16

Strong interactions and field theory

16.1 Overview

16.1.1 Phenomenological approach to hadron scattering

We have considered the phenomenological theory of strong interactions. In the last 20 years the theory was developing in two parallel directions. The questions were:

- how the interaction takes place in a relativistic situation, irrespectively to *who* is interacting;
- how the mass spectrum of hadrons is built up; what type of laws can be extracted not knowing the microscopic dynamics?

Qualitatively, high-energy processes are understood. Such general consequences of the complex angular momentum theory as the uniformity of particle production in rapidity, the shrinkage of the diffractive cone and the logarithmic multiplicity growth, are in good qualitative agreement with the experiment. Why is there no good *quantitative* agreement?

In the expression for the Regge radius,

$$R^2 = R_0^2 + \alpha' \ln \frac{s}{m^2} \,,$$

the pomeron slope – the impact parameter diffusion coefficient – turned out to be *numerically small*:

$$\alpha' / R_0^2 \sim 1/12 \,.$$

The hadron radius increases very slowly with the energy, and hence different simplifying properties that we have discussed in these lectures

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

will never manifest themselves; the true high-energy asymptotics will never be reached. We do not understand the *reason* why this is so.

Nevertheless, the *qualitative* agreement exists, since, although the growth of the radius is not sufficient, a relatively homogeneous distribution in rapidity is established in nature.

In what concerns the possible structure of the mass spectrum of hadrons, exploring this problem has led to the notion of 'duality'.*

16.1.2 Duality

Take some field theory based on a certain number of interacting fields (bare particles). If the coupling constant is sufficiently large, bound states (and resonances) will appear and enter the theory on equal footing with the input particles. A complicated, but legitimate, structure of the physical particle spectrum will emerge, driven by unitarity and analyticity.

We know that each Regge pole in the j-plane feels unitarity thresholds which, generally speaking, essentially deform its trajectory. However, experimentally this is not true. In a large interval of t all known Regge trajectories are *linear* with quite a good accuracy.

This fact looks rather strange and calls for explanation. How is it that all along a distance of several GeV the linearity persists?

One possibility is that a large mass scale may be there, hidden at the quark level (quark masses?)† the fact that it is mesons and baryons – bound states of quarks – that propagate at macroscopic distances is a secondary thing, less important for the dynamics.

Another attempt to explain the linearity consists of introducing into the theory, from the very beginning, an infinitely large number of particles and connecting their spins (placing them on the Regge trajectories). This way we make a step beyond the QFT framework not by abandoning locality but by introducing an infinite number of fields.

Why does this lead to linear trajectories specifically?

We know that all the singularities (including trajectories) are determined by unitarity, i.e. by the interaction. Hence, to insert, e.g. a $\sqrt{t}$ singularity by hand, does not seem clever or harmless.

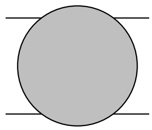
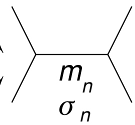
The simplest analytic choice is a straight line.

* The concept of duality gave birth to string theory.

† On the appearance of a large mass scale in the vacuum channel, see Shifman and Vainshtein, 2005 (ed.).

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

Let us approximate the scattering amplitude by the sum over s-channel particle poles:

$$A(s,t) =$$

$$= \sum_{n}^{\infty}$$


$$+ \text{ corrections.} \qquad (16.1)$$

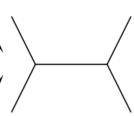
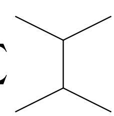
One could think that the hadron interaction that we observe is *strong* already because there are many particles to exchange in (16.1). Within this logic, one might hope the corrections to such 'Born' approximation to be small.

Did we draw in (16.1) everything we had to? Should we not include also the poles in the t- and u-channels?

This would have been necessary if the number of particles was finite, or, at least, if the series in (16.1) converged well. However, the situation here is not so simple: since we are planning to include particles with arbitrarily high spins σ_n, the behaviour of $P_\sigma(\cos\theta_s)$ at $\sigma \to \infty$ is a worry. At $t > 0$ ($\cos\theta_s > 1$) the Legendre polynomials P_σ *increase* as a power of σ. Hence, the series makes sense only in the physical region of the s-channel ($t < 0$) and diverges at $t > 0$.

What does the divergence of the series mean for the amplitude $A(s,t)$? It is just a singularity in t. We know, however, that all the singularities are either poles or thresholds. We want to construct the Born amplitude which does not take into account the interactions; consequently, there may be only poles. Thus, we are unable to write an amplitude that would have poles *only* in s.

Here lies the basic idea of *duality*: the requirement that the series has in t just the necessary poles, those corresponding to particles (resonances) that can be exchanged in the t-channel of the reaction. The scattering amplitude can be alternatively expanded in terms of the t-channel poles:

$$A(s,t) = \sum_{n}^{\infty}$$

$$= \sum_{n}^{\infty}$$

$$. \qquad (16.2)$$

The equality of the sums over s- and t-channel resonances (*duality* of the two representations) guarantees the self-consistency of the construction. Essentially, the duality idea is the only way to introduce an infinite number of particles in the theory.

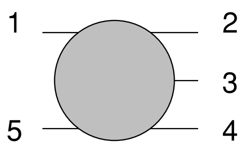
Obviously, the duality relation (16.2) imposes severe restrictions on the possible structure of the particle spectrum. How can we write a suitable meromorphic function having only poles? As it turns out, the problem

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

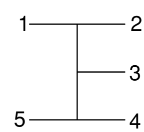
can be solved quite easily (Veneziano, 1968):

$$\begin{aligned} A(s,t) &= \frac{\Gamma(-\alpha(s))\Gamma(-\alpha(t))}{\Gamma(-\alpha(s)-\alpha(t))} \\ &= B(-\alpha(s),-\alpha(t)) \equiv \int_0^1 dx(1-x)^{-\alpha(s)-1}x^{-\alpha(t)-1} \end{aligned} \tag{16.3}$$

where α is a linear trajectory: $\alpha(t) = t$ (in units $\alpha' = 1$). This function has poles in each integer point both in the s- and t-channels. It is clear that having written a function symmetric in s and t, I have satisfied the duality relation automatically.



Actually, one can go even further and construct, in particular, a $2 \to 3$ scattering amplitude. It contains many pair invariants; continuing into each particular channel, $s_{ik} > 0$, we must obtain a series of the corresponding resonance poles.



It is easy to draw a diagram having poles, say, in s_{12} and s_{45}. Can one invent a contribution that would have, in addition, poles in s_{35}? It is clear that this is impossible: as soon as particles 4 and 5 combine into a pole, there is no information left about the momentum p_5 that could correlate with p_3 (only information about spins can be carried up the graph). As a result, dual multi-particle amplitudes consist of a sum of all possible tree diagrams with poles in all sub-channels.

There is a technical difficulty: the amplitude must factorize; moreover, the residues must be *positive*. The task of constructing the meson amplitude close to the real hadron spectrum is almost completed. The case of baryons turns out to be more difficult.

How will the Born dual amplitude behave at $s \to \infty$? If the asymptotic behaviour were arbitrary, there would be no hope that the pole approximation is any good. In this case the correction terms to (16.1) which have the form of *loops* and are responsible for the *unitarization* of the amplitude would have to be significant.

From the point of view of the t-channel pole expansion in (16.2), the reggeon behaviour $A \propto s^{\alpha(t)}$ seems to be very natural. Indeed, the Veneziano formula matches this expectation, meaning that the theory embedded a priori not just many particles but the true Regge families. To see this we take the limit of large $(-s)$ in the integral representation

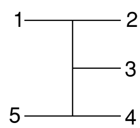
https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

(16.3) and make use that $-\alpha(s) \gg 1$:

$$\begin{aligned}\int_0^1 dx(1-x)^{-\alpha(s)-1}x^{-\alpha(t)-1} &= \int^1 dx\, x^{-\alpha(t)-1}\,\mathrm{e}^{-\alpha(s)\ln(1-x)}\\ &\approx \int_0^\infty \frac{dx}{x}\,x^{-\alpha(t)}\,\mathrm{e}^{\alpha(s)x} = \int_0^\infty \frac{dy}{y}\left(\frac{y}{-\alpha(s)}\right)^{-\alpha(t)}\mathrm{e}^{-y} \\ &\propto (-\alpha(s))^{\alpha(t)} = (-s)^{\alpha(t)}.\end{aligned} \tag{16.4}$$

Actually, we could use the well-known Stirling formula for Γ in order to derive the asymptotics of the B function. The integral representation is, however, more suitable for generalizations.

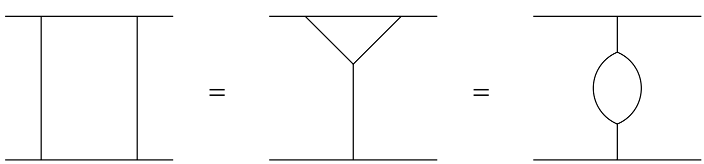
Multi-particle diagrams also exhibit multi-regge behaviour:



$$\sim\ (-s_{23})^{\alpha(t_{12})}\cdot(-s_{34})^{\alpha(t_{45})}$$

As for the problem of unitarization, almost no progress has been achieved apart from a telling technical achievement.

It turns out that in the case of an infinite number of particles, in the same way as in the usual QFT, instead of inserting the imaginary part in the dispersion integral one can draw a Feynman diagram automatically satisfying the unitarity conditions. However, here the loop diagrams have more subtle properties. Due to (16.2), there is a variety of remarkable equalities between absolutely different graphs, for example,



The sharp turn from the attempts to do without the internal structure of hadrons (local features of the objects) was, I think, due to experiments, namely, those on deep inelastic scattering.

16.2 Parton picture

The concept of Feynman–Bjorken partons was invented to explain the deep inelastic scattering phenomenon but may have a deeper significance.

Using perturbative language, we have seen that at high energies there emerged significant simplifications in a wide class of hadron interaction

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

processes. Is it possible to see these simplifications without appealing to the perturbation theory, not with the help of reggeons, but directly from the field theory?

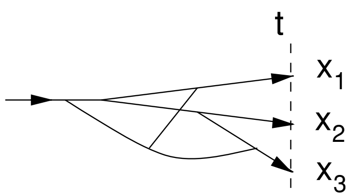
16.2.1 Parton wave function of a high-energy hadron

How does field-theoretical description relate to the usual quantum mechanics? Why do we use Feynman graphs rather than the wave function as in the non-relativistic theory? Can one return to the wave function in QFT?

In a field theory, if we make a snapshot of a particle we will see many particles whose number changes with time. The NQM wave function depends on the coordinates of all the particles in the system; even if we manage to invent a QFT wave function, it would be a multi-component object.

What is a physical particle?

Even if I start with a bare object, it 'dresses up', and in the course of propagation develops into a multi-particle system:



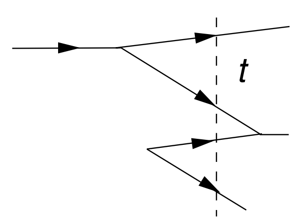
$$\Psi = \begin{pmatrix} \psi_1(t, \mathbf{x}_1) \\ \psi_2(t, \mathbf{x}_1, \mathbf{x}_2) \\ \cdots \\ \psi_n(t, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \\ \cdots \end{pmatrix}.$$

An ensemble of such Green functions determines probability amplitudes for finding 2, 3, etc. particles at a given time in definite points in space.

A wave function is a convenient object to use; it is normalized in a definite way, the integrals of the squared wave function determine probabilities of various physical processes.

So why don't we use a wave function in quantum field theory?

Recall that the QFT diagrams necessarily contain the space–time configurations in which some particles originate directly from the *vacuum* rather than from the original particle itself. Even in the simplest case of the self-energy correction diagram, the order of interaction times is arbitrary, and I would have to include into consideration all sort of virtual process going on in the vacuum.



In a quantum field theory, a free particle is not a dynamically closed system. Rather, it is like an object in a medium – in the 'external' field of vacuum fluctuations.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

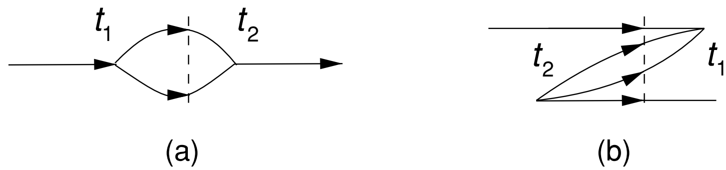
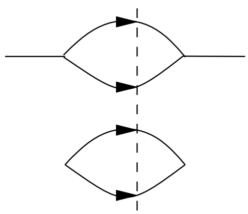


Fig. 16.1 Configurations $t_{21} > 0$ and $t_{21} < 0$ in the self-energy graph.



For certain graphs, it is straightforward to distinguish vacuum processes that do not affect the particle propagation. However, as soon as our particle and a vacuum fluctuation interact with one another (even in the future!), the picture becomes confusing; it is no longer possible to tell what belongs to the particle under consideration, and what to the vacuum.

If the longitudinal momenta of particles in the intermediate state of a virtual decay of their energetic parent, $E \gg m$, in Fig. 16.1(a) are of the same order, $x_1/x_2 = \mathcal{O}(1)$, the energy difference becomes small:

$$\Delta E = \sqrt{m^2 + \mathbf{p}^2} - \sqrt{m_1^2 + \mathbf{p}_1^2} - \sqrt{m_2^2 + \mathbf{p}_2^2} \simeq \frac{m^2}{2p} - \frac{m_{1\perp}^2}{2x_1 p} - \frac{m_{2\perp}^2}{2x_2 p}.$$

If you now integrate over t_1, t_2, an essential time interval between the interaction points will be very large – proportional to the initial energy, $t_2 - t_1 \sim 1/\Delta E \propto E$. As for the configuration of Fig. 16.1(b), the energy defect here is enormous:

$$\Delta E = 0 - \left(\sqrt{[1]} + \sqrt{[2]} + \sqrt{[3]}\right) \propto E,$$

and the lifetime of such a fluctuation is instead small: $|t_2 - t_1| \sim 1/E$. In order for our conclusion about the first graph (Fig. 16.1(a)) to hold, transverse momenta of intermediate state particles must be limited; otherwise, there would be no cancellation.

Let us sketch the calculation of contributions of these two time-ordered regions. The self-energy diagram contains two-particle Green functions:

$$\Sigma(p) = \int d^4x_{12}\, \mathrm{e}^{ipx_{21}} G(x_{21}) G(x_{21}).$$

If $x_0 = t > 0$, we close the integration contour in the lower half-plane of the energy component k_0 to obtain

$$G(x) = \int \frac{d^4k}{(2\pi)^4 i} \frac{\mathrm{e}^{-ikx}}{m^2 - k^2 - i\epsilon} = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{\mathrm{e}^{-i\left(t\sqrt{m^2+\mathbf{k}^2} - \mathbf{x}\cdot\mathbf{k}\right)}}{2\sqrt{m^2 + \mathbf{k}^2}}. \qquad (16.5)$$

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

Thus, when we integrate over *positive* times, $t_{21} > 0$ (Fig. 16.1(a)),

$$\int_0^\infty dt_{21}\, \mathrm{e}^{it_{21}\left(p_0 - \sqrt{[1]} - \sqrt{[2]}\right)} \sim \frac{1}{p_0 - \sqrt{m_1^2 + \mathbf{k}_1^2} - \sqrt{m_2^2 + \mathbf{k}_2^2}}.$$

When $t_{21} < 0$, the sign of the phase in the Green functions (16.5) flips:

$$\int_{-\infty}^0 dt_{21}\, \mathrm{e}^{it_{21}\left(p_0 + \sqrt{[1]} + \sqrt{[2]}\right)} \sim \frac{1}{p_0 + \sqrt{m_1^2 + \mathbf{k}_1^2} + \sqrt{m_2^2 + \mathbf{k}_2^2}}.$$

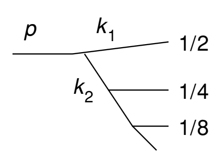
(Clearly, the integral over $d^3\mathbf{k}$ must converge for our estimate to make sense.) With the growth of energy, the real time-ordered processes of the type (a) give larger and larger contributions as compared to vacuum fluctuations (b).

You may ask: does it make sense to separate two time-ordering regions in the diagram which is relativistic invariant *as a whole*? If I choose to sit in a reference frame where the energy is finite, $E = \mathcal{O}(m)$, the vacuum processes would again be inseparable from the particle, and the simplification would disappear.

The point is, in a high-energy scattering process there is always at least one fast particle, and our consideration is valid. Already at this point it becomes apparent that the wave function Ψ that we are about to introduce will be by no means a Lorentz invariant object.

16.2.2 Feynman scaling

Let us see what will happen in more complicated diagrams. I want to stress again that our first and basic assumption is that the transverse momenta of the particles involved are bounded from above: $\mathbf{k}_\perp^2 \lesssim m^2$.

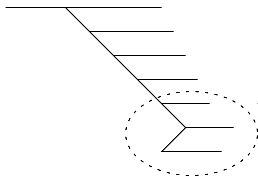


In Lectures 5 and 10 we have discussed the ‘ladder’ (multiperipheral) kinematics and saw that it is most efficient to share the large momentum roughly equally at each step of the particle multiplication: $k_{1\|} \sim k_{2\|} \sim \frac{1}{2}p_\|$. Obviously, this is so only on average; there are always fluctuations, imbalanced configurations which give, however, a smaller contribution. Thus, when we consider more and more complicated diagrams with an increasing number of particles, we get an ever-slower particle in the intermediate state. After

$$\bar{n} \sim \ln\frac{p}{m}\Big/\ln 2$$

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

steps we get a slow particle with $k_n \sim m$, and my logic of unimportance of the interaction with the vacuum breaks down.



We come to the picture of a system of point-like particles – a 'comb' of partons inside an incident fast hadron – which 'scratches' the vacuum by its soft end (the so-called 'wee' partons).

Such a picture can be characterized by a probability amplitude ψ which is *almost* a wave function. In the coordinate space, it can be normalized:

$$\psi_n(t; \mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n), \qquad \sum_n \int \prod_{i=1}^{n} d^3\mathbf{r}_i \cdot |\psi_n(t, \{\mathbf{r}_i\})|^2 = 1.$$

In the momentum space, we have

$$\psi_n(\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_n), \qquad \sum_i \mathbf{k}_i = \mathbf{p}\,.$$

The introduction of such an object does break the Lorentz invariance, but only slightly – at the level of the *slowest parton* in the ensemble. To see how this comes about, let us try to invent a *Hamiltonian* for our multi-component wave function. We take for simplicity the $\lambda\varphi^3$ interaction, represent the field as the sum of creation and annihilation operators, $\varphi = \varphi^\dagger + \varphi$, and drop the two terms corresponding to the vacuum processes:

$$(\varphi^\dagger + \varphi)^3 = (\varphi^\dagger)^3 + \varphi^3 + 3\left[(\varphi^\dagger)^2\varphi + \varphi^\dagger\varphi^2\right] \Longrightarrow (\varphi^\dagger)^2\varphi + \varphi^\dagger\varphi^2.$$

The two remaining terms describe the splitting, $1 \to 2$, and the fusion processes, $2 \to 1$. Let us look at a stationary state,

$$E\psi = \hat{\mathrm{H}}\psi,$$

and write down the dynamical equation for the n-parton wave function component:

$$\begin{aligned} E\,\psi_n = \sum_{i=1}^{n} E_i\,\psi_n \;+\; g\sum_i \psi_{n-1}(\mathbf{k}_1, \ldots, \mathbf{k}_i + \mathbf{k}_{i+1}, \ldots, \mathbf{k}_n) \\ + g\sum_i \int d\mathbf{p}'\,\psi_{n+1}(\mathbf{k}_1, \ldots, \mathbf{k}_{i-1}, \mathbf{p}', \mathbf{k}_i - \mathbf{p}', \ldots, \mathbf{k}_n). \end{aligned}$$

Here the first – kinetic – term on the r.h.s. is the sum of parton energies, the second one is responsible for the splitting of one among $(n-1)$ partons

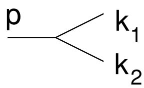
https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

into two, and the last term – for the fusion of two partons in the $(n+1)$ state into one.

Let us examine the simplest, two-parton state:

$$\big[E(\mathbf{p}) - E_1(\mathbf{k}_1) - E_2(\mathbf{k}_2)\big]\psi_2(\mathbf{k}_1, \mathbf{k}_2) = h \cdot \psi_1(\mathbf{k}_1 + \mathbf{k}_2) + h' \cdot \psi_3\,. \quad (16.6)$$

Calculating the splitting term h,



$$= g\,\frac{d^3\mathbf{k}_1}{2k_{01}(2\pi)^3}\frac{d^3\mathbf{k}_2}{2k_{02}(2\pi)^3}\frac{d^3\mathbf{p}_1}{2p_0(2\pi)^3}\cdot(2\pi)^3\delta(\mathbf{p}-\mathbf{k}_1-\mathbf{k}_2)$$

$$\Longrightarrow\ \frac{g}{2p}\cdot d\Gamma(\mathbf{k}_1)\,d\Gamma(\mathbf{k}_2),$$

and expanding the difference of energies,

$$E(\mathbf{p}) - E_1(\mathbf{k}_1) - E_2(\mathbf{k}_2) \simeq \frac{1}{2p}\left[m^2 - (m^2+\mathbf{k}_{1\perp}^2)\frac{p}{k_{1\|}} - (m^2+\mathbf{k}_{2\perp}^2)\frac{p}{k_{2\|}}\right],$$

we observe that the common factor $1/2p$ cancels, leaving us with

$$\left[m^2 - (m^2+\mathbf{k}_{1\perp}^2)\frac{p}{k_{1\|}} - (m^2+\mathbf{k}_{2\perp}^2)\frac{p}{k_{2\|}}\right]\psi_2 \;=\; g\,\psi_1 + \cdots\,.$$

We conclude that the Hamiltonian depends not on the parton longitudinal momenta themselves, but on their *ratio* to the initial momentum, $k_{i\|}/p$. In other words, the wave function is a function of the *rapidity difference*

$$\xi - \eta_i \;\simeq\; \ln\frac{p}{k_{i\|}}; \qquad \eta = \tfrac{1}{2}\ln\frac{k_0+k_\|}{k_0-k_\|}, \quad \xi = \tfrac{1}{2}\ln\frac{E+p}{E-p}.$$

This is the manifestation of the (partial) relativistic invariance: as long as the partons are *fast enough* for the expansion of the energy roots E_i in (16.6) to be applicable, the parton wave function remains invariant under the *boosts* along the direction of the large initial momentum $\mathbf{p}$.

What next? Since the incident particle is a cloud of virtual particles – the partons, we can investigate the parton wave function by studying the integrals, representing the parton number density, two-parton correlations, in other words – the density matrix characterizing the multi-parton state:

$$\begin{aligned}\sum_n \int \psi_n(\{k_i\})\psi_n^*(\{k_i\}) \prod_{j\neq\ell} d\Gamma(k_j) \;&=\; n(k_\ell),\\ \sum_n \int \psi_n(\{k_i\})\psi_n^*(\{k_i\}) \prod_{j\neq\ell,m} d\Gamma(k_j) \;&=\; n(k_\ell, k_m), \quad \text{etc.}\end{aligned} \qquad (16.7)$$

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

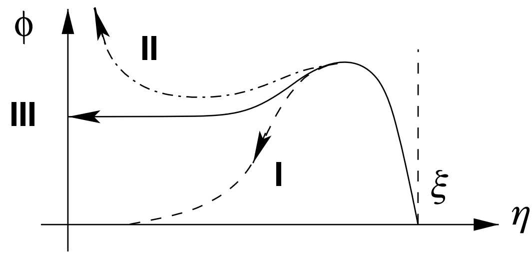


Fig. 16.2 Possible regimes of the parton-density behaviour at small momenta.

16.2.3 Parton density inside a hadron

Let us fix the momentum $\mathbf{k}$ of one of the partons and integrate over all the others, to obtain an inclusive *parton density*:

$$\phi(k_{\|}, \mathbf{k}_{\perp}) = \sum_n \frac{1}{(n-1)!} \int \prod_{i=1}^{n-1} \frac{d^3\mathbf{k}_i}{2k_{0i}(2\pi)^3} \cdot |\psi_n(\mathbf{k}, \mathbf{k}_1, \ldots, \mathbf{k}_{n-1})|^2 \,.$$

One can imagine different possibilities.

(1) Multi-parton fluctuations 'bounce' from the vacuum and collapse back into the original particle, not producing enough 'wee' partons, line I in Fig. 16.2. Such a regime is typical for weakly interacting systems when the coupling is small and the perturbation theory works.

(2) A fast growth of the wee-parton density (line II) is a signal of an instability. This shows that the interaction does not stabilize the system; the density of particles in unit volume increases indefinitely, signalling an intrinsic instability. In particular, in the $\lambda\varphi^3$ theory the situation is most likely like this (no vacuum state).

(3) The vacuum plays the rôle of the boundary, but with account of production and re-absorption of partons, a certain constant density of slow partons emerges (line III).

The key hypothesis of the parton model is that, one way or another, as a result of a balance between parton emission and recombination processes, the mean parton density does occur in nature. In essence, this is the condition for a particle to exist 'independently of the vacuum', in the sense that the fast part of the parton 'comb' does not depend on the reference frame (is invariant under $\eta \to \eta + \text{const}$), and hence, does not

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

know about the vacuum:

$$dn \;=\; \phi(\xi-\eta)\,d\eta \;\to\; \phi_\infty\,d\eta, \quad \xi-\eta \gg 1.$$

As for the transverse momentum dependence of the parton density, not much can be said about it from the first principles. At the same time, it is important to bear in mind the hypothesis which is crucial for the parton picture, namely that the transverse momentum integrals converge at some finite scale $\left\langle \mathbf{k}_\perp^2 \right\rangle \sim m^2$.

A plausible picture for the double differential distribution

$$dn(\eta,\mathbf{k}_\perp^2) = \phi(\xi-\eta,\mathbf{k}_\perp^2)\,\frac{d\eta\,d^2\mathbf{k}_\perp}{2(2\pi)^3}$$

can be drawn based on two observations. Firstly, if successive parton emissions are independent of each other, then, as we have discussed in Lectures 5 and 9, the random walk pattern emerges, and in the impact parameter space we have

$$\phi(\xi-\eta,\boldsymbol{\rho}^2) \;\propto\; \exp\left\{-\frac{\boldsymbol{\rho}^2}{\gamma(\xi-\eta)}\right\}. \tag{16.8a}$$

Secondly, a fluctuation with a typical lifetime $1/\mu$ may have a *total number* (multiplicity) of slow partons of order unity. Therefore, the distribution (16.8a) has to be properly normalized:

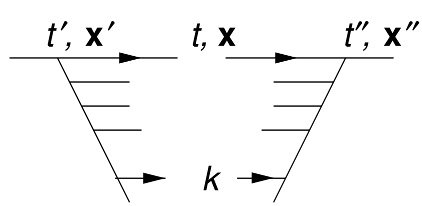
$$\phi(\Delta\eta,\boldsymbol{\rho}^2) \;=\; \frac{C}{\gamma\,\Delta\eta}\exp\left\{-\frac{\boldsymbol{\rho}^2}{\gamma\,\Delta\eta}\right\}, \qquad \int d^2\boldsymbol{\rho}\,\phi(\Delta\eta,\boldsymbol{\rho}^2) = C. \tag{16.8b}$$

No surprise, this is nothing but the vacuum pole amplitude, with the pomeron slope $\alpha' = \frac{1}{4}\gamma$.

16.2.4 Partons and reggeons

Importantly, an analysis in terms of the partons essentially coincides with the analysis of high-energy scattering amplitudes that we carried out in these lectures.

Indeed, let us take the simplest ladder-type diagram, fix the momentum k of one of the partons and integrate over all the others.

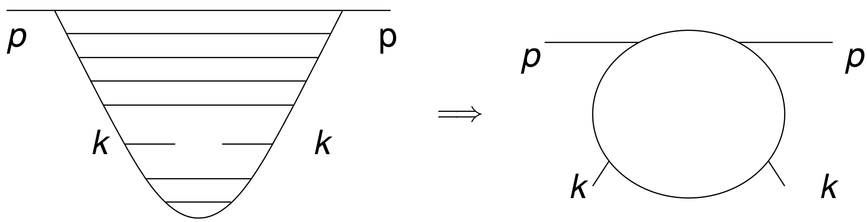


https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

An integration of the product of the amplitude and the conjugate amplitude over the position of the parton $\mathbf{x}$ yields (in the time-ordered region, $t' < t < t''$) the Green function describing the propagation from $\mathbf{x}'$ to $\mathbf{x}''$:

$$\int G(t-t', \mathbf{x}-\mathbf{x}')\, \overleftrightarrow{\partial_t} G^*(t''-t, \mathbf{x}''-\mathbf{x})\, d^3\mathbf{x} \;=\; G(t''-t, \mathbf{x}''-\mathbf{x}').$$

Hence, integrating over all particles, we arrive at the diagram



which is, essentially, the scattering amplitude. Thus, studying the parton density, we in fact investigate the scattering amplitude. If we substitute the Regge expression for the forward scattering amplitude, $s^\alpha \sim (p/k)^\alpha$, the pomeron pole, $\alpha(0) = 1$, gives us a homogeneous parton density distribution,

$$dn(k) \sim \frac{dk}{k} = \frac{dx}{x} = d\eta,$$

– the famous *Feynman plateau* in rapidity.

Our theory of the asymptotic behaviour of hadron scattering amplitudes provides an insight into the structure of the hadron wave function, and such a duality has a general physical significance.

The deuteron is a long-living object; the proton and the neutron inside the deuteron collide from time to time, but most of the time they are well separated. Does this picture apply to arbitrarily large energies?

Yes: this would mean that the deuteron always behaves as an object that consists of two independently interacting particles.

No: the parton clouds of the two nucleons eventually overlap at high-enough energies, so that by looking at the results of the collision with the target, it is impossible to tell if we had a deuteron for a projectile.

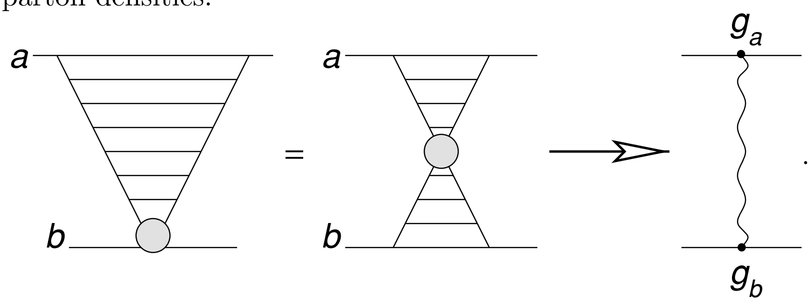
As we already know, this dilemma is related to the choice between the strong- or weak-coupling regimes of the reggeon field theory.

The knowledge of the properties of the strong interaction between slow partons is lacking. In spite of this, from the point of view of the parton

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

picture, the asymptotic equality of the cross sections of *all* hadron processes looks rather natural.

In the laboratory reference frame a slow parton from the incident hadron interacts with the target; in the centre-of-mass frame it is slow partons from the wave functions of the colliding hadrons that interact with each another, in which case the hadron–hadron cross section is given by the product of the universal slow-parton interaction cross section and two parton densities.



In either picture the density of slow partons at the tip of a long parton comb is universal – independent of the parent hadron (factorization). Moreover, according to the logic of the parton picture, a hadron is almost never in a sterile state: it is *always* represented by a 'comb'.

In the reggeon language this means that the probability of emitting a pomeron is one. Therefore, the pomeron–hadron residues, and thus the total hadron–hadron interaction cross sections,

$$\sigma_{ab} \propto g_a\, g_b \propto \left\langle \sum_{s,s'} \sigma_{ss'}(x,x') \right\rangle ,$$

turn out to be independent of the types of hadrons.‡

16.2.5 Hadrons 'inside' a parton

What is new in what the parton picture tells us about the strong interaction?

When a soft parton interacts with the target, the coherence of the system breaks down. The partons of the 'comb' get released and, emitting their own 'ladders', separate from one another producing some number of hadrons in the final state.

The most interesting question is how the transition from partons to hadrons occurs. One can formulate some sort of an *uncertainty principle.*

‡ For a detailed discussion of hadron–hadron and lepton–hadron interactions in the framework of the parton picture see Gribov's lecture entitled 'Space–time description of the hadron interactions' in Gribov (2003).

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

We say that a hadron consists, on average, of $\langle n\rangle \sim \ln p$ partons. It is clear that the opposite should also hold; the wave function of a single parton 'contains' $\langle n\rangle$ hadrons:

$$\Delta n_{\rm part} \times \Delta n_{\rm hadr} \sim \ln p\,.$$

How to visualize the conversion of a single parton into hadrons?

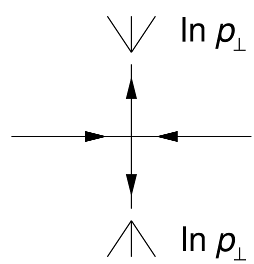
I have told you that the interaction of slow partons determines strong processes, while to explore the nature of the point-like constituents one has to employ electromagnetic and weak probes. This is not entirely true.

If we look at *rare* processes with large momentum transfers, significant simplification occur in a purely strong interaction as well. In particular, the parton → hadrons transition can be studied in a rare process of a large-angle scattering of *energetic* partons.

Such scattering produces a constituent with large transverse momentum $p_\perp$. With the increase of $p_\perp$ it becomes less and less likely that the other partons from the wave function of the incident hadron will follow suit and recombine with the struck parton. Therefore, an isolated parton will fragment on its own, producing a shower of hadrons.

Once we measure, say, a single energetic pion at 90° in the cms of the hadron collision, with a unit probability it has to be accompanied by a bunch of hadrons with logarithmic multiplicity.

Moreover, by the conservation of the transverse momentum, in the direction opposite to that of the triggered particle in the transverse plane there must be also $\ln p_\perp$ particles that originate from the recoiling parton.



Hadrons with large transverse momenta are rarely produced in hadron collisions: the inclusive distribution falls fast with $p_\perp$. However, as soon as we have triggered one such particle, there is no additional suppression for having other large-$p_\perp$ hadrons in the final state.

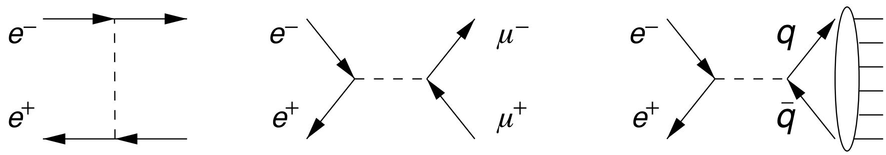
Observation of hadron jets, and especially of a recoiling jet in hadron collisions with the production of large $p_\perp$ particles verifies the parton picture.

16.2.6 e^+e^- annihilation into hadrons

The annihilation of e^+ and e^- into hadrons is the cleanest process from the point of view of the parton picture. In the first order in $\alpha_{\rm em}$ an electron and a positron may either scatter or *annihilate*, producing any charged

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

particle and its antiparticle, e.g. $\mu^+\mu^-$.



The corresponding cross section one calculates in a standard way in quantum electrodynamics. What character should have the process of e^+e^- annihilation into *hadrons*?

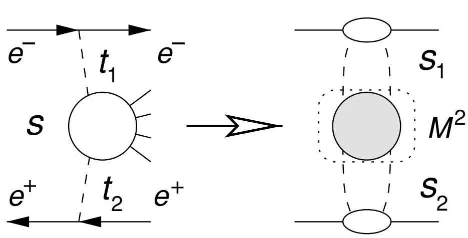
In quantum field theory the photon interacts with a point-like charge. If the energy is large, an intermediate-state photon has a huge virtual mass that can produce either a pair of energetic leptons or, equally well, a pair of electrically charged *partons* (quark and antiquark) flying in the opposite directions in the centre-of-mass of the collision. The strong interaction switches on, and two bare partons convert into hadrons. If transverse momenta in this transition process are, once again, limited, we will see two jets of hadrons emerging from the annihilation point.

This process is especially interesting in the sense that its cross section is easy to *calculate* since it has a purely electromagnetic nature. The cross section falls with the energy as $\sigma \propto \alpha_{\text{em}}^2/s$, but its *ratio* to the standard QED $\mu^+\mu^-$ production cross section tends to a constant given by a simple expression

$$R(s) \equiv \frac{\sigma_{e^+e^-\to\text{hadrons}}(s)}{\sigma_{e^+e^-\to\mu^+\mu^-}(s)} \simeq \frac{\sum_i e_i^2}{e^2}\,. \tag{16.9}$$

The sum runs over the species of partons that can be produced at a given energy, i.e. with masses $m_i < \frac{1}{2}\sqrt{s}$.

Let us remark that at *very high energies* the production of hadrons in e^+e^- collisions is dominated by another process in which hadrons originate from the 'collision' of virtual photons belonging to electromagnetic coats of incident leptons:



$$\frac{d\sigma}{dM^2} \propto \frac{\alpha_{\text{em}}^4}{M^4}\frac{dt_1}{t_1}\frac{dt_2}{t_2}. \tag{16.10}$$

The corresponding cross section is much smaller, of the order of α_{em}^4. However, it does fall with energy; since photon spin is one, the energy behaviour is similar to the case of the pomeron exchanges, $s_1 M^2 s_2 \sim s$. Moreover, the total hadron production cross section actually *grows* with

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

energy. Integrations over momentum transfers t_1 and t_2 in a broad interval $t_{\min} \sim M^4/s \ll |t_i| \ll s$ give rise to two logarithms $\ln s$; one more logarithmic enhancement originates from the integration over the rapidity of the hadron block: $M^4\, d\sigma/dM^2 \propto \ln^3 s$.

At small energies $R(s)$ of (16.9) was exhibiting resonance structures (markedly, the ρ meson peak), and then froze at $R \simeq 2$. This value came as a gift to the hypothesis of three *coloured quarks*:

$$\frac{1}{e^2}\left(e_u^2 + e_d^2 + e_s^2\right) \times 3 = \left(\left(\tfrac{2}{3}\right)^2 + \left(-\tfrac{1}{3}\right)^2 + \left(-\tfrac{1}{3}\right)^2\right) \times 3 = 2.$$

Now it seems that above $\sqrt{s} \simeq 3\,\mathrm{GeV}$ a new threshold opens up and a new heavier quark ('charm') enters the game. Independently of the theory, in the near future we will be witnessing an avalanche of new particles. A new spectroscopy starts following that of the 1960s when the model based on three light quarks managed to classify hadrons.

16.3 Deep inelastic scattering

Now that we have gained certain knowledge about strong interactions, let us discuss some aspects of the lepton–hadron scattering that we did not touch on in Lecture 4.

Recall the essence of the deep inelastic scattering (DIS) phenomenon: a virtual photon seems to interacts with a point-like particle ('parton') inside a target hadron, which parton has spin $\frac{1}{2}$ and a limited transverse momentum.

16.3.1 Photon interaction with nucleons and nuclei

To penetrate deep into the proton interior, the momentum transfer $|q^2|$ has to be large. The energy transferred by the photon from the lepton to the proton should also be large enough, $W^2 = |q^2|(\omega - 1) \gg m^2$, in order to have a multi-state hadron system produced.

Imagine hadrons were point-like. Then the direct interaction process of Fig. 16.3(a) dies out at high energies as $\sigma \propto 1/s$.

It is easy to see, however, that there are certain unusual processes due to which the photon–hadron interaction cross section, although small, is *constant* in the high-energy limit.

The photon may fluctuate into a hadron pair, as shown in Fig. 16.3(b). The lifetime of such a state may be very large at high energies, $t \sim q_0/\mu^2$. This does not happen often because the electromagnetic coupling is numerically small; we may find the photon in a 'hadron state' only in one of

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

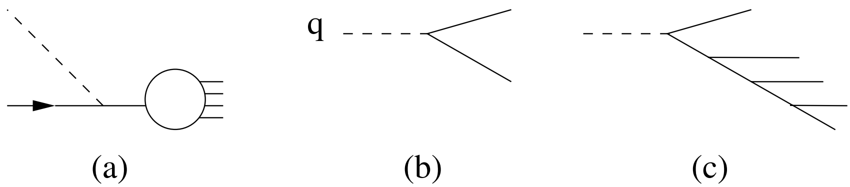


Fig. 16.3 Direct photon–proton interaction (a); energetic photon fluctuating into hadrons (b, c).

137 occasions, so to speak. However, once a fluctuation like that occurred, with respect to the *hadron multiplication* there is no perturbation theory, so that the photon will develop a full-scale parton comb, as a hadron does.

Thus, with the probability $\mathcal{O}(\alpha_{\rm em})$ the photon would approach the target as an ensemble of hadrons. Among them there are slow ones which will interact with the target providing $\sigma_{\rm tot}^{\gamma N}(s) \simeq$ const. The only difference with the hadron–hadron scattering case is in the size of the corresponding photon–pomeron coupling: $g_\gamma \propto \alpha_{\rm em}$.

There exists a rather unexpected experimental check of the validity of this picture. What would you expect for the cross section of photon interaction with a *nucleus*? A nucleus with atomic number A has a radius $R \sim A^{1/3}$. Total hadron–nucleus cross sections behave as $\sigma_{\rm tot}^{hA} \sim \pi R^2$; strongly interacting particles get stuck, with a large probability, and hence the cross section is proportional to the *surface* of the target.

On the other hand, since the electromagnetic interaction is weak, there is no large absorption and therefore the photon may freely pass through the nucleus. Therefore one would expect

$$\sigma_{\rm tot}^{\gamma A} = \sigma_{\rm tot}^{\gamma N} \cdot A \sim \sigma_{\rm tot}^{\gamma N} \cdot R^3.$$

But if our picture of a fluctuating photon is correct, at sufficiently high energy the photon must interact with the nucleus with a *hadronic* cross section $\sigma \propto R^2$!

The condition reads $q_0/\mu^2 \gg R$, that is, the lifetime of the hadron ladder fluctuation exceeding the time it takes to traverse the nucleus (Ioffe, 1969).

If we are sitting in this kinematical domain, we do not explore, with the help of a photon probe, the interior of the hadron target but simply observe a hadron–hadron interaction. We would see nothing like the Rutherford scattering in this case.

If our goal is instead to study the internal structure of the nucleon, we must cut off hadron-like configurations inside the photon. How to do that? By restricting the lifetime of the photon in such a way that it would

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

have no time to develop a shower while passing through the nucleon, that is, by making the photon sufficiently virtual:

$$\frac{q_0}{|q^2|} \lesssim \frac{1}{\mu} \quad \Longrightarrow \quad \frac{2mq_0}{|q^2|} = \omega \lesssim \frac{2m}{\mu} \sim 10.$$

So, the dimensionless parameter $\omega = -2pq/q^2$ that we have introduced in Lecture 4 must be of the order of unity (but not too large) in order to see the photon interacting with point-like partons inside the nucleon, to see the Bjorken scaling, etc.

16.3.2 DIS in the parton picture and quarks as partons

We are going to apply the parton model to the DIS process. Since we have a picture of the parton content of a fast hadron, let us choose a reference frame $q_0 = 0$ in which a fast hadron collides with a static electromagnetic field along some direction $\mathbf{z}$:

$$q^2 = -q_z^2, \;\; 2pq = -2p_z q_z, \quad \omega = -\frac{2pq}{q^2} = -\frac{2p_z}{q_z}.$$

The photon will be absorbed by one of the partons with momentum k. By which? It has to be a fast parton since otherwise an overlap of its wave function with the photon field of the size $\Delta z \sim 1/\left|q_z\right|$ would be small. The time the hadron passes through the field does not exceed $1/\mu$. For a long-living fast parton with the 'lifetime' $k/\mu^2 \gg 1/\mu$ this is an instantaneous interaction, and such interaction occurs with *conservation of the energy*. This condition selects a parton with a definite momentum:

$$k' = k + q, \;\; k_0' = k_0 \;\Longrightarrow\; |k_z| \simeq \left|k_z'\right| \;\Longrightarrow\; k_z = -k_z' = -\tfrac{1}{2}q_z. \qquad (16.11)$$

After absorbing the photon, the struck parton flies in the opposite direction, $k_z' = -k_z$. What will happen with it afterwards is unimportant for the DIS cross section. It is determined simply by the product of the Born cross section for the photon absorption by a parton and the density of partons with the given rapidity η:

$$\frac{d\sigma}{dq^2\,d\omega} = \frac{d\sigma_B}{dq^2} \cdot \int \phi(\eta, \mathbf{k}_\perp^2) \frac{d^2\mathbf{k}_\perp}{(2\pi)^2}, \qquad (16.12a)$$

$$\xi - \eta = \ln\frac{p_z}{k_z} = \ln\frac{2p_z}{|q_z|} = \ln\omega. \qquad (16.12b)$$

Recalling that the parton density in the parton model depends only on the rapidity distance to the parent proton (16.12b), the DIS cross section

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

takes the form

$$\frac{d\sigma}{dq^2\,d\omega} = \frac{4\pi\alpha_{\rm em}^2}{q^4}\left\{\phi_0(\omega)\cdot e_{\rm part}^2\left[1-\frac{pq}{pp_e}\right]\right. \\ \left. + \phi_{\frac{1}{2}}(\omega)\cdot e_{\rm part}^2\left[1-\frac{pq}{pp_e}+\frac{1}{2}\left(\frac{pq}{pp_e}\right)^2\right]\right\}, \tag{16.13}$$

where we have introduced the densities of partons with spin-$\frac{1}{2}$ and spin-zero, accompanied by the corresponding electron scattering cross sections.

Apart from simple QED factors, this expression contains two unknown functions ϕ of one variable ω. This is the *Bjorken scaling*. Experiments, as we have already discussed in Lecture 4, show no presence of spin-zero charges, $\phi_0 \simeq 0$.

Thus, from the parton model we have derived the Bjorken scaling using two hypotheses: $\phi = \phi(\xi - \eta)$, that is the existence of the stationary parton density, and, certainly, the hypothesis of limited transverse momenta of partons inside the hadron wave function.

The expression (16.13) contains, obviously, the parton charge $e_{\rm part}$ (in units of the electron charge). If partons are *quarks*, these charges are fractional numbers. Is there a way to measure them experimentally?

The parton density ϕ is not entirely arbitrary; it obeys certain normalization conditions. If we integrate $\phi(\eta)$ over rapidity we obtain the multiplicity of partons of a given species i:

$$\int_0^{\xi} d\eta\,\phi_i(\eta) \;=\; N_i.$$

More informative is another sum rule, namely

$$\sum_i \int_0^{\xi} d\eta\,\phi_i(\eta)\cdot \mathrm{e}^{\eta} \;=\; \frac{p}{m}, \tag{16.14}$$

which expresses the fact that the total longitudinal momentum (energy) of all the partons equals the momentum of the hadron they belong to.

The DIS cross section (16.13) contains the product $e_i^2\phi_i$, and one cannot extract the charges without knowing the normalization of ϕ. Fortunately, there is an additional source of information.

One may study *weak* processes like deep inelastic neutrino scattering, $\nu p \to \mu^- + X$. The weak interaction is insensitive to electric charges but feels the presence of spin-$\frac{1}{2}$ partons in the proton. The mass of the intermediate boson $W^{\pm}$ is apparently too large to make itself felt at presently available energies, so that the behaviour of partons in weak DIS processes can be described by means of the standard four-fermion Fermi interaction.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

By combining the information coming from electromagnetic and weak DIS, one attempts to extract electric charges of partons. Modulo some uncertainties in the neutrino–quark scattering amplitudes, the results are consistent with

$$\sum_{i \in \text{proton}} e_i^2 \simeq 1 \qquad \left(= 2 \cdot \left[\tfrac{2}{3}\right]^2 + \left[\tfrac{1}{3}\right]^2 \right),$$

$$\sum_{i \in \text{neutron}} e_i^2 \simeq \frac{2}{3} \qquad \left(= 2 \cdot \left[\tfrac{1}{3}\right]^2 + \left[\tfrac{2}{3}\right]^2 \right).$$

Moreover, integrating and adding up the densities of quarks–partons of different species, one can estimate the total *energy-momentum* carried by charged partons inside the nucleon. The result is telling: about 50% of the nucleon momentum in (16.14) belongs to some electrically neutral fields (gluons?).

16.3.3 Structure of the final state

Let us address a very important question: what happens with the parton system after one parton is kicked off.

We turned around one of the partons of what was a coherent system. The relative invariant energy between the struck parton and its neighbours becomes large, and therefore it is unlikely to interact with the rest of the system. This means that we have prepared an isolated parton – a bare field-theoretical object. Flying in the opposite direction, it will 'decay' into $\ln(q_z/2)$ hadrons.

What about the rest of the parton comb that keeps moving in the initial direction of the nucleon?

Coherence of the *bottom part* of the parton fluctuation in Fig. 16.4(a) is not disturbed by the scattering. First slow partons with $k_i \sim \mu$ and then, successively, faster partons will be absorbed; they will revert to assembling the initial coherent proton. However, at the level of $k_i \sim k_z = \frac{1}{2}|q_z|$ the parton ensemble becomes aware that its coherence had been broken; the *upper part* of the fluctuation will get released, turning into hadrons.

The resulting rapidity distribution of final-state hadrons is sketched in Fig. 16.4(b). It has a characteristic *hole* in rapidity.

This would be the structure of the final state if quarks were *not confined.* The answer may be different if there is some specific dynamics (beyond our naive field-theoretical approach based on the locality of the interaction in rapidity) that would force the struck quark to interact with the rest of the parton ensemble in order to prevent the two separating hadron

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

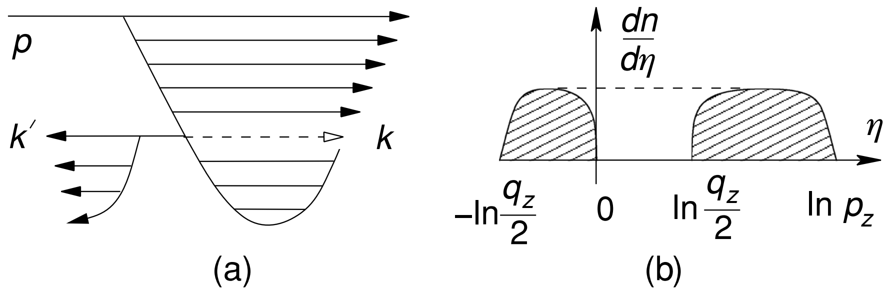


Fig. 16.4 (a) Coherent collapse of an undisturbed part of the parton wave function in DIS; (b) rapidity distribution of final-hadrons with a 'hole'.

systems from having *fractional electric charges*. Such an interaction could fill in the gap between the negative rapidity ('current fragmentation') and positive rapidity hadrons ('target fragmentation'), leading to a uniform hadron plateau – the dashed line in Fig. 16.4(b).

The existence of the hole in the hadron distribution is a key test for quarks in the rôle of partons. Experimentally, there is no sign of a hole in the proton fragmentation.

Thus we face a strange situation: on the one hand, the picture of quasi-free quark-partons works in DIS; on the other hand, the mechanism that forces the flying-away quark to 'communicate' with the rest of the proton is unclear.

16.4 The problem of quarks

There were times when quarks were thought to have large proper masses in order to explain their non-observation as free particles. Now, from the DIS experiments we know that the masses of $(u,\, d)$ quarks must be smaller than 1 GeV. Thus, there has to be a special reason for quarks not appearing in the physical spectrum – the *confinement*.

Whatever the reason for quarks to be confined inside hadrons, looking at what happens in the e^+e^- annihilation it is clear that to ensure the quark confinement in such a process is not easy.

The quark and the antiquark fly too fast, and separate too far, to be able to interact with one another. Nevertheless, we expect two 'jets' of hadrons in the final state. Therefore, the production of multiple $q\bar{q}$ pairs which then recombine to form mesons, Fig. 16.5, cannot be explained by the re-interaction between the quarks created in the annihilation point. It has to be the result of the reaction of the *vacuum* on the production of two relativistic charges.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

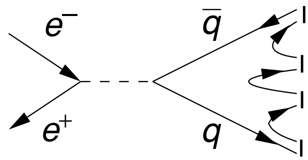


Fig. 16.5 Recombination of quarks into final-state hadrons in e^+e^- annihilation.

In quantum electrodynamics such a process is accompanied by radiation of soft photons (bremsstrahlung). Maybe in strong interactions an analogous phenomenon also plays a rôle.‡

16.4.1 One-dimensional electrodynamics

We have an example of a field theory with confinement – the Schwinger model. It is one-dimensional electrodynamics with massless fermions. It does not answer the question why quarks are confined in the real world, but it demonstrates nevertheless how the reaction of the vacuum results in the formation of 'hadrons' (Schwinger, 1962).

In one spatial dimension two charges cannot be separated because they interact as two infinite planes in our world, with their electromagnetic energy increasing linearly with the distance:

$$\mathcal{E} = \frac{\mathbf{E}^2 V}{8\pi}, \quad |\mathbf{E}| = \text{const.}$$

Classically, a pair of planes with opposite charges will oscillate. In the quantum case the system cannot have an arbitrary energy; it has to be quantized. So, a boson spectrum with a definite mass must appear. If we produce two planes with a large energy (as in e^+e^- annihilation), we expect that many oppositely charged pairs of planes will be produced giving rise to many bosons in the final state.

Let us study this theory in a more formal way. We have massless fermions and a photon, and the standard electromagnetic interaction between them. Since our space is a line, a massless fermion moves with the speed of light either in the positive or in the negative direction along it.

As we shall see shortly, an amplitude for the photon transfer into the fermion pair is different from zero only when the fermions move in the

‡ Indeed, it is radiation of soft gluons that fills in the 'Gribov hole' (Gribov *et al.*, 1987). QCD bremsstrahlung plays a key role in the formation of final hadron states (Amati and Veneziano, 1979; Marchesini and Webber, 1984).

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

same direction (this is the consequence of the helicity conservation in the electromagnetic interaction).

An invariant mass of a pair of massless particles moving in the same direction is zero,

$$(k_1+k_2)^2 = k_1^2 + k_2^2 + 2(k_{10}k_{20} - k_{1z}k_{2z}) = 0+0+0, \quad k_{1z}/k_{10} = k_{2z}/k_{20}.$$

Therefore, we have two massless states – the photon and the $q\bar{q}$ pair – that mix, so there must be a splitting of the degenerate levels. (In fact, $q\bar{q}$ is not one state but many, with different quark energies, k_{10}/k_{20}.)

As a result, the discrete state – the photon – must acquire a finite mass. Let us see how this occurs.

$$\Pi_{\mu\nu}(k) = -e^2 \int \frac{d^2p}{(2\pi)^2 i} \,\mathrm{Tr}\left[\gamma_\mu \frac{1}{\hat{p}}\gamma_\nu \frac{1}{\hat{p}-\hat{k}}\right] = (g_{\mu\nu}k^2 - k_\mu k_\nu)\Pi(k^2).$$

Let us calculate the imaginary part of the specific component Π_{00} of the polarization tensor:

$$\mathrm{Im}\,\Pi_{00} = -e^2 \int \frac{d^2p}{2} \,\mathrm{Tr}\left[\gamma_0\hat{p}\gamma_0(\hat{p}-\hat{k})\right]\delta_+(p^2)\delta_+((k-p)^2). \tag{16.15a}$$

The trace equals

$$\begin{aligned} -\frac{1}{2}\,\mathrm{Tr}\left[\gamma_0\hat{p}\gamma_0(\hat{p}-\hat{k})\right] &= 2p_0(k_0-p_0) - p(k-p)] \\ &= p_0(k_0-p_0) + p_z(k_z-p_z). \end{aligned} \tag{16.15b}$$

Since $p_0 = |p_z|$ and $k_0 - p_0 = |k_z - p_z|$, the trace vanishes when the fermions fly in opposite directions. When the momenta are parallel, (16.15) yields

$$\mathrm{Im}\,\Pi_{00}(k^2) = e^2 \int_0^{k_0} \frac{dp_0}{2p_0} 2p_0(k_0-p_0)\delta(k^2) = e^2k_0^2\delta(k^2) \;\equiv\; k_z^2 \cdot \mathrm{Im}\,\Pi(k^2).$$

We get

$$\mathrm{Im}\,\Pi(k^2) = e^2\delta(k^2) \quad\Longrightarrow\quad \Pi(k^2) = \frac{e^2/\pi}{k^2 - i\epsilon}.$$

The singularity at $k^2 = 0$ of the polarization operator results in

$$D_{\mu\nu} = \left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) d_t(k^2), \quad d_t(k^2) = \frac{1}{k^2(1-\Pi(k^2))} = \frac{1}{k^2 - m^2}.$$

Thus, the gauge invariance is spontaneously broken and the photon acquires a finite mass $m^2 = e^2/\pi$. (In the one-dimensional QED the charge has the dimension of mass.) Massive photon is that very bosonic state that represents a pair of planes.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

Given a finite photon mass, electric fields fall exponentially with distance, $E \sim \exp\{-mr\}$. This would contradict the Maxwell equation $\operatorname{div}\mathbf{E} = 4\pi\rho$, unless the charge density ρ were in fact zero.

Indeed, by explicitly solving the theory one can show that the introduction of external currents provokes an appearance of the *polarization current*, which results in a local compensation of electric charges. An 'e^+e^- annihilation' process in one-dimensional electrodynamics causes the production of multiple fermion pairs, resulting in the final-state structure guessed at in Fig. 16.5.

16.4.2 A field theory for strong interactions?

The quark model has demonstrated that there is a simplicity in the spectrum of hadrons – mesons and baryons. Maybe there is a certain simplicity not only in the mass spectrum but in a deeper sense. The apparent complexity of hadron interactions does not exclude the simplicity at short distances.

In spite of the fact that the quantum mechanics of electrons in the Coulomb field is perfectly known, we would not dare to attempt to quantitatively describe the structure of a final state in, say, the collision of two atoms of mercury. We just get a mess. But at high energies, and in specific observables, the simplicity of the internal structure is manifest.

Should we approach the hadron dynamics from a short-distance side?

Although, as was mentioned in Lecture 4, the experimentally observed Bjorken scaling is not described by any field theory, we cannot imagine anything but a field theory based on simple point-like objects. Today progress (if it may be called so) goes in the direction of returning to quantum field theory with, possibly, a *small coupling constant*.

We saw in Lecture 2 that the effective pion–nucleon interaction coupling constant is large, $g_{\pi N}^2/4\pi \simeq 14$. But it may be that the coupling is large for *composite* objects – hadrons – while the quarks (which sit inside and by some reason cannot be freed) interact with a smaller constant. Otherwise, why would a parton absorb a virtual photon in the DIS process like a quasi-free particle?

Could it be possible to find a quantum field theory in which the interaction between the objects with small wavelengths is weak and the interaction with large wavelengths is strong? If such a theory will be found, we will see a leap of interest; the complexity of hadrons will cease to be the object of attention (as does the complexity of the Hg atom). Under focus will be, rather, processes in which a relatively simple internal structure of hadrons reveals itself.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

16.5 Zero charge in QED and elsewhere

The best developed quantum field theory is quantum electrodynamics. What do we know about QED? On the one hand, it contains divergences, on the other hand, it is *renormalizable*. What does this, essentially, mean? Let us imagine that either the perturbation theory, or the theory itself, is wrong at short distances. Hence, integrating over momenta, we can do this correctly only up to a certain large momentum scale $k < \Lambda$.

What was the essence of progress in the 1950s? The uncertainty which appears due to the existence of such a cutoff, is contained by two quantities: by the observable charge and the mass of the electron which, therefore, remain uncalculable. It is obvious, however, that one cannot simply stop at that. Increasing the momenta of the external particles, sooner or later Λ will make itself felt.

There are two ways to formulate the problem.

(1) Let the interaction have a simple (point-like) form at $k \sim \Lambda$ (and at $k > \Lambda$ no interaction whatsoever). At momentum scales $k \sim \Lambda$ there are no corrections at all, and with the decrease of momenta, the corrections become relevant.

(2) The second approach is, though less transparent, essentially equivalent. Assume a given interaction at $k \sim m$; I want to see what happens when k increases. This is a semi-phenomenological approach; obviously, in a quantum field theory large distances are determined by small ones and not the other way round.

The basic quantities of the theory that contain ultraviolet divergences are the vertex part and the photon and electron Green functions. The integrals that determine all the other quantities are converging.[§]

16.5.1 First approach

In the first approach one introduces the bare charge e_0 at the ultraviolet momentum scale Λ.

Each subsequent correction to the diagram adds the factor e_0^2, one photon and two fermion Green functions and the momentum integration d^4k. However, it is not reasonable just to list all the corrections order by order in perturbation theory.

As we have seen in the QED course (Gribov and Nyiri, 2001), the set of Feynman diagrams can be rearranged into the series of the *skeleton*

[§] For details of the renormalization programme in quantum electrodynamics see Gribov lectures (Gribov and Nyiri, 2001) (ed.).

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

graphs that are built of the exact vertices (Γ) and exact photon (D) and fermion Green functions (G). The topology of the skeleton diagrams is such that they do not contain sub-diagrams that could be attributed to the vertex or propagator corrections.

Skeleton diagrams have a remarkable feature: each of them contains a single power of the logarithm of the ultraviolet cutoff, $\ln\Lambda$. The reason for this is simple: the integration momenta are mixed up, each momentum enters a large number of lines, and hence, the multiple integrals converge up to the last step. This is just the property of renormalizability of quantum electrodynamics.

For large momenta k we write $G \simeq -g(k^2)/\hat{k}$ and $D \simeq d(k^2)/k^2$, and the magnitude of the correction is determined in fact by the combination

$$e_0^2\,\Gamma^2 G^2 D\,d^4k \;\to\; e_0^2\,\Gamma^2(k^2)\,g^2(k^2)d(k^2)\,d\ln k^2.$$

The quantity

$$e^2(k^2) = e_0^2\Gamma^2(k^2)g^2(k^2)\,d(k^2) \tag{16.16}$$

is the *invariant charge* which characterizes the interaction strength at the momentum scale k. This structure is common for all renormalizable theories.

How can the perturbative series be summed up? Initially we assumed the bare charge to be small, $e_0^2 \ll 1$. Let us decrease the external momenta p not too much, so that $e_0^2\ln\Lambda^2/p^2 \lesssim 1$. In this situation all powers of the parameter $e_0^2\ln\Lambda^2/p^2$ have to be resummed. My exact propagators and the vertex function have the following structure:

$$\Gamma(p^2) = \gamma_1\left(e_0^2\ln\frac{\Lambda^2}{p^2}\right) + e_0^2\gamma_2\left(e_0^2\ln\frac{\Lambda^2}{p^2}\right) + e_0^4\gamma_3\left(e_0^2\ln\frac{\Lambda^2}{p^2}\right) + \cdots\,. \tag{16.17}$$

Neglecting the corrections of the type $e_0^4\ln(\Lambda^2/p^2) \sim e_0^2 \ll 1$ we get the so-called leading logarithmic approximation (LLA) in which the problem was first solved by Landau *et al.* (1956).

The higher functions γ_n, with $n \geq 2$ are rather complicated; they are given by skeleton diagrams with the exact LLA vertices ($\Gamma = \gamma_1$) being propagators. Imagine that $\gamma_i(x)$ in (16.17) are of the same order not only for $x \lesssim 1$ but for any x. In this case everything seems to be ideal. If, however, γ_1 turns out to be singular or zero (e.g. at large x values), the higher-order terms have to be investigated seriously.

Actually, due to the Ward identity,

$$\frac{\partial G^{-1}(p)}{\partial p_\mu} \;=\; \Gamma_\mu(p,p,0),$$

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

the ultraviolet logarithms in Γ and g cancel, $\Gamma = g^{-1}$, and one is left with the photon renormalization function only in (16.16):

$$e^2(k^2) = e_0^2\, d(k^2).$$

Corrections to the propagator and to the vertex depend on the properties of the charged particle. Therefore, if not for the Ward identity, the charge would not be such a universal quantity.

We have calculated the invariant charge in Chew and Low (1959):

$$e^2(k^2) = \frac{e_0^2}{1 + \frac{e_0^2}{12\pi^2} \ln \frac{\Lambda^2}{k^2}} \,. \tag{16.18}$$

What does this expression mean? We have a simple picture: the invariant charge decreases monotonously with the increase of the wavelength (the situation is just the opposite of what we would like to have in strong interactions).

The expression (16.18) is valid for large virtual photon momenta. At small $|k^2| \lesssim m^2$, under the logarithm the mass m appears, and for the physical charge that determines, e.g. the Coulomb scattering we get

$$e_c^2 = e^2(k^2 = 0) \;=\; \frac{e_0^2}{1 + \frac{e_0^2}{12\pi^2} \ln \frac{\Lambda^2}{m^2}} \,. \tag{16.19}$$

Suppose that I have introduced in the theory a finite charge at small distances corresponding to the large momentum scale Λ. Then at larger distances the effective charge becomes smaller. Moreover, for a point-like particle there is a *complete screening*; if we decrease the size of the region over which the bare charge e_0 is smeared, the on-mass-shell charge vanishes: $e_c^2 \to 0$ with $r_0 \sim 1/\Lambda \to 0$. The polarization of the vacuum screens the point-like charge totally, so that the observable electric charge has to be zero; $e_c^2/4\pi = 0$ instead of $1/137$.

How can this be explained?

We used the hypothesis that the bare charge e_0^2 is small. The LLA cannot be blamed for such an unphysical result. Our approximation of neglecting the higher-order corrections in (16.17) becomes even *better*, since the true measure of the interaction strength – the effective charge $e^2(k^2)$ – is *decreasing* with the decrease of the virtuality.

There remains one unsolved problem: what, if e_0^2 was large initially? More than 20 years has passed since the discovery of the running QED charge and of the zero-charge problem, but nobody has been able to explain why the large charge would be harder to screen than the small one.

May be there is a real ultraviolet cutoff Λ in nature? From the very beginning, it was assumed that Λ could be related to the *gravitational radius*. In such a scenario one would need to have $\nu = 13$ elementary

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

fermions (presuming that only fermions are polarizing the QED vacuum).

In any case, pure quantum electrodynamics is a contradictory theory.

16.5.2 Second approach

The property of renormalizability means that all the divergences (dependence on the cutoff Λ) have to enter the observable charge e_c:

$$e_c^2 \equiv e_0^2 \cdot Z_3(e_0^2;\Lambda^2), \quad d(e_0^2;k^2,\Lambda^2) = Z_3(e_0^2;\Lambda^2)\cdot d_c(e_c^2;k^2).$$

Does our solution satisfy the renormalizability property? It certainly does:

$$\begin{aligned} d(k^2) &= \frac{1}{1+\frac{e_0^2}{12\pi^2}\ln\frac{\Lambda^2}{k^2}} = \frac{1}{\left[1+\frac{e_0^2}{12\pi^2}\ln\frac{\Lambda^2}{m^2}\right] - \frac{e_0^2}{12\pi^2}\ln\frac{k^2}{m^2}} \\ &= \frac{1}{1+\frac{e_0^2}{12\pi^2}\ln\frac{\Lambda^2}{m^2}} \times \frac{1}{1-\frac{e_c^2}{12\pi^2}\ln\frac{k^2}{m^2}} \equiv Z_3 \times d_c(k^2), \end{aligned} \tag{16.20}$$

where e_c^2 is given by (16.19). The photon renormalization function,

$$Z_3 = \frac{1}{1+\frac{e_0^2}{12\pi^2}\ln\frac{\Lambda^2}{m^2}}\ ,$$

embeds the whole dependence on Λ and on the bare charge e_0, while the renormalized photon Green function $d_c(k^2)$ contains only the renormalized charge $e_c^2 = Z_3 e_0^2$, as it has to be, owing to renormalizability.

Of course, this result for the Green function could have been obtained in the renormalized perturbation theory, not introducing Λ at all. We would notice that the fermion loop behaves at $k^2 \gg m^2$ as

$$(g^{\mu\nu}k^2 - k^\mu k^\nu)\cdot e_c^2 \ln\frac{k^2}{m^2},$$

i.e. perturbation theory breaks down when $e_c^2\ln(k^2/m^2) \sim 1$. Then, repeating the logics of the logarithmic approximation,

$$e_c^2 \ll 1; \qquad e_c^2 \ln\frac{k^2}{m^2} \sim 1,$$

we would arrive at a geometrical progression, and the running coupling $e^2(k^2)$ would be expressed directly in terms of renormalized quantities:

$$e^2(k^2) = e_c^2\,\Gamma_c^2 g_c^2 d_c = e_c^2\, d_c = \frac{e_c^2}{1-\frac{e_c^2}{12\pi^2}\ln\frac{k^2}{m^2}}, \tag{16.21}$$

where we have used the Ward identity.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

In such a chain of thoughts we keep e_c^2 to be fixed. If so, we get a stupid result: $e^2(k^2)$ develops a singularity ('Landau pole') with the growth of k^2. The reason for that is that we have made a contradictory assumption, namely, that the observable charge e_c^2 can be taken as a finite quantity.

The effective charge $e^2(k^2)$ is the true expansion parameter. Near the pole, where $e_c^2 \ln\frac{k^2}{m^2} \simeq 1$, it becomes large, and the higher corrections become important and may help to solve the problem. The question is not cleared yet.

There is a simple relation between the two methods: $e_0^2 = e^2(\Lambda^2)$.

This situation has also influenced other theories. For a long time it seemed that the screening, and therefore the zero-charge problem, was unavoidable in any quantum field theory, and this was the reason owing to which the interest in field theories was lost.

Let us suppose, for a minute, that the sign in the denominator of (16.21) would be the opposite:

$$e^2(k^2) = \frac{e_c^2}{1 - \frac{e_c^2}{12\pi^2}\ln\frac{k^2}{m^2}} \quad \Longrightarrow \quad \frac{g_c^2}{1 + c\, g_c^2 \ln\frac{k^2}{m^2}}\,, \quad c > 0. \tag{16.22}$$

How crazy is such an expression? It looks impossible at first sight. Indeed, how could it be that, when we move away from the charge, we see it not *screened* by the medium but, on the contrary, *growing* as if it pulled on other charges of the same sign?

Indeed, in electrodynamics this does not happen. However, for the *gravitational* interaction where all masses attract each other, this scenario can be easily imagined.

Such a system would be, obviously, unstable. Something must happen at *large* distances; the expression (16.22) contains an unphysical *ghost* pole at small momenta k^2. On the other hand, an instability may be a welcome feature which might help us to understand why the quarks cannot be separated at large distances but are confined inside hadrons.

16.6 Looking for a better QFT

According to deep inelastic scattering experiments, for strong interactions we need a theory with the interaction that *weakens* at small distances (e.g. as (16.22) does). We also want the theory to be renormalizable. This requirement does not follow from any physics. Nevertheless, to have a non-renormalizable theory is for us, unfortunately, the same as not having a theory at all.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

In order to see whether a given theory is renormalizable, it suffices to look at the *dimension* of the coupling constant: $[g] = [m^\alpha]$. At large virtual momenta p, a higher-order correction then has the structure $(g^2/p^2)^\alpha$. If $\alpha > 0$, the correction falls in the ultraviolet region, and the theory is superconvergent (as is the $\lambda\varphi^3$ theory). If, on the contrary, $\alpha < 0$, the loop integral diverges in the ultraviolet, and the theory is non-renormalizable (as the four-fermion Fermi interaction). When g is dimensionless, corrections are logarithmic, $g^2 \ln p^2$, and the theory is renormalizable.

16.6.1 Fermion and scalar fields

There was a time when the field theory with the interaction $g\bar{\psi}\gamma_5\psi\varphi$ was considered as being related to reality, with nucleons and pions as elementary fields. We have seen already that the coupling constant g is then very large, $g^2/4\pi \simeq 14$. This is, however, on the mass shell, $g(m_\pi, m_N)$. What is the form of the effective interaction? Unfortunately, the same as in electrodynamics:

$$g^2(k^2) = \frac{g_0^2}{1 + c\,g_0^2 \ln\frac{\Lambda^2}{k^2}}, \quad c > 0.$$

At small distances the interaction is even stronger. We face the same zero-charge problem as in electrodynamics. In reality one has to add to this interaction the quartic point-like interaction between pions, $\lambda\varphi^4$. Even if we do not introduce it from the beginning, it is induced by a logarithmically divergent diagram with a four-scalar field attached to the fermion loop. Thus, the new vertex is necessary to achieve renormalizability of all amplitudes, including that of the four-scalar particle interaction.

16.6.2 Vector fields: how to construct a renormalizable theory

Let us consider now a vector particle (as in electrodynamics). Will such a theory be renormalizable? Generally speaking, vector theories are *not renormalizable.*

The Green function for a vector particle with mass m is

$$D_{\mu\nu}(k) = \frac{1}{m^2 - k^2}\left[\frac{k_\mu k_\nu}{m^2} - g_{\mu\nu}\right]. \tag{16.23a}$$

Where this structure comes from? The propagator of a vector particle,

$$-\sum_{\lambda=1}^{3} \frac{e_\mu^\lambda(k) e_\nu^\lambda(k)}{m^2 - k^2}, \tag{16.23b}$$

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

contains the sum over three polarizations (the overall minus sign guarantees that the residue is positive, since $(e_\mu)^2 < 0$). Three polarization vectors out of four, $e^\lambda_\mu k^\mu = 0$, $\lambda = 1, 2, 3$, correspond to a spin one particle; in the rest frame there are three spin projections, $e^\lambda_\mu = (0, \mathbf{e}^\lambda)$. The fourth vector, $e^0_\mu = (1, \mathbf{0})$, describes a particle with spin *zero* – a scalar.

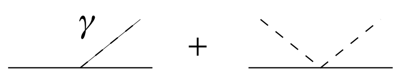
In order to eliminate the term proportional to $k_\mu k_\nu / m^2$ in (16.23a), we could try to extend the sum in (16.23b) to all four polarizations, by allowing the propagator to describe degenerate states of spin-zero and spin-one objects. In this case, however, the scalar would be a ghost (would enter with negative probability).

Thus, if we want to preserve unitarity, we must keep the $k_\mu k_\nu$ term in (16.23a), and this leads to a catastrophe: superfluous momenta multiply the vertices, and integrals diverge at large loop momenta.

Why then is electrodynamics renormalizable? There is no necessity to take special care about the 'scalar' polarization, since, owing to *current conservation*, it is not produced, and the $k_\mu k_\nu$ term can be simply dropped.

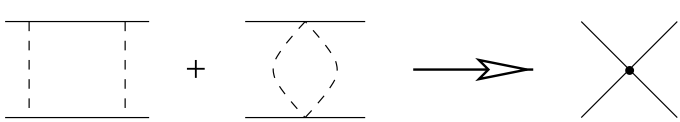
Is the fact that the photon is massless important for the renormalizability of QED? Not at all. A theory containing a massive photon, and conserved current, barely differs from electrodynamics. Unfortunately, it is also no different from QED in what concerns us, namely, the problem of zero charge; the mass in the photon propagator is unimportant in the ultraviolet momentum region, $k \to \infty$.

A few years ago it seemed that there are no other renormalizable quantum field theories. This situation looks strange: we have a theory of fermions interacting with the electromagnetic field. What about the other particles? Can they interact with the electromagnetic field? We know that the answer is yes for scalar particles, with only one subtlety: there have to be two vertices,



(16.24)

However, the scattering of scalar particles contains logarithmic divergences,



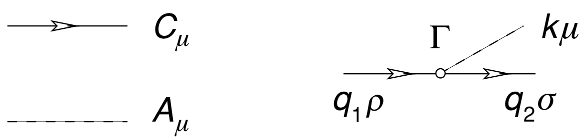
and the local four-scalar interaction has to be introduced. Thus, scalar particles cannot interact with the electromagnetic field *only*.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

A scalar particle cannot be considered as truly point-like with respect to the electromagnetic radiation; we are forced to introduce an additional interaction ($\lambda\phi^4$) which makes the particle somewhat 'smeared'.

One wonders, whether only fermions can be point-like? This would be in agreement with the fact that particles that are 'sterile' with respect to strong interactions (point-like with respect to the weak interaction) are only fermions (e, μ, ν).

What if we take a charged vector particle ?



(16.25)

For renormalizability we have to require

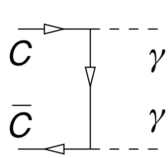
$$q_1^\rho \Gamma_{\mu\rho\sigma}(q_1, q_2) = q_2^\sigma \Gamma_{\mu\rho\sigma}(q_1, q_2) = 0;$$

in addition, conservation of the electromagnetic current requires $k^\mu \Gamma_{\mu\rho\sigma} = 0$. It can be easily seen that no vertex can satisfy the current conservation relations simultaneously with respect to all three momenta. Does this mean that charged vector particles do not exist at all?

We can, of course, imagine a vector particle as a bound state of a fermion and an antifermion. In this case the electromagnetic interaction will not reduce to that with a point-like particle with spin $\sigma = 1$, and the renormalizability will not be broken. However, we do not know how to write down such an interaction phenomenologically; moreover, we are not able to calculate bound states of relativistic objects.

In terms of dispersion relations, how would the problem manifest itself?

We have the term $q_\mu q_\nu/m^2$ in the charged-particle propagator, and, if the vertices do not give zero, the amplitude increases. It turns out that the condition $q\Gamma = 0$ ($k\Gamma = 0$) can be satisfied if the other two vector particles in the vertex are *real*, that is, on-mass-shell and with physical polarizations. (We shall see this later explicitly.) If so, the term $q_\mu q_\nu/m^2$ in the $C\bar{C} \to \gamma\gamma$ amplitude can be dropped as long as the external particles are real.



https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

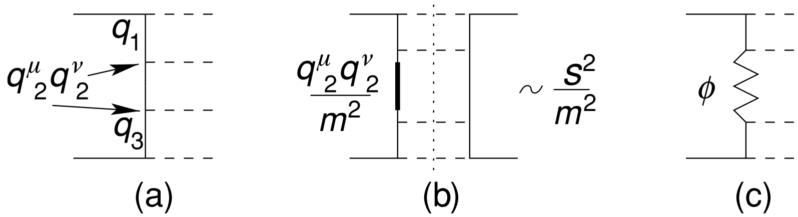
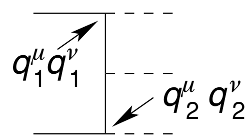
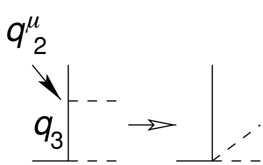


Fig. 16.6 (a) Longitudinal term in the multi-photon annihilation amplitude. (b) Energy growth of the imaginary part of the scattering amplitude. (c) Adding a scalar-particle exchange.

Let us consider more complicated processes. In the process of $C\bar{C}$ annihilation into *three* photons, everything is all right too: $q^\mu q^\nu$ pieces of the propagators of the internal virtual lines q_1 and q_2 multiply the external vertices, and their contributions vanish on the mass shell.



As to the production of four photons (Fig. 16.6) the situation changes. In the amplitude Fig. 16.6(a), the term $q_2^\mu q_2^\nu$ coming from the propagator surrounded by two *virtual* lines q_1 and q_3, does not disappear. As a result, the imaginary part of the amplitude of $C\bar{C}$ scattering via photons displayed in Fig. 16.6(b) grows with energy, and the restoration of the amplitude requires a subtraction term – an arbitrary constant. This means non-renormalizability of the theory, in the language of dispersion relations. Since the convolution of the momentum vector q_2^μ with the vertex $\Gamma(q_2, q_3)$ must be zero when the particle 3 is on the mass shell,



$$q_2^\mu \Gamma_{\mu\rho\sigma}(q_2, q_3) \propto (m^2 - q_3^2). \qquad (16.26)$$

Hence, the propagator of the virtual particle 3 actually cancels in the diagram, giving rise to a reduced diagram with two photons effectively emerging from one point.

As I have already mentioned, the sign of the residue coming from the badly behaving polarization is of 'ghost' type. Therefore, in principle, this unwanted contribution can be cancelled by adding to the theory a normal charged scalar particle φ as shown in Fig. 16.6(c). A simple tuning allows one to make the sum of the diagrams (a) and (c) to have good asymptotics (i.e. to cancel the divergences).

In the electrodynamics of *scalar* particles (as well as in the QFT with interacting fermions and (pseudo)scalars discussed above), in order to

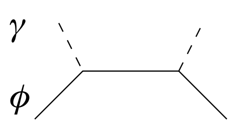
https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

have a renormalizable theory, we have to introduce an *additional interaction.* In the electrodynamics of *vector* particles the introduction of an *additional scalar field* is required.

It turns out that the class of renormalizable theories can be enlarged considerably by generalizing electrodynamics in a more serious way than by just making the photon massive. What is the idea? How could one approach the problem of renormalizability in a more general way, avoiding the necessity of inventing the counter-term diagrams?

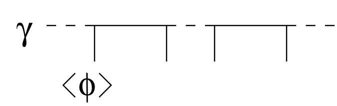
In electrodynamics a 'scalar' component of the vector field cannot be produced even *virtually* due to the condition $k_\mu A_\mu = 0$, which selects three components of the photon field out of the four-vector A_μ. In fact a real photon has not three, but two polarizations; the 'longitudinal' component, $A_\mu(k) \propto k_\mu$, is not produced either (current conservation). Turning to a theory of a vector field with a finite mass, three components appear. Where do they come from? We inserted them into the massive Lagrangian just by hand. Would it not be possible to write the third component separately, so that it would delicately join in, adding to the initially massless two-component photon?

Let us ask ourselves whether a scalar field with zero mass can exist. One may push in an arbitrary number of such particles with $\mathbf{k} = 0$ in the vacuum. The first thing that is likely to emerge is a constant field of ϕ particles – a Bose condensate.



As soon as I introduce the interaction with photons, a photon propagating in a constant scalar field will experience Thomson scattering and continuously accumulate a scattering phase, in other words – it will acquire a finite mass.

Diagrammatically, a photon mixes up with the scalar field; after diagonalization the propagating states possess a non-zero mass. Two components of the photon and the (gradient of the) scalar field combine into a three-component massive vector field.



On the basis of this simple example one can try to construct a theory of massive charged vector particles.

16.6.3 Conservation of current and cancelling diagrams

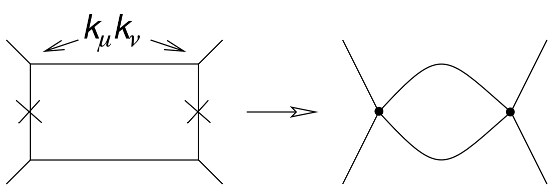
- When we try to construct renormalizable field theories, for particles with spins $\sigma \geq 2$ ultraviolet divergences appear which we do not know how to deal with.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

- In the case of vector particles, $\sigma = 1$, QED has taught us that the conservation of current can help to eliminate divergences. In particular, it is straightforward to construct a theory of a neutral massive vector field.
- Turning to electrodynamics of charged spin one particles, we see that there is no vertex that would provide current conservation with respect to all three vector lines simultaneously.

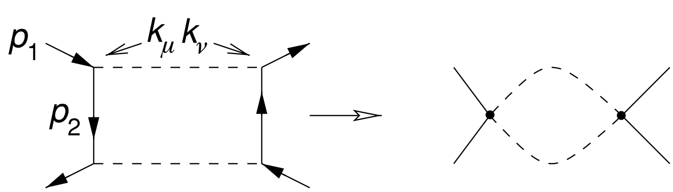
Maybe, there exists a deeper symmetry which provides a stronger current conservation than it can be seen just from the structure of the vertex?

Arriving at the vertices with virtual lines, $k_\mu k_\nu$ gives not zero but a quantity proportional to the departure from the mass shell, cf. (16.26). This cancels the propagators and leads to a topologically simpler diagram:



Is it not possible to invent a theory which would have such a diagram a priori, to help to cancel the 'longitudinal' term?

In electrodynamics this was just the case. Let us take, for example, a box diagram with e^+e^- annihilating into two photons and consider what happens with the 'longitudinal' part of the photon Green function:



(16.27a)

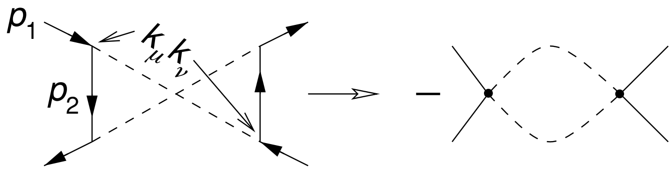
The photon momentum k_μ, multiplying the vertex, produces

$$k_\mu \gamma^\mu = \hat{p}_1 - \hat{p}_2 = -(m - \hat{p}_1) + (m - \hat{p}_2);$$

the first term gives zero when it acts on a spinor describing an on-mass-shell electron, $(m - \hat{p}_1)u(p_1) = 0$ (Dirac equation), while the second term cancels the virtual fermion propagator, resulting in a reduced graph,

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

(16.27a). However, in QED we have also the second Feynman diagram,



(16.27b)

Two diagrams cancel in the sum.

In the case of *scalar* charges, the cancellation between reduced graphs (16.27) is incomplete (because of the momentum dependence of the $\gamma\varphi\varphi$ vertex). In this situation the diagrams with the *four-vertex*, $\gamma^2\varphi^2$, see (16.24), participate to make the $k_\mu k_\nu$ piece go away.

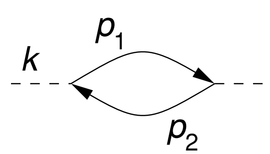
I wanted to demonstrate in terms of diagrams, what is required from the new theories.

16.7 Yang–Mills theory

The current conservation in quantum electrodynamics had one more very important interpretation: QED was constructed in such a way that a photon mass could not appear there due to gauge invariance.

We want to construct a theory where, again, summing the full set of diagrams, we get cancellation of dangerous $k_\mu k_\nu$ terms. This would ensure current conservation in spite of the impossibility to have conservation directly in the three-particle vertex.

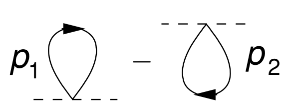
What is the connection to the photon mass?

$\Pi^{\mu\nu}(k) =$ 

Multiplying the polarization operator by the photon momentum,

$$k_\mu\gamma^\mu = \hat{p}_1 - \hat{p}_2 = (m - \hat{p}_2) - (m - \hat{p}_1),$$

we obtain, diagrammatically,

$k_\mu\Pi^{\mu\nu}(k) =$ 

$= 0.$

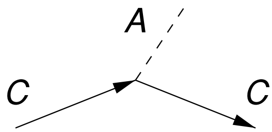
The transversality of the photon polarization operator and, therefore, non-appearance of photon mass, is a particular case of the cancellation of divergences.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

16.7.1 Electrodynamics of massless vector particles

Guided by the idea that the absence of masses is already a hint to serious cancellations, our first hope is to try to build up a theory of *massless* charged particles.

$$\mathcal{L}_{\text{int}} \sim \bar{C}CA \tag{16.28}$$

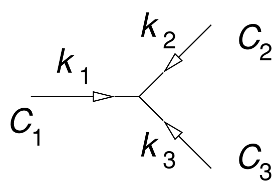


What is a charged particle? The charged field C can be represented as a linear combination of two neutral fields, C_1 and C_2,

$$C = \frac{C_1 + C_2}{\sqrt{2}}, \quad \bar{C} = \frac{C_1 - C_2}{\sqrt{2}},$$

with positive- and negative-charge parity, correspondingly. Since a photon has negative charge parity, inserting this in the vertex (16.28), there remain the transitions $C_1 + \gamma \to C_2$ and $C_2 + \gamma \to C_1$.

If the charged particles are massless as the photon is, it is natural to consider three neutral fields C_1, C_2, C_3 ($C_3 \equiv A$) on equal footing, with the interesting interaction



(16.29)

which is just another representation for the electromagnetic vertex (16.25).

How to write an interaction so that the current is conserved? Let us construct the tensor

$$\Gamma^{\mu_1\mu_2\mu_3}(k_1, k_2, k_3) = g^{\mu_1\mu_2} p_{(3)}^{\mu_3} + g^{\mu_1\mu_3} p_{(2)}^{\mu_2} + g^{\mu_2\mu_3} p_{(1)}^{\mu_1},$$

where p is a linear combination of the momenta k_i, e.g.

$$p_{(1)}^{\mu} = a_{11}k_1^{\mu} + a_{12}k_2^{\mu} + a_{13}k_3^{\mu}.$$

(It is dangerous to use higher powers of momenta, since this would increase the divergence in the loop integrals.) Multiplying by the momentum, say, k_3, we get

$$k_{3,\mu_3}\Gamma^{\mu_1\mu_2\mu_3} = g^{\mu_1\mu_2}(k_3\, p_{(3)}) + k_3^{\mu_1} p_{(2)}^{\mu_2} + k_3^{\mu_2} p_{(1)}^{\mu_1}. \tag{16.30}$$

We need this expression to vanish upon multiplication by the *physical polarization* vectors $e_{\mu_i}^{\lambda_i}(k_i)$, $i = 1, 2$, which satisfy $(e_{\mu_i}^{\lambda_i}\, k_i^{\mu_i}) = 0$, when

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

particles 1, 2 are on the mass shell: $k_1^2 = k_2^2 = 0$,

$$e_{\mu_1}^{\lambda_1}(k_1) e_{\mu_2}^{\lambda_2}(k_2) \cdot k_{3,\mu_3} \Gamma^{\mu_1\mu_2\mu_3}(k_1, k_2, k_3) = 0.$$

Since we may choose the polarization vectors $e(k_1)$, $e(k_2)$ orthogonal to the plane formed by the four-momenta $\{k_1, k_2, k_3\}$, we must have $(k_3 p_{(3)}) = 0$ in (16.30). Representing $p_{(3)}$ as a linear combination

$$\begin{aligned} p_{(3)} &= a(k_1 - k_2) + b\, k_3 \qquad (-k_3 = k_1 + k_2), \\ (k_3 p_{(3)}) &= -a[\, k_1^2 - k_2^2\,] \;+\; b\, k_3^2 = 0, \end{aligned}$$

we conclude that $b = 0$ (recall that $k_1^2 = k_2^2 = 0$). Hence,

$$p_{(3)}^{\mu} = a(k_1^{\mu} - k_2^{\mu}),$$

and the resulting form of the vertex is (up to an overall constant)

$$\Gamma^{\mu_1\mu_2\mu_3} = g^{\mu_1\mu_2}(k_1 - k_2)^{\mu_3} + g^{\mu_1\mu_3}(k_3 - k_1)^{\mu_2} + g^{\mu_2\mu_3}(k_2 - k_3)^{\mu_1}. \tag{16.31}$$

This is the sum of three electromagnetic vertices for a scalar particle.

Let us look again at the product (16.30):

$$k_{3,\mu_3}\Gamma^{\mu_1\mu_2\mu_3} = g^{\mu_1\mu_2}(-k_1^2 + k_2^2) + k_3^{\mu_1}(k_3 - k_1)^{\mu_2} + k_3^{\mu_2}(k_2 - k_3)^{\mu_1}.$$

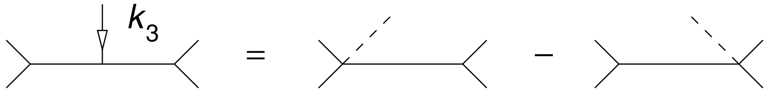
Substituting $k_3 = -(k_1 + k_2)$ we obtain

$$\begin{aligned} k_{3,\mu_3}\Gamma^{\mu_1\mu_2\mu_3} &= g^{\mu_1\mu_2}(-k_1^2 + k_2^2) + k_1^{\mu_1}k_1^{\mu_2} - k_2^{\mu_2}k_2^{\mu_1} \\ &= -(G^{-1})^{\mu_1\mu_2}(k_1) + (G^{-1})^{\mu_1\mu_2}(k_2), \end{aligned}$$

where G^{-1} is the inverse transverse propagator of a vector particle,

$$(G^{-1})^{\mu_1\mu_2}(k) = g^{\mu_1\mu_2}k^2 - k^{\mu_1}k^{\mu_2}.$$

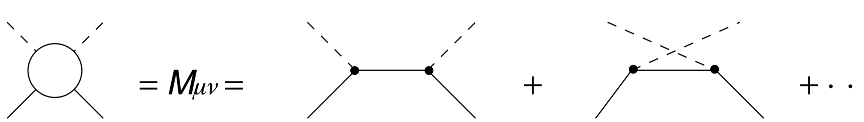
Thus, multiplying the vertex by the momentum of one of the particles ('photon line'), we have a difference of inverse propagators of two others ('charged lines'). We get the difference of the 'reduced' diagrams,



similar to that in quantum electrodynamics.

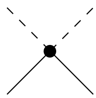
Is this everything we need? As often happens in bosonic theories, one type of the vertex, (16.29), is not enough.

Recall the second-order QED interaction amplitude:



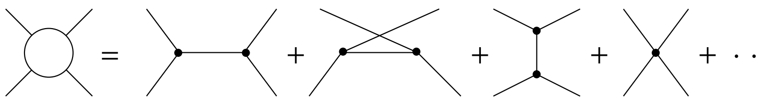
https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

For spinor particles, $\sigma = \frac{1}{2}$, the sum of two Born amplitudes satisfied the current conservation condition, $k^\mu M_{\mu\nu} = 0$. In the case of scalar charges,



however, it was necessary to add a four-particle point interaction, because the triple vertex was *linear* in momenta.

Exactly the same situation is there in our vector theory: for vector currents to be conserved, one has to introduce a dimensionless four-particle vertex:

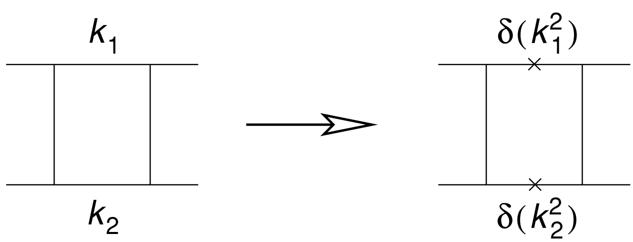


(16.32)

16.7.2 Feynman diagrams and unitarity: Faddeev–Popov ghost

Having done so, the theory is formulated. As soon as we have Born terms, perturbation theory can be constructed. How do we do this? A method which always works is to use the unitarity conditions. We may take the on-mass-shell Born amplitude on the r.h.s. of (16.32), square it, and reconstruct the higher-order scattering amplitude by its imaginary part, employing the dispersion relation (with one subtraction). In principle, there is no problem in carrying out this procedure; still, this is a rather cumbersome way. Usually, we draw a Feynman diagram instead, which coincides with the dispersion expression. However, in the present case *this is not true.*

To see what goes wrong, let us draw a diagram and take the imaginary part by putting the cut particle lines on the mass shell, replacing the cut propagators by the delta functions:



Which polarization states will emerge in the intermediate state? As a rule, we expect two transverse polarizations to appear from the left and the right, see Fig. 16.7(a), while the contributions of the *longitudinal* polarizations have to be zero, owing to current conservation. However, our four-particle amplitude (as well as the three-particle vertex Γ above) satisfies the current conservation with respect to each external line only

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

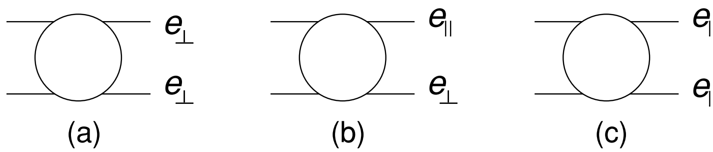


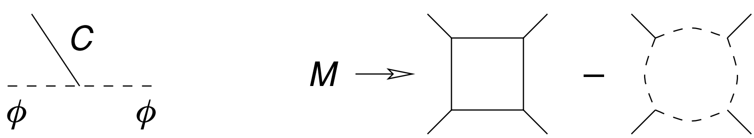
Fig. 16.7 Various combinations of intermediate state polarizations.

if *all* the other lines are on-mass-shell and with physical polarizations:

$$k_1^{\mu_1} M_{\mu_1\mu_2\mu_3\mu_4} \cdot e^{\mu_2}(k_2) e^{\mu_3}(k_3) e^{\mu_4}(k_4) = 0; \;\; (e(k_i)k_i) = 0, \;\; k_i^2 = 0, \;\; i = 2, 3, 4.$$

Therefore, a situation like Fig. 16.7(b) (three transverse polarizations out of four) will not appear in the imaginary part. The intermediate state with *two* longitudinal polarizations shown in Fig. 16.7(c) will, however, be present. Writing unitarity conditions, we summed, of course, only over physical states, $e_\perp$. Thus, the imaginary part of the Feynman diagram differs from that in unitarity conditions. In fact we face here the first case when the Feynman diagrams are not correct literally.

What can we do? A technical problem appears: is it possible, nevertheless, to represent the correct dispersion result in terms of diagrams, which would provide us with the means to carry out calculations? The answer, at the one-loop level, was given by Feynman. He suggested to introduce a fictitious massless scalar particle (essentially, a single state $e_\parallel$ can be considered as a particle with $\sigma = 0$), and *subtract* the corresponding loop,



A general prescription was given by Faddeev and Popov, namely: one should introduce in the theory an additional field ϕ, with a normal vertex, and handle it as a *fermion*, i.e. count every ϕ loop with a minus sign.

This example demonstrates that the Feynman technique ceases to be transparent.

16.7.3 Gauge invariance and Lagrangian

Now we are going to discuss a different approach that leads to the theory of interacting vector fields – the Yang–Mills theory.

So far we have considered only vector mesons. It is easy to introduce also fermions. We should not forget, however, that up to now our construction was symmetrical with respect to the indices $1 \to 2 \to 3$ of the vector fields. The symmetry of the theory with respect to the rotation in the 'space' of

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

the three field components can be preserved if we introduce a doublet of fermions, similar to a nucleon with two isospin states, $N = (p, n)$.

But first let us elucidate the relation of the Yang–Mills theory with electrodynamics. In the free fermion Lagrangian,

$$\mathcal{L}_{\psi 0} = \bar{\psi}(x)\left(i\gamma_\mu \frac{\partial}{\partial x_\mu} - m\right)\psi(x),$$

one can substitute

$$\psi \to \psi' = \mathrm{e}^{i\alpha}\psi, \quad \bar{\psi} \to \bar{\psi}' = \mathrm{e}^{-i\alpha}\bar{\psi}, \tag{16.33a}$$

with $\alpha = \text{const}$, without affecting the equation of motion. This is, essentially, the expression of the fact that a fermion and an antifermion can not transform one into the other.

A prominent feature of quantum electrodynamics is that in this theory one is allowed to carry out the transformation (16.33a) of the fermion field, with the phase depending on the space–time point, $\alpha = \alpha(x)$. (In QED one can tell a fermion from an antifermion *locally*.) To keep the action invariant under such transformation,

$$\mathcal{L}_{\psi 0} \to \mathcal{L}'_{\psi 0} = \bar{\psi}'\left(i\gamma_\mu \frac{\partial}{\partial x_\mu} - m\ - \gamma_\mu \frac{\partial \alpha(x)}{\partial x_\mu}\right)\psi',$$

one has to add to the Lagrangian the interaction with a *vector field*,

$$\mathcal{L}_{\psi 0} \implies \mathcal{L}_\psi = \bar{\psi}\left(i\gamma_\mu \frac{\partial}{\partial x_\mu} - m + \gamma_\mu A_\mu\right)\psi,$$

and simultaneously transform this field as

$$A_\mu(x) \ \to\ A'_\mu(x) = A_\mu(x) + \frac{\partial \alpha(x)}{\partial x_\mu}. \tag{16.33b}$$

The field A_μ changes, turns into some other field A'_μ. Hence, if A is the dynamical variable itself, its proper action must be invariant with respect to the gradient transformation (16.33b). Such an invariant photon Lagrangian, as you know, is given by the square of the antisymmetric electromagnetic stress tensor:

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu}, \qquad \mathcal{L} = \mathcal{L}_\psi + \frac{1}{4e^2} F_{\mu\nu}F^{\mu\nu}.$$

From the point of view of such an approach the Yang–Mills theory is just an exact repetition of the same logic.

Let us introduce two fermions in the theory:

$$\mathcal{L}_\psi = \mathcal{L}_{\psi_1} + \mathcal{L}_{\psi_2} = \bar{\boldsymbol{\psi}}\left(i\gamma_\mu \frac{\partial}{\partial x_\mu} - m\right)\boldsymbol{\psi} \tag{16.34}$$

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

where $\boldsymbol{\psi}$ is a column,

$$\boldsymbol{\psi}(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix},$$

like a nucleon with two isospin components, $N = (p, n)$.

If there is degeneracy in the system (even with account of the interaction), I can consider these two levels as one field. With the help of a unitary matrix, I can redefine (globally) who I call the 'proton' and who the 'neuteron'. Let us try to construct a theory in which such a 'redefinition' can be carried out locally.

Consider a transformation,

$$\psi'(x) = S(x)\psi(x), \quad \bar{\psi}'(x) = \bar{\psi}'(x)S^{-1}(x); \qquad S^\dagger(x) = S^{-1}(x),$$

with $S(x)$ a unitary matrix.

Let us take the Lagrangian

$$\mathcal{L}_\psi = \bar{\psi}(x)\left(i\gamma_\mu \frac{\partial}{\partial x_\mu} - m + i\hat{A}(x)\right)\psi(x), \quad \hat{A} = \gamma_\mu \mathbf{A}_\mu,$$

with $\mathbf{A}_\mu$ an *anti-hermitean* 2×2 matrix of vector fields (an analogue of photon), and rotate the spinor fields in it:

$$\mathcal{L}_\psi \to \bar{\psi}(x)\left(i\gamma_\mu \frac{\partial}{\partial x_\mu} - m + i\gamma_\mu S^{-1}(x)\frac{\partial S(x)}{\partial x_\mu} + iS^{-1}\hat{A}S\right)\psi(x). \quad (16.35)$$

The Lagrangian stays invariant if the $\mathbf{A}$ field also transforms as follows:

$$\mathbf{A}_\mu \;\to\; \mathbf{A}'_\mu \;=\; S\mathbf{A}_\mu S^{-1} - \frac{\partial S}{\partial x_\mu}S^{-1}; \qquad (16.36a)$$

$$\mathbf{A}_\mu \;=\; S^{-1}\mathbf{A}'_\mu S + S^{-1}\frac{\partial S}{\partial x_\mu}. \qquad (16.36b)$$

In order to have an invariant theory, the invariance of the action of the fields $\hat{A}$ is required with respect to the transformation (16.36). One observes that the field strength tensor

$$\mathbf{G}_{\mu\nu} = \frac{\partial \mathbf{A}_\nu}{\partial x_\mu} - \frac{\partial \mathbf{A}_\mu}{\partial x_\nu} + [\mathbf{A}_\mu\, \mathbf{A}_\nu] \;=\; \mathbf{F}_{\mu\nu} + [\mathbf{A}_\mu\, \mathbf{A}_\nu]$$

transforms *homogeneously*, $G'_{\mu\nu} = SG_{\mu\nu}S^{-1}$, and therefore the vector field Lagrangian

$$\mathcal{L}_A = \frac{1}{4g^2}\,\mathrm{Tr}\big(\mathbf{G}_{\mu\nu}\mathbf{G}^{\mu\nu}\big) \qquad (16.37)$$

is invariant under the gauge transformation (16.36).

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

How is the field $\mathbf{A}$ built up? The 2×2 matrix has four components:

$$\mathbf{A}_\mu = I \cdot A_\mu^0 + \sum_{a=1}^{3} \tau^a A_\mu^a,$$

with τ^a the Pauli matrices. The field A^0 here corresponds to the usual electromagnetic interaction. It is generated by the (Abelian) transformation subgroup $U(1)$ contained in the group of unitary 2×2 matrices: $U(2) = U(1) \times SU(2)$. Let us exclude the photon and restrict ourselves to fields $\operatorname{Tr} \mathbf{A} = 0$, i.e. to A_μ^a with $a = 1, 2, 3$. The full Lagrangian of the theory will have the structure

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_\psi + \mathcal{L}_A = \mathcal{L}_{0\psi} + \bar{\psi}\gamma_\mu \sum_a i A_\mu^a \tau_a \psi + \mathcal{L}(F_{\mu\nu}) \\ &\quad + \frac{1}{2g^2} \operatorname{Tr}(F_{\mu\nu}[A_\mu A_\nu]) + \frac{1}{4g^2} \operatorname{Tr}([A_\mu A_\nu][A_\mu A_\nu]). \end{aligned} \tag{16.38}$$

The last two terms generate those two vertices that we have drawn above in (16.32). In terms of rescaled fields $\tilde{A} = A/g$, the three-boson and fermion–vector boson vertices are proportional to the coupling constant g, and the four-boson vertex proportional to g^2.

We considered here $SU(2)$ gauge symmetry. The discussed scheme can be applied, however, to any unitary group $SU(N)$. In particular, for $N = 3$ we introduce quarks of three *colours* and obtain $N^2 - 1 = 8$ *gluons*, – fundamental fields of the *quantum chromodynamics*.

16.7.4 Essential gauge invariance and cyclic variables

How can one work with such gauge theories? One of the ways to construct a quantum theory is to use the functional integral:

$$\int \mathrm{e}^{i \int \mathcal{L}[A,\psi] dx}\, d\psi\, d\bar{\psi}\, dA_\mu. \tag{16.39}$$

Let us say a few words about the appearing problems. Making an effort to have a gauge invariant theory, we arrived at an undefined system, since we introduced more variables than it has in reality.

We operate with fields $\mathbf{A}_\mu$ which can be substituted by other fields, $S^{-1}\mathbf{A}'_\mu S + S^{-1}\frac{\partial S}{\partial x_\mu}$, not changing the action. Hence, the Lagrangian does not depend on S, and this shows that in the functional integral (16.39) there is integration over a large number of superfluous unphysical variables. The relation (16.36a) allows us to choose new physical fields in different ways, e.g. so that $A'_0 \equiv 0$ (by solving the equation $\mathbf{A}_0 = S^{-1}\frac{\partial S}{\partial x_0}$).

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

Let us set $A_0^a = 0$ and keep only the integration over three-vector potentials A_i^a. Then, $G_{00}^a = 0$, $G_{0i}^a = \partial A_i^a/\partial x_0 \equiv \dot{A}_i^a$, and what remains is a quite reasonable Lagrangian:

$$\mathcal{L}(A_i, \psi) = \mathcal{L}_\psi - \frac{1}{2g^2}\Big[\mathrm{Tr}\big(\dot{\mathbf{A}}_i\dot{\mathbf{A}}_i\big) - \tfrac{1}{2}\,\mathrm{Tr}\big(\mathbf{G}_{ik}\mathbf{G}_{ik}\big)\Big]; \quad (16.40\text{a})$$

$$\mathbf{G}_{ik} = \frac{\partial \mathbf{A}_k}{\partial x_i} - \frac{\partial \mathbf{A}_i}{\partial x_k} + \big[\mathbf{A}_i\,\mathbf{A}_k\big]. \quad (16.40\text{b})$$

(The minus sign in front of the boson Lagrangian is due to my choice of $\mathbf{A}_i$ as anti-hermitean matrices.) Since $\mathbf{G}_{ik}$ do not contain the time derivative, the last term in (16.40a) plays the rôle of the potential energy of the self-interacting fields A_i^a; the term $(\dot{A}_i^a)^2$ is their kinetic energy.

Formally, this is not a relativistically invariant description, and usually one chooses an invariant gauge fixing condition $k_\mu A_\mu^a = 0$.

We saw already in electrodynamics that there are two types of gauge invariance.

- Firstly, the current conservation makes it possible to impose the condition $k_\mu A_\mu = 0$ in order to select three components of four.
- Secondly, a deeper consequence of the gauge invariance lies in the fact that the photon is massless, and therefore a real photon has only *two* field components.

The condition $(\mathbf{k}\cdot\mathbf{A}^\lambda) \equiv k_i A_i^\lambda = 0$ $(\lambda = 1, 2)$ applies only to *real photons*. In diagrams with *virtual particles* the longitudinal component of the electromagnetic potential, $(\mathbf{k}\cdot\mathbf{A}^{\|}) \neq 0$, is acting and represents the Coulomb field. This *essential* invariance allows us to write

$$A_i(x) = B_i(x) + \frac{\partial\varphi(x)}{\partial x_i}; \quad \mathrm{div}\,\mathbf{B} \equiv \frac{\partial}{\partial x_i}B_i = 0.$$

The field φ here is absolutely essential inside the diagrams but does not correspond to a free particle (does not 'fly away'); real photons are described solely by the two-component field B_i.

Our aim is to show that the theory with the Lagrangian (16.40) describes indeed massless vector particles. The equation $\mathbf{A}_0 = S^{-1}\frac{\partial S}{\partial x_0}$ that we have used to eliminate the scalar component of the potential, fixed S up to unitary matrices not depending on time. Arbitrary rotations depending on x_i are still at our disposal. This allows us to look for an additional invariance in our Lagrangian where no A_0 is present and everything is

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

determined. Let us write A_i in the form

$$\mathbf{A}_i \;=\; v^{-1}\mathbf{B}_i\, v + v^{-1}\frac{\partial v}{\partial x_i} \tag{16.41}$$

and fix B_i^a by the condition

$$\frac{\partial}{\partial x_i} B_i^a = 0 \tag{16.42}$$

in the same way as in electrodynamics. It is important to stress that this is *not* a gauge transformation, just an attempt to separate physical degrees of freedom. Note that since (16.41) has the *form* of a gauge transformation, $\mathrm{Tr}(\mathbf{G}_{ik}^2)$ does not depend on v. From the point of view of classical mechanics, this makes v a *cyclic variable*, i.e. the one that has the kinetic but not the potential energy (n the same way as the Coulomb field in electrodynamics). A cyclic variable q in mechanics,

$$\mathcal{L} = \dot q^2 + \sum_k \dot q_k^2 - U(q_k),$$

changes linearly with time:

$$\dot q = \text{const}, \quad q = ct + b.$$

This is actually the real difficulty of quantizing a gauge theory: there are variables which do not oscillate but increase with time. It goes without saying that if I measure components of the *field strength*, $\dot A_i^a$, G_{ik}^a, all will be fine. Nevertheless, a problem remains: perturbation theory cannot be applied to such a Lagrangian possessing a cyclic variable.

Let us omit fermions and start with a free Lagrangian

$$\mathcal{L}_{A0} \;\propto\; \frac{1}{2}\sum_i \dot A_i^2 - \frac{1}{4}\sum_{i,k}\left(\frac{\partial A_i}{\partial x_k} - \frac{\partial A_k}{\partial x_i}\right)^2. \tag{16.43}$$

In the momentum representation,

$$A_i(x) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}} a_i(x_0,\mathbf{k})\,\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}, \quad a_i(x_0,-\mathbf{k}) = a_i^*(x_0,\mathbf{k}),$$

we get

$$\mathcal{L} \propto \tfrac{1}{2}\big(\dot a_i \dot a_i^* - |[\mathbf{k}\,\mathbf{a}]|^2\big) = \tfrac{1}{2}\big(\dot a_i \dot a_i^* - (\delta_{ij}\mathbf{k}^2 - k_i k_j)a_i a_j^*\big), \tag{16.44}$$

showing that the longitudinal component of the field, $\mathbf{a} \propto \mathbf{k}$, does not enter the potential energy. On the definite energy states, $\mathbf{a}(x_0,\mathbf{k}) = \mathrm{e}^{-ik_0x_0}\mathbf{C}(\mathbf{k})$, the Lagrangian (16.44) turns into

$$\mathcal{L} \;\propto\; \tfrac{1}{2}\big[\delta_{ij}\,k_0^2 \;-\; (\delta_{ij}\mathbf{k}^2 - k_i k_j)\big]C_i C_j^*.$$

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

The expression in the square brackets is the inverse propagator:

$$D^{-1}_{ij}(k) \;=\; \delta_{ij}\, k^2 + k_i k_j, \tag{16.45a}$$

$$D_{ij}(k) \;=\; \frac{1}{k^2}\left(\delta_{ij} - \frac{k_i k_j}{k_0^2}\right). \tag{16.45b}$$

You can verify that (16.45b) is indeed an inverse tensor to (16.45a):

$$\sum_{\alpha=1}^{3} D^{-1}_{i\alpha}(k)\, D_{\alpha j}(k) \;=\; \delta_{ij}.$$

The free Green function acquires a pole in energy k_0, and all the integrals will diverge at $k_0 \to 0$, which divergence corresponds to the linear growth with time of the longitudinal component of the vector field. The Green function (16.45b) can be split into the sum of the transverse and longitudinal contributions:

$$D_{ij}(k) = \frac{1}{k^2}\left(\delta_{ij} - \frac{k_i k_j}{\mathbf{k}^2}\right) \;+\; \frac{k_i k_j}{k_0^2\, \mathbf{k}^2}\,. \tag{16.46}$$

The transverse part of the propagator, $k_i \cdot D^{\perp}_{ij}(k) = 0$, looks reasonable; the longitudinal part contains an infrared divergence. Evidently, the latter has to be defined additionally using some sort of $i\epsilon$ (or 'principal value') prescription of how to deal with the singularity at $k_0 = 0$.

We conclude that a formulation without superfluous degrees of freedom faces some difficulties owing to the increase of the longitudinal component of the field with time.

What can we expect from the cyclic variable φ in quantum mechanics?

In classical mechanics it can be fixed arbitrarily. Since, however, the equation of motion is $\ddot{\varphi} = \dot{\pi} = 0$, the momentum is conserved, and the wave function with a definite momentum $\mathrm{e}^{i\pi\varphi}$ is a stationary state. The ground state corresponds to $\pi = 0$; it is the S-state in the φ variable. In other words, the ground state does not depend on the cyclic coordinate. In terms of operators,

$$\frac{\delta \mathcal{L}}{\delta \dot{v}} \big|\Omega\big\rangle \;=\; 0,$$

where Ω is any state of physical fields.

How does this S-state appear in terms of the action? We integrate over all possible field trajectories in the functional integral. The usual saying goes as follows: whichever field configuration is introduced at $t = -\infty$, in an infinite time the system arrives at the ground state. The functional

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

integral up to $A(t)$,

$$\int_{-\infty}^{A(t)} \mathrm{e}^{i\int \mathcal{L}[A(x)]\,dx}\, dA,$$

is just the wave function of the vacuum. If, however, a certain field A does not enter anything but the terms with *derivatives* in the Lagrangian, the integral does not depend on the value of A; the integrand contains only the *differences* over which the integral is taken. Hence, the system occurs automatically in a S-state in cyclic variables which enter only the kinetic energy.

16.7.5 Vacuum in the Yang–Mills theory

What does this tell us about the vacuum in our gauge theory? Let us calculate the time derivative of the field in the form (16.41):

$$\dot{\mathbf{A}}_i = v^{-1}\dot{\mathbf{B}}_i v + v^{-1}\mathbf{B}_i \dot{v} + (\dot{v^{-1}})\mathbf{B}_i v + \frac{\partial}{\partial t}\left(v^{-1}\frac{\partial v}{\partial x_i}\right).$$

Making use of unitarity, $v \cdot v^{-1} = 1 \Rightarrow v(\dot{v^{-1}}) + \dot{v}v^{-1} = 0$,
we obtain

$$\dot{\mathbf{A}}_i = v^{-1}\big[\dot{\mathbf{B}}_i + \nabla_i(B)f\big]v, \tag{16.47}$$

where $f \equiv \dot{v}v^{-1}$ and

$$\nabla_i(B)f \;\equiv\; \frac{\partial f}{\partial x_i} + [\mathbf{B}_i f]. \tag{16.48}$$

Roughly speaking, f is the time derivative of the 'phase' of the unitary variable v. In electrodynamics where v is a number, $v = \exp i\alpha(x)$, this is literally true. The outstanding factors v, v^{-1} cancel under the trace in (16.40a), and for the kinetic energy we derive

$$T = -\tfrac{1}{2}\,\mathrm{Tr}(\mathbf{E}_i)^2; \qquad \mathbf{E}_i \equiv \dot{\mathbf{B}}_i + \nabla_i(B)f. \tag{16.49}$$

Here $E_i^a = \delta\mathcal{L}/\delta\dot{B}_i^a$ are analogous to 'electric fields'. Since f enters the kinetic energy (16.49) only, the momentum corresponding to the cyclic coordinate reads

$$\boldsymbol{\pi} \equiv \frac{\delta T}{\delta f} \;=\; \nabla_i(B)\big(\dot{\mathbf{B}}_i + \nabla_i(B)f\big) = \nabla_i(B)\mathbf{E}_i. \tag{16.50}$$

Here we have used the definition of the 'long' (covariant) derivative (16.48) and the transversality condition for the B field (16.42). As we have discussed above, the momentum $\boldsymbol{\pi}$ is conserved and should be set to zero in the vacuum state.

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

We want to extract the *transversal part* of the 'electric fields' and treat them as canonical momenta $\boldsymbol{\pi}_i$ conjugated to dynamical coordinates $\mathbf{B}_i$:

$$\mathbf{E}_i = \boldsymbol{\pi}_i + \frac{\partial \phi}{\partial x_i}, \quad \frac{\partial}{\partial x_i}\boldsymbol{\pi}_i = 0. \tag{16.51}$$

To exclude the longitudinal part, we combine (16.51) and (16.49):

$$\left(\frac{\partial}{\partial x_i}\right)^2 \phi \;=\; \frac{\partial}{\partial x_i}\mathbf{E_i} = \Box(B) f, \tag{16.52a}$$

where the operator $\Box$ is defined as

$$\Box(B) \;\equiv\; \frac{\partial}{\partial x_i}\nabla_i(B) = \nabla_i(B)\frac{\partial}{\partial x_i}. \tag{16.52b}$$

Now we are ready to set the cyclic momentum (16.50) to zero:

$$0 = \nabla_i(B)\mathbf{E}_i = \nabla_i(B)\left(\boldsymbol{\pi}_i + \frac{\partial \phi}{\partial x_i}\right) = [\mathbf{B}_i\,\boldsymbol{\pi}_i] + \Box(B)\varphi\,,$$

which gives

$$\phi = \frac{1}{\partial_i^2}\Box(B) f \;=\; -\frac{1}{\Box(B)}\rho, \quad \rho = [\mathbf{B}_i\,\boldsymbol{\pi}_i]. \tag{16.53}$$

This is an analogue of the usual Coulomb equation $\phi = -(\partial_i^2)^{-1}\rho$, with ρ the charge density. When the momentum π is different from zero, after excluding the cyclic variable, an additional centrifugal energy appears. This is nothing but the Coulomb interaction energy. While in electrodynamics this is the energy of *external charges*, $\rho = \rho_{\mathrm{ext}}$, in our case the fields A are charged themselves, and we have the Coulomb energy of the self-interacting Yang–Mills fields, produced by the 'charge density'

$$\rho = [\mathbf{B}_i\boldsymbol{\pi}_i]. \tag{16.54}$$

We want to find the Hamiltonian of the system,

$$H = \int d^3x\,\mathcal{H}(x), \quad \mathcal{H}(x) = -\frac{1}{2g^2}\,\mathrm{Tr}\big(\mathbf{E}_i^2(x) + \tfrac{1}{2}\mathbf{G}_{ik}^2(x)\big)\,.$$

For that we square the electric field (16.51):

$$\int d^3x\,\mathbf{E}_i^2(x) \;=\; \int d^4x\,\big(\boldsymbol{\pi}_i^2(x) - \phi(x)\partial^2\phi(x)\big)\,.$$

Finally, substituting (16.53) renders the Hamiltonian density:

$$\mathcal{H} \;=\; -\frac{1}{2g^2}\,\mathrm{Tr}\left(\boldsymbol{\pi}_i^2 + \rho\,\frac{1}{\boldsymbol{\Delta}(B)}\,\rho + \tfrac{1}{2}\mathbf{G}_{ik}^2\right); \tag{16.55}$$

$$\boldsymbol{\Delta}(B) \;\equiv\; -\Box(B)\,\partial^{-2}\,\Box(B). \tag{16.56}$$

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

This calculation completes the usual verification of the fact that the initial Lagrangian describes interacting massless particles. This is not the end of the story, however.

Usually, the vacuum state is characterized by small oscillations of transverse fields. In our case it also contains *randomly oscillating* (S-state!) longitudinal Coulomb fields. We see that the effective interaction turns out to be rather unusual, non-local, because we tried to get rid of these random longitudinal fields.

If we neglect B fields in (16.56), we get the usual Coulomb interaction between charges:

$$\mathbf{\Delta}(0) = -\partial^2, \quad \frac{1}{\mathbf{\Delta}(0)} \Rightarrow \frac{1}{|\mathbf{x}-\mathbf{x}'|}.$$

The dependence on B shows that the *instantaneous* Coulomb interaction is modified by the presence of vacuum fluctuations of transverse (physical) fields. Instead of being purely dynamical, our problem of the structure of the vacuum state becomes an almost *statistical* one. It is this specificity that reverses the behaviour of the effective coupling of the theory (asymptotic freedom), see (16.22).

16.8 Asymptotic freedom

Now we are going to find out how the effective charge behaves in the Yang–Mills theory. We will follow the path suggested by Khriplovich (1969), who has calculated the invariant charge in Coulomb gauge even before the discovery of asymptotic freedom in non-Abelian theories (Gross and Wilczek, 1973; Politzer, 1973). He looked at the first $\mathcal{O}(g^2)$ correction to the vacuum energy in the presence of two infinitely heavy charges.

We add an external source $\rho_h = g\delta^3(\mathbf{x}-\mathbf{x}_0)$ to the charge density (16.54), substitute $\rho+\rho_h$ into the second term of the Hamiltonian density (16.55) and evaluate the correlator between two static sources placed at $\mathbf{x}_0 = \mathbf{x}_1$ and $\mathbf{x}_2$.

The vacuum energy of the system contains two contributions quadratic in ρ_h. The first is just 'classical' Coulomb energy of the sources given by the correlator of the external charge densities, $\rho_h \cdot \rho_h$ averaged over transverse fields $B_\perp$ in the vacuum:

$$V^c_{\text{Coul}} = \rho_h \Big\langle 0 \Big| \frac{1}{\mathbf{\Delta}(B)} \Big| 0 \Big\rangle \rho_h. \tag{16.57}$$

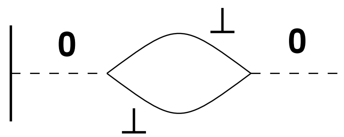
The second is quantum correction due to the mixing term $\rho_h \cdot \rho$ in the Hamiltonian. It leads to transition of the Coulomb field of the external charge into a pair of transverse fields (gluons with physical polarizations),

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

and back again:

$$V^q_{\text{Coul}} = \sum_n \frac{|V_{0n}|^2}{E_0 - E_n}, \quad V_{0n} = \rho_h \Big\langle 0 \Big| \frac{1}{\mathbf{\Delta}(B)}\, \rho \Big| n \Big\rangle,$$

where the sum runs over the energies of the intermediate two-gluon state. Here we can replace $\mathbf{\Delta}(B) \simeq \mathbf{\Delta}(0)$, and this contribution reduces to the Feynman diagram with two gluons in the intermediate state:



A standard calculation yields in the momentum space

$$V^q_{\text{Coul}} = \frac{g^2}{\mathbf{k}^2}\left(-\frac{C_2}{3}\cdot\frac{g^2}{16\pi^2}\ln\frac{\Lambda^2_{\text{UV}}}{\mathbf{k}^2}\right), \tag{16.58a}$$

with C_2 the number appearing from the square of the matrix commutator (the Casimir operator $C_2 = N$ for the $SU(N)$ group). If we add fermion fields (n_f families of quarks) into the game, additional quark loops appear, and the coefficient in (16.58a) gets modified as follows

$$-\frac{N}{3} \implies -\left(\frac{N}{3} + \frac{2}{3}n_f\right). \tag{16.58b}$$

This quantum correction is due to a virtual decay into physical states and corresponds to *screening* all right, having the same sign as in QED and elsewhere.

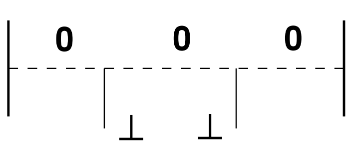
Now we return to the ‘classical’ piece (16.57). To find the $\mathcal{O}(g^2)$ correction due to vacuum fields we need to expand the operator $\mathbf{\Delta}^{-1}$ to the second order in B. First we calculate approximately the inverse of the operator $\Box$,

$$\Box^{-1}(B) = \partial^{-2} - \partial^{-2}\big[\mathbf{B}_i\,\partial_i\big]\partial^{-2} + \partial^{-2}\big[\mathbf{B}_i\,\partial_i\big]\partial^{-2}\big[\mathbf{B}_j\,\partial_j\big]\partial^{-2} + \cdots,$$

and substitute into (16.57):

$$\begin{aligned}\mathbf{\Delta}^{-1}(B) &= -\Box^{-1}(B)\partial^2\,\Box^{-1}(B)\\ &\simeq -\partial^{-2} + 2\,\partial^{-2}\big[\mathbf{B}_i\,\partial_i\big]\partial^{-2} - 3\,\partial^{-2}\big[\mathbf{B}_i\,\partial_i\big]\partial^{-2}\big[\mathbf{B}_j\,\partial_j\big]\partial^{-2}.\end{aligned}$$

The term linear in B disappears upon averaging, and we are left with the equal-time vacuum average of two transverse fields:



$$\langle 0|B_iB_j|0\rangle = \frac{1}{2|\mathbf{k}'|}\left(\delta_{ij} - \frac{k'_ik'_j}{\mathbf{k}'^2}\right).$$

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press

In the momentum space, the Coulomb energy takes the form

$$V^c_{\rm Coul} = \frac{g^2}{\mathbf{k}^2}\left\{1 + 3\cdot\int\frac{d^3\mathbf{k}'}{(2\pi)^3}\frac{N\,g^2}{2\,|\mathbf{k}'|\,(\mathbf{k}'-\mathbf{k})^2}\left(1 - \frac{(\mathbf{k}\cdot\mathbf{k}')^2}{\mathbf{k}^2\,\mathbf{k}'^2}\right)\right\}.$$

The angular integration produces

$$\int\frac{d\varphi\,d\cos\theta}{(2\pi)^3}\big(1-\cos^2\theta\big) = \frac{1}{(2\pi)^2}\left(2-\frac{2}{3}\right),$$

and from the large momentum region we get a logarithmically divergent correction

$$V^c_{\rm Coul} = \frac{g^2}{\mathbf{k}^2}\left(1 + 4C_2\cdot\frac{g^2}{4\pi^2}\ln\frac{\Lambda^2_{\rm UV}}{\mathbf{k}^2}\right). \tag{16.59}$$

Combining with the quantum contribution (16.58), we finally obtain the first correction to the vacuum energy,

$$\frac{g^2}{\mathbf{k}^2} \implies \frac{g^2}{\mathbf{k}^2}\left\{1 + \left(\left[4-\frac{1}{3}\right]C_2 - \frac{2}{3}n_f\right)\cdot\frac{g^2}{4\pi^2}\ln\frac{\Lambda^2_{\rm UV}}{\mathbf{k}^2}\right\},$$

which corresponds to the invariant coupling of Yang–Mills fields,

$$g^2(k^2) = \frac{g^2}{1-\beta_0\frac{g^2}{4\pi^2}\ln\frac{\Lambda^2_{\rm UV}}{k^2}} = \frac{4\pi^2}{\beta_0\ln\frac{k^2}{\Lambda^2}}, \qquad \beta_0 = \frac{11}{3}N - \frac{2}{3}n_f, \tag{16.60}$$

which *decreases* with the increase of k^2.

We see that the *anti-screening* (asymptotic freedom) is entirely due to vacuum fluctuations of the gluon fields affecting the Coulomb interaction. This effect is also likely to have other serious consequences related to the infrared instability of the quantum theory of Yang–Mills fields (Gribov, 1978).

https://doi.org/10.1017/9781009290227.017 Published online by Cambridge University Press