form.
6. R.P. Feynman, Acta Phys. Polon. 24, 297 (1962).
7. L.D. Fadeev and V.N. Popov, Phys.Letters 25B, 29 (1967).
8. Perhaps one should call the FFP ghosts for ghouls in order to distinguish them from the metric ghosts. They eat dead probability.
9. For a more general proof along these lines see G. 't Hooft, Nucl.Phys. B33, 173 (1971).
10. G. 't Hooft and M. Veltman, Diagrammar (CERN 73-9).
11. See section 12.
12. A.S. Goldhaber and M.N. Nieto, Sci.Am. 234, 86 (1976).
13. This should not be confused with (L) $\varepsilon^\lambda_\mu = (0, \vec{k}/\omega)$ introduced in the massless case although we use the same notation.

14. H. van Dam and M. Veltman, Nucl.Phys. <u>B22</u>, 397 (1970).

15. E.S. Abers and B.W. Lee, Phys.Rep. <u>96</u>, 1 (1973).

16. In principle we could calculate the Green functions with four different couplings. For the discussion below it is convenient to make explicit the dependence on the four couplings.

17. J.C. Taylor, Nucl.Phys. <u>B33</u>, 436 (1971)
A. Slavnov, Theor. and Math.Phys. <u>10</u>, 99 (1972).

18. T. Kinoshita, J.Math.Phys. <u>3</u>, 650 (1962).

19. These are the Callan-Symanzik equations.

20. H.D. Politzer, Phys.Rep. <u>14C</u>, 129 (1974).

12.76.200
3.77.100
5.77.100
2.78.50
9.79.50