Preface

Quantum Chromodynamics or QCD was developed and defined over a brief period from 1972-73. One of us (EK) wrote an article early in 2021 on the scalar glueball and searched the literature to find where glueballs were first mentioned. This was at the 16th International Conference on High-Energy Physics (ICHEP 72). In the winter of 2021/2022 he thought it was time to prepare a volume dedicated to 50 Years of QCD. He got approval from the EJPC, and asked FG to join the effort. Here is the result.

It's been quite an adventure to guide and prepare this volume. From the start it was to be published as a single article, organized and edited by the two coeditors, with integrated contributions from invited scientists familiar with all aspects of the subject. Our initial outline included only eight sections, but as we got advice from our conveners and early contributors, the number of sections grew to the 14 you see here, and in some cases the number of subsections in each section also grew. The subject is both beautiful and vast, and keeping this volume "limited" in length was a real challenge.

Our goal was to prepare a volume for young Ph.Ds and postdocs that could serve as a readable resource and introduction to specialties outside of their own field of research – a shortcut to acquiring the broad familiarity that usually takes time to acquire. We also invited our contributors to reflect on how they developed their ideas/insights, usually discouraged in scientific articles. We believe that what has resulted is truly unique.

The volume begins with the personal reflections of two scientists who were contributors to the foundations of QCD (Sec. 1), and follows with three early developments that quickly showed that QCD as on the right track (Sec. 2). Prominent among these was the "November revolution," where the discovery and explanation of the states of charmonium lead quickly to the Cornell potential and an early description of why quarks could not be seen, convincing may doubters that quarks were real.

After establishing that the QCD fine structure constant, α_s, is too large at hadronic scales for perturbation theory to work (Sec. 3), we describe in some detail Lattice QCD (Sec. 4), believed now to be the only method that can give *exact* predictions for QCD (with numerical errors, of course, which are decreasing rapidly as the computations and computers improve). Unfortunately, Lattice QCD does not give much of an intuitive picture of how the physics works, so approximate analytic methods are needed (and will probably always be needed) and these are summarized in Secs. 5 and 6, including effective field theories, a powerful tool with many applications. Perhaps some day we will have exact analytic solutions, but not today.

From there our account turns to experimental manifestations of QCD (with theoretical support), starting with the exploration of the QCD phase diagram in heavy ion collisions and in dense matter (Sec. 7), followed by the study of mesons (Sec. 8) and baryons (Sec. 9) that reveal the existence of "exotic" states like glueballs, hybrids, hadronic molecules, and tetra- and pentaquarks. A special focus is given to the nucleon and its structure (Sec. 10). Then, collisions at high energies are discussed, from the hard scattering of two partons followed by their hadronization (Sec. 11); the production and identification of jets of particles culminated in the discovery of the Higgs boson and measurements of its properties (Sec. 12); weak decays, precision analyses of the quark mixing matrix, and the anomalous magnetic moment of muons that show the first hints of New Physics beyond the Standard Model (Sec. 13). The volume concludes with a brief account of experimental projects under construction or already funded (Sec. 14). We do not discuss the many exciting theoretical or experimental ideas that are currently in the drawing board, or as theorists sometimes say, on the "second sheet" (when they are joking about wonderful ideas still in an imaginative state). These we save for the next volume!

It has been a great experience for us to work on this volume; we hope you will find some pleasure in skimming through it.

1 Theoretical Foundations

Conveners:
Franz Gross and Eberhard Klempt

This section contains two personal accounts of the early development or "discovery" of QCD.

Leutwyler's contribution starts with a broad picture of the chaotic state of "theories" of the strong interactions in the 1960's and carries us through to the present day. He describes how many thought field theory could not work for the description of "nuclear forces." They thought the use of dispersion relations and unitarity would provide a better approach, but now we know that these are only useful tools. His discussion of how the exact and approximate symmetries of QCD lead to an understanding of the mass scales of the quarks shows how much the development of QCD and the standard model have brought order out of chaos, and have led to a deep understanding of the physics.

The second contribution by Fritzsch gives a more focused and personal account of how some issues that had to be surmounted before QCD became the accepted theory of the strong forces. He describes several arguments that them to the necessity for three colors of quarks (and the SU(3) color symmetry). He reminds us that QCD and the existence of quarks did not become widely accepted until the discovery of the J/ψ, among the topics discussed in the following Sec. 2.

Both of these accounts of the history and the physics are exciting to read, and a broad introduction to this volume. We hope you will enjoy them as much as we have.

1.1 The strong interactions[1]

Heinrich Leutwyler

1.1.1 Beginnings

The discovery of the neutron in 1932 [2] may be viewed as the birth of the strong interaction: it indicated that the nuclei consist of protons and neutrons and hence the presence of a force that holds them together, strong enough to counteract the electromagnetic repulsion. Immediately thereafter, Heisenberg introduced the notion of isospin as a symmetry of the strong interaction, in order to explain why proton and neutron nearly have the same mass [3]. In 1935, Yukawa pointed out that the nuclear force could be generated by the exchange of a hypothetical spinless particle, provided its mass is intermediate between the masses of proton and electron – a *meson* [4]. Today, we know that such a particle indeed exists: Yukawa predicted the pion. Stueckelberg pursued similar ideas, but was mainly thinking about particles of spin 1, in analogy with the particle that mediates the electromagnetic interaction [5].

In the thirties and fourties of the last century, the understanding of the force between two nucleons made considerable progress, in the framework of nonrelativistic potential models. These are much more flexible than quantum field theories. Suitable potentials that are attractive at large distances but repulsive at short distances do yield a decent understanding of nuclear structure: Paris potential, Bonn potential, shell model of the nucleus. In this framework, nuclear reactions, in particular the processes responsible for the luminosity of the sun, stellar structure, α-decay and related matters were well understood more than sixty years ago.

These phenomena concern interactions among nucleons with small relative velocities. Experimentally, it had become possible to explore relativistic collisions, but a description in terms of nonrelativistic potentials cannot cover these. In the period between 1935 and 1965, many attempts at formulating a theory of the strong interaction based on elementary fields for baryons and mesons were made. In particular, uncountable PhD theses were written, based on local interactions of the Yukawa type, using perturbation theory to analyze them. The coupling constants invariably turned out to be numerically large, indicating that the neglect of the higher order contributions was not justified. Absolutely nothing worked even half way.

Although there was considerable progress in understanding the general principles of quantum field theory (Lorentz invariance, unitarity, crossing symmetry, causality, analyticity, dispersion relations, CPT theorem, spin and statistics) as well as in renormalization theory, faith in quantum field theory was in decline, even concerning QED (Landau pole). To many, the renormalization procedure – needed to arrive at physically meaningful results – looked suspicious, and it appeared doubtful that the strong interaction could at all be described by means of a local quantum field theory. Some suggested that this framework should be replaced by S-matrix theory – heated debates concerning this suggestion took place at the time [6]. Regge poles were considered as a promising alternative to the quantum fields (the Veneziano model is born in 1968 [7]). Sixty years ago, when I completed my studies, the quantum field theory of the strong interaction consisted of a collection of beliefs, prejudices and assumptions. Quite a few of these turned out to be wrong.

1.1.2 Flavor symmetries

Symmetries that extend isospin to a larger Lie group provided the first hints towards an understanding of the structure underneath the strong interaction phenomena. The introduction of the strangeness quantum number and the Gell-Mann-Nishijima formula [8, 9] was a significant step in this direction. Goldberger and Treiman [10] then showed that the *axial vector current* plays an important role, not only in the weak interaction (the pion-to-vacuum matrix element of this current – the pion decay constant F_π – determines the rate of the weak decay $\pi \to \mu\nu$) but also in the context of the strong interaction: the nucleon matrix element of the axial vector current, g_A, determines the strength of the interaction between pions and nucleons:

$$g_{\pi N} = g_A M_N / F_\pi \,.$$

At low energies, the main characteristic of the strong interaction is that the energy gap is small: the lightest state occurring in the eigenvalue spectrum of the

[1] The present section is an extended version of my lecture notes *On the history of the strong interaction* [1]

Hamiltonian is the pion, with[2] $M_\pi \simeq 135$ MeV, small compared to the mass of the proton, $M_p \simeq 938$ MeV. In 1960, Nambu found out why that is so: it has to do with a hidden, approximate, continuous symmetry [11]. Since some of its generators carry negative parity, it is referred to as a *chiral symmetry*. For this symmetry to be consistent with observation, it is essential that an analog of spontaneous magnetization occurs in particle physics: for dynamical reasons, the state of lowest energy – the vacuum – is not symmetric under chiral transformations. Consequently, the symmetry cannot be seen in the spectrum of the theory: it is *hidden* or *spontaneously broken*. Nambu realized that the spontaneous breakdown of a continuous symmetry entails massless particles analogous to the spin waves of a magnet and concluded that the pions must play this role. If the strong interaction was strictly invariant under chiral symmetry, there would be no energy gap at all – the pions would be massless.[3] Conversely, since the pions are not massless, chiral symmetry cannot be exact – unlike isospin, which at that time was taken to be an exact symmetry of the strong interaction. The spectrum does have an energy gap because chiral symmetry is not exact: the pions are not massless, only light. In fact, they represent the lightest strongly interacting particles that can be exchanged between two nucleons. This is why, at large distances, the potential between two nucleons is correctly described by the Yukawa formula.

The discovery of the *Eightfold Way* by Gell-Mann and Ne'eman paved the way to an understanding of the mass pattern of the baryons and mesons [13, 14]. Like chiral symmetry, the group SU(3) that underlies the Eightfold Way represents an approximate symmetry: the spectrum of the mesons and baryons does not consist of degenerate multiplets of this group. The splitting between the energy levels, however, does exhibit a pattern that can be understood in terms of the assumption that the part of the Hamiltonian that breaks the symmetry transforms in a simple way. This led to the Gell-Mann-Okubo formula [14, 15] and to a prediction for the mass of the Ω^-, a member of the baryon decuplet which was still missing, but was soon confirmed experimentally, at the predicted place [16].

1.1.3 Quark Model

In 1964, Gell-Mann [17] and Zweig [18] pointed out that the observed pattern of baryons can qualitatively be understood on the basis of the assumption that these particles are bound states built with three constituents, while the spectrum of the mesons indicates that they contain only two of these. Zweig called the constituents "aces". Gell-Mann coined the term "quarks", which is now commonly accepted. The Quark Model gradually evolved into a very simple and successful semi-quantitative framework, but gave rise to a fundamental puzzle: why do the constituents not show up in experiment? For this reason, the existence of the quarks was considered doubtful: "Such particles [quarks] presumably are not real but we may use them in our field theory anyway ..." [19]. Quarks were treated like the veal used to prepare a pheasant in the royal french cuisine: the pheasant was baked between two slices of veal, which were then discarded (or left for the less royal members of the court). Conceptually, this was a shaky cuisine.

If the flavor symmetries are important, why are they not exact? Gell-Mann found a beautiful explanation: *current algebra* [14, 19]. The charges form an exact algebra even if they do not commute with the Hamiltonian and the framework can be extended to the corresponding currents, irrespective of whether or not they are conserved. Adler and Weisberger showed that current algebra can be tested with the sum rule that follows from the nucleon matrix element of the commutator of two axial vector charges [20, 21]. Weinberg then demonstrated that even the strength of the interaction among the pions can be understood on the basis of current algebra: the $\pi\pi$ scattering lengths can be predicted in terms of the pion decay constant [22].

1.1.4 Behavior at short distances

Bjorken had pointed out that if the nucleons contain point-like constituents, then the *ep* cross section should obey scaling laws in the deep inelastic region [23]. Indeed, the scattering experiments carried out by the MIT-SLAC collaboration in 1968/69 did show experimental evidence for such constituents [24]. Feynman called these *partons*, leaving it open whether they were the quarks or something else.

The operator product expansion turned out to be a very useful tool for the short distance analysis of the theory – the title of the paper where it was introduced [25], "Nonlagrangian models of current algebra," reflects the general skepticism towards Lagrangian quantum field theory that I mentioned in Section 1.1.1.

1.1.5 Color

The Quark Model was difficult to reconcile with the spin-statistics theorem which implies that particles of spin $\frac{1}{2}$ must obey Fermi statistics. Greenberg proposed

[2] I am using natural units where $\hbar = c = 1$.

[3] A precise formulation of this statement, known as the *Goldstone theorem*, was given later [12].

that the quarks obey neither Fermi-statistics nor Bose-statistics, but "para-statistics of order three" [26]. The proposal amounts to the introduction of a new internal quantum number. Indeed, Bogolyubov, Struminsky and Tavkhelidze [27], Han and Nambu [28] and Miyamoto [29] independently pointed out that some of the problems encountered in the quark model disappear if the u, d and s quarks occur in 3 states. Gell-Mann coined the term "color" for the new quantum number.

One of the possibilities considered for the interaction that binds the quarks together was an abelian gauge field analogous to the e.m. field, but this gave rise to problems, because the field would then interfere with the other degrees of freedom. Fritzsch and Gell-Mann pointed out that if the gluons carry color, then the empirical observation that quarks appear to be confined might also apply to them: the spectrum of the theory might exclusively contain color neutral states [30].

In his lectures at the Schladming Winter School in 1972 [31], Gell-Mann thoroughly discussed the role of the quarks and gluons: theorists had to navigate between Scylla and Charybdis, trying to abstract neither too much nor too little from models built with these objects. The basic tool at that time was *current algebra on the light cone*. He invited me to visit Caltech. I did that during three months in the spring break of 1973 and spent an extremely interesting period there. The personal recollections of Harald Fritzsch (see Section 1.2) describe the developments that finally led to *Quantum Chromodynamics*.

As it was known already that the electromagnetic and weak interactions are mediated by gauge fields, the idea that color might be a local symmetry as well does not appear as far fetched. The main problem at the time was that for a gauge field theory to describe the hadrons and their interaction, it had to be fundamentally different from the quantum field theories encountered in nature so far: all of these, including the electroweak theory, have the spectrum indicated by the degrees of freedom occurring in the Lagrangian: photons, leptons, intermediate bosons, ... The proposal can only make sense if this need not be so, that is if the spectrum of physical states in a quantum field theory can differ from the spectrum of the fields needed to formulate it: gluons and quarks in the Lagrangian, hadrons in the spectrum. This looked like wishful thinking. How come that color is confined while electric charge is free?

1.1.6 Electromagnetic interaction

The final form of the laws obeyed by the electromagnetic field was found by Maxwell, around 1860 – these laws survived relativity and quantum theory, unharmed. Fock pointed out that the Schrödinger equation for electrons in an electromagnetic field,

$$\frac{1}{i}\,\frac{\partial\psi}{\partial t}-\frac{1}{2m_e^2}\,(\vec{\nabla}+i\,e\vec{A})^2\psi-e\,\varphi\,\psi=0\,, \tag{1.1.1}$$

is invariant under a group of local transformations:

$$\vec{A}\,'(x)=\vec{A}(x)+\vec{\nabla}\alpha(x),\qquad \varphi\,'(x)=\varphi(x)-\frac{\partial\alpha(x)}{\partial t}$$
$$\psi(x)'=e^{-ie\alpha(x)}\,\psi(x)\,, \tag{1.1.2}$$

in the sense that the fields $\vec{A}',\varphi',\psi'$ describe the same physical situation as $\vec{A},\varphi,\psi$ [32]. Weyl termed these *gauge transformations* (with gauge group U(1) in this case). In fact, the electromagnetic interaction is fully characterized by symmetry with respect to this group: gauge invariance is the crucial property of this interaction.

I illustrate the statement with the core of Quantum Electrodynamics: photons and electrons. Gauge invariance allows only two free parameters in the Lagrangian of this system: e, m_e. Moreover, only one of these is dimensionless: $e^2/4\pi=1/137.035\,999\,084\,(21)$. U(1) symmetry and renormalizability fully determine the properties of the e.m. interaction, except for this number, which so far still remains unexplained.

1.1.7 Nonabelian gauge fields

Kaluza [33] and Klein [34] had shown that a 5-dimensional Riemann space with a metric that is independent of the fifth coordinate is equivalent to a 4-dimensional world with *gravity*, a *gauge field* and a *scalar field*. In this framework, gauge transformations amount to a shift in the fifth direction: $x^{5'}=x^5+\alpha(\vec{x},t)$. In geometric terms, a metric space of this type is characterized by a group of isometries: the geometry remains the same along certain directions, indicated by Killing vectors. In the case of the 5-dimensional spaces considered by Kaluza and Klein, the isometry group is the abelian group U(1). The fifth dimension can be compactified to a circle – U(1) then generates motions on this circle. A particularly attractive feature of this theory is that it can explain the quantization of the electric charge: fields living on such a manifold necessarily carry integer multiples of a basic charge unit.

Pauli noticed that the Kaluza-Klein scenario admits a natural generalization to higher dimensions, where larger isometry groups find place. Riemann spaces of dimension >5 admit nonabelian isometry groups that reduce the system to a 4-dimensional one with *gravity*, *nonabelian gauge fields* and several *scalar fields*. Pauli was motivated by the isospin symmetry of the meson-nucleon interaction and focused attention on a Riemann space of dimension 6, with isometry group SU(2).

Pauli did not publish the idea that the strong interaction might arise in this way, because he was convinced that the quanta of a gauge field are massless: gauge invariance does not allow one to put a mass term into the Lagrangian. He concluded that the forces mediated by gauge fields are necessarily of long range and can therefore not mediate the strong interaction, which is known to be of short range. More details concerning Pauli's thoughts can be found in [35]. The paper of Yang and Mills appeared in 1954 [36]. Ronald Shaw, a student of Salam, independently formulated nonabelian gauge field theory in his PhD thesis [37]. Ten years later, Higgs [38], Brout and Englert [39] and Guralnik, Hagen and Kibble [40] showed that Pauli's objection is not valid in general: in the presence of scalar fields, gauge fields can pick up mass, so that forces mediated by gauge fields can be of short range. The work of Glashow [41], Weinberg [42] and Salam [43] then demonstrated that nonabelian gauge fields are relevant for physics: the framework discovered by Higgs et al. does accommodate a realistic description of the e.m. and weak interactions.

1.1.8 Asymptotic freedom

Already in 1965, Vanyashin and Terentyev [44] found that the renormalization of the electric charge of a vector field is of opposite sign to the one of the electron. In the language of SU(2) gauge field theory, their result implies that the β-function is negative at one loop.

The first correct calculation of the β-function of a nonabelian gauge field theory was carried out by Khriplovich, for the case of SU(2), relevant for the electroweak interaction [45]. He found that β is negative and concluded that the interaction becomes weak at short distance. In his PhD thesis, 't Hooft performed the calculation of the β-function for an arbitrary gauge group, including the interaction with fermions and Higgs scalars [46, 47]. He demonstrated that the theory is renormalizable and confirmed that, unless there are too many fermions or scalars, the β-function is negative at small coupling.

In 1973, Gross and Wilczek [48] and Politzer [49] discussed the consequences of a negative β-function and suggested that this might explain Bjorken scaling, which had been observed at SLAC in 1969. They pointed out that QCD predicts specific modifications of the scaling laws. In the meantime, there is strong experimental evidence for these.

1.1.9 Arguments in favor of QCD

The reasons for proposing QCD as a theory of the strong interaction are discussed in [50]. The idea that the observed spectrum of particles can fully be understood on the basis of a theory built with quarks and gluons still looked rather questionable and was accordingly formulated in cautious terms. In the abstract, for instance, we pointed out that "...there are several advantages in abstracting properties of hadrons and their currents from a Yang-Mills gauge model based on colored quarks and color octet gluons." Before the paper was completed, the papers by Gross, Wilczek and Politzer quoted above circulated as preprints - they are quoted and asymptotic freedom is given as argument #4 in favor of QCD. Also, important open questions were pointed out, in particular, the U(1) problem.

Many considered QCD a wild speculation. On the other hand, several papers concerning gauge field theories that include the strong interaction appeared around the same time, for instance [51, 52].

1.1.10 November revolution

The discovery of the J/ψ was announced simultaneously at Brookhaven and SLAC, on November 11, 1974. Three days later, the observation was confirmed at ADONE, Frascati and ten days later, the ψ' was found at SLAC, where subsequently many further related states were discovered. We now know that these are bound states formed with the c-quark and its antiparticle which is comparatively heavy and that there are two further, even heavier quarks: b and t.

At sufficiently high energies, quarks and gluons do manifest themselves as jets. Like the neutrini, they have left their theoretical place of birth and can now be seen flying around like ordinary, observable particles. Gradually, particle physicists abandoned their outposts in no man's and no woman's land, returned to the quantum fields and resumed discussion in the good old *Gasthaus zu Lagrange*, a term coined by Jost. The theoretical framework that describes the strong, electromagnetic and weak interactions in terms of gauge fields, leptons, quarks and scalar fields is now referred to as the Standard Model - this framework clarified the picture enormously.[4]

[4] Indeed, the success of this theory is amazing: Gauge fields are renormalizable in four dimensions, but it looks unlikely that the Standard Model is valid much beyond the explored energy range. Presumably it represents an effective theory. There is no reason, however, for an effective theory to be renormalizable. One of the most puzzling aspects of the Standard Model is that it is able to account for such a broad range of phenomena that are characterized by very different scales within one and the same renormalizable theory.

1.1.11 Quantum chromodynamics

If the electroweak gauge fields as well as the leptons and the scalars are dropped, the Lagrangian of the Standard Model reduces to QCD:

$$\mathcal{L}_{QCD} = -\tfrac{1}{4}F^A_{\mu\nu}F^{A\,\mu\nu} + i\bar{q}\gamma^\mu(\partial_\mu + ig_s\tfrac{1}{2}\lambda^A\mathcal{A}^A_\mu)q \\ -\bar{q}_{\mathrm{R}}\mathcal{M}q_{\mathrm{L}} - \bar{q}_{\mathrm{L}}\mathcal{M}^\dagger q_{\mathrm{R}} - \theta\,\omega\,. \tag{1.1.3}$$

The gluons are described by the gauge field $\mathcal{A}^A_\mu$, which belongs to the color group $\mathrm{SU_c}(3)$ and g_s is the corresponding coupling constant. The field strength tensor $F^A_{\mu\nu}$ is defined by

$$F^A_{\mu\nu} = \partial_\mu\mathcal{A}^A_\nu - \partial_\nu\mathcal{A}^A_\mu - g_s f_{ABC}\mathcal{A}^B\mathcal{A}^C\,, \tag{1.1.4}$$

where the symbol f_{ABC} denotes the structure constants of SU(3). The quarks transform according to the fundamental representation of $\mathrm{SU_c}(3)$. The compact notation used in (1.1.3) suppresses the labels for flavor, colour and spin: the various quark flavors are represented by Dirac fields, $q = \{u,\, d,\, s,\, c,\, b,\, t\}$ and $q_{\mathrm{R}} = \frac{1}{2}(1+\gamma_5)q$, $q_{\mathrm{L}} = \frac{1}{2}(1-\gamma_5)q$ are their right- and left-handed components. The field u(x), for instance, contains 3×4 components. While the 3×3 Gell-Mann matrices λ^A act on the color label and satisfy the commutation relation

$$[\lambda^A, \lambda^B] = 2if_{ABC}\lambda^C\,, \tag{1.1.5}$$

the Dirac matrices γ^μ operate on the spin index. The mass matrix $\mathcal{M}$, on the other hand, acts in flavor space. Its form depends on the choice of the quark field basis. If the right- and left-handed fields are subject to independent rotations, $q_{\mathrm{R}} \to V_{\mathrm{R}}\,q_{\mathrm{R}}$, $q_{\mathrm{L}} \to V_{\mathrm{L}}\,q_{\mathrm{L}}$, where $V_{\mathrm{R}}, V_{\mathrm{L}} \in \mathrm{SU}(N_f)$ represent $N_f \times N_f$ matrices acting on the quark flavour, the quark mass matrix is replaced by $\mathcal{M} \to V_{\mathrm{R}}^\dagger\mathcal{M}V_{\mathrm{L}}$. This freedom can be used to not only diagonalize $\mathcal{M}$, but to ensure that the eigenvalues are real, nonnegative and ordered according to $0 \le m_u \le m_d \le \ldots \le m_t$.

The constant θ is referred to as the vacuum angle and ω stands for the winding number density

$$\omega = \frac{g_s^2}{32\pi^2}F^A_{\mu\nu}\tilde{F}^{A\,\mu\nu}\,, \tag{1.1.6}$$

where $\tilde{F}^{A\,\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F^A_{\rho\sigma}$ is the dual of the field strength.

As it is the case with electrodynamics, gauge invariance fully determines the form of the chromodynamic interaction. The main difference between QED and QCD arises from the fact that the corresponding gauge groups, U(1) and SU(3), are different. While the structure constants of U(1) vanish because this is an abelian group, those of $\mathrm{SU_c}(3)$ are different from zero. For this reason, gauge invariance implies that the Lagrangian contains terms involving three or four gluon fields: in contrast to the photons, which interact among themselves only via the exchange of charged particles, the gluons would interact even if quarks did not exist.

The term involving ω can be represented as a derivative, $\omega = \partial_\mu f^\mu$. Since only the integral over the Lagrangian counts, this term represents a contribution that only depends on the behavior of the gauge field at the boundary of space-time. In the case of QED, where renormalizability allows the presence of an analogous term, quantities of physical interest are indeed unaffected by such a contribution, but for QCD, this is not the case. Even at the classical level, nonabelian gauge fields can form instantons, which minimize the Euclidean action for a given nonzero winding number $\nu = \int d^4x\,\omega$.

1.1.12 Theoretical paradise

In order to briefly discuss some of the basic properties of QCD, let me turn off the electroweak interaction, treat the three light quarks as massless and the remaining ones as infinitely heavy:

$$m_u = m_d = m_s = 0\,, \qquad m_c = m_b = m_t = \infty\,. \tag{1.1.7}$$

The Lagrangian then contains a single parameter: the coupling constant g_s, which may be viewed as the net color of a quark. Unlike an electron, a quark cannot be isolated from the rest of the world – its color g_s depends on the radius of the region considered. According to perturbation theory, the color contained in a sphere of radius r grows logarithmically with the radius[5]:

$$\alpha_s \equiv \frac{g_s^2}{4\pi} = \frac{2\pi}{9\,|\ln(r\,\Lambda)|}\,. \tag{1.1.8}$$

Although the classical Lagrangian of massless QCD does not contain any dimensionful parameter, the corresponding quantum field theory does: the strength of the interaction cannot be characterized by a number, but by a dimensionful quantity, the intrinsic scale Λ.

The phenomenon is referred to as *dimensional transmutation.* In perturbation theory, it manifests itself through the occurrence of divergences – contrary to what many quantum field theorists thought for many years, the divergences do not represent a disease, but are intimately connected with the structure of the theory. They are a consequence of the fact that a quantum field theory does not inherit all of the properties of the corresponding classical field theory. In the case of massless Chromodynamics, the classical Lagrangian does not contain any dimensionful constants and hence

[5] The formula only holds if the radius is small, $r\,\Lambda \ll 1$.

remains invariant under a change of scale. This property, which is referred to as conformal invariance, does not survive quantization, however. Indeed, it is crucial for Quantum Chromodynamics to be consistent with what is known about the strong interaction that this theory does have an intrinsic scale.

Massless QCD is how theories should be: the Lagrangian does not contain a single dimensionless parameter. In principle, the values of all quantities of physical interest are predicted without the need to tune parameters (the numerical value of the mass of the proton in kilogram units cannot be calculated, of course, because that number depends on what is meant by a kilogram, but the mass spectrum, the width of the resonances, the cross sections, the form factors, ... can be calculated in a parameter free manner from the mass of the proton, at least in principle).

1.1.13 Symmetries of massless QCD

The couplings of the u-, d- and s-quarks to the gauge field are identical. In the *chiral limit*, where the masses are set equal to zero, there is no difference at all – the Lagrangian is symmetric under SU(3) rotations in flavor space. Indeed, there is more symmetry: for massless fermions, the right- and left-handed components can be subject to independent flavor rotations. The Lagrangian of QCD with three massless flavors is invariant under $\mathrm{SU}(3)_R{\times}\mathrm{SU}(3)_L$. QCD thus explains the presence of the mysterious chiral symmetry discovered by Nambu: an exact symmetry of this type is present if some of the quarks are massless.

Nambu had conjectured that chiral symmetry breaks down spontaneously. Can it be demonstrated that the symmetry group $\mathrm{SU}(3)_R{\times}\mathrm{SU}(3)_L$ of the Lagrangian of massless QCD spontaneously break down to the subgroup $\mathrm{SU}(3)_{R+L}$? To my knowledge an analytic proof is not available, but the work done on the lattice demonstrates beyond any doubt that this does happen. In particular, for $m_u = m_d = m_s$, the states do form degenerate multiplets of $\mathrm{SU}(3)_{R+L}$ and, in the limit $m_u, m_d, m_s \to 0$, the pseudoscalar octet does become massless, as required by the Goldstone theorem.

1.1.14 Quark masses

The 8 lightest mesons, $\pi^+, \pi^0, \pi^-, K^+, K^0, \bar{K}^0, K^-, \eta$, do have the quantum numbers of the Nambu-Goldstone bosons, but massless they are not. The reason is that we are not living in the paradise described above: the light quark masses are different from zero. Accordingly, the Lagrangian of QCD is only approximately invariant under chiral rotations, to the extent that the symmetry breaking parameters m_u , m_d , m_s are small. Since they differ, the multiplets split. In particular, the Nambu-Goldstone bosons pick up mass.

Even before the discovery of QCD, attempts at estimating the masses of the quarks were made. In particular, nonrelativistic bound state models for mesons and baryons where constructed. In these models, the proton mass is dominated by the sum of the masses of its constituents: $m_u + m_u + m_d \simeq m_p$, $m_u \simeq m_d \simeq 300\,\mathrm{MeV}$.

With the discovery of QCD, the mass of the quarks became an unambiguous concept: the quark masses occur in the Lagrangian of the theory. Treating the mass term as a perturbation, one finds that the expansion of $m^2_{\pi^+}$ in powers of m_u, m_d, m_s starts with $m^2_{\pi^+} = (m_u + m_d)B_0 + \ldots$ The constant B_0 also determines the first term in the expansion of the square of the kaon masses: $m^2_{K^+} = (m_u + m_s)B_0 + \ldots$, $m^2_{K^0} = (m_d + m_s)B_0 + \ldots$ Since the kaons are significantly heavier than the pions, these relations imply that m_s must be large compared to m_u, m_d.

The first crude estimate of the quark masses within QCD relied on a model for the wave functions of π, K, ρ, which was based on SU(6) (spin-flavor-symmetry) and led to $B_0 \simeq \frac{3}{2} m_\rho F_\rho / F_\pi$. Numerically, this yields $B_0 \simeq 1.8\,\mathrm{GeV}$. For the mean mass of the two lightest quarks, $m_{ud} \equiv \frac{1}{2}(m_u{+}m_d)$, this estimate implies $m_{ud} \simeq 5\,\mathrm{MeV}$, while the mass of the strange quark becomes $m_s \simeq 135\,\mathrm{MeV}$ [53]. Similar mass patterns were found earlier, within the Nambu-Jona-Lasinio model [54] or on the basis of sum rules [55].

1.1.15 Breaking of isospin symmetry

From the time Heisenberg had introduced isospin symmetry, it was taken for granted that the strong interaction strictly conserves isospin. QCD does have this symmetry if and only if $m_u = m_d$. If that condition were met, the mass difference between proton and neutron would be due exclusively to the e.m. interaction. This immediately gives rise to a qualitative problem: why is the charged particle, the proton, lighter than its neutral partner, the neutron?

The Cottingham formula [56] states that the leading contribution of the e.m. interaction to the mass of a particle is determined by the cross section for electron scattering on this particle. We evaluated the formula on the basis of Bjorken scaling and of the experimental data for electron scattering on protons and neutrons available at the time. Since we found that the electromagnetic self energy of the proton is larger than the one of the neutron, we concluded that the strong interaction does not conserve isospin: even if the e.m. interaction is turned off, m_u must be different from m_d. In fact, the first crude estimate for the masses of the light quarks [57],

$$m_u \simeq 4\,\mathrm{MeV}, \quad m_d \simeq 7\,\mathrm{MeV}, \quad m_s \simeq 135\,\mathrm{MeV}\,, \tag{1.1.9}$$

indicated that m_d must be almost twice as large as m_u.

It took quite a while before this bizarre pattern was generally accepted. The Dashen theorem [58] states that, in a world where the quarks are massless, the e.m. self energies of the kaons and pions obey the relation $m^{2\,em}_{K^+} - m^{2\,em}_{K^0} = m^{2\,em}_{\pi^+} - m^{2\,em}_{\pi^0}$. If the mass differences were dominated by the e.m. interaction, the charged kaon would be heavier than the neutral one. Hence the mass difference between the kaons cannot be due to the electromagnetic interaction, either. The estimates for the quark mass ratios obtained with the Dashen theorem confirm the above pattern [59].

1.1.16 Approximate symmetries are natural in QCD

At first sight, the fact that m_u strongly differs from m_d is puzzling: if this is so, why is isospin such a good quantum number? The key observation here is the one discussed in Section 1.1.12: QCD has an intrinsic scale, Λ. For isospin to represent an approximate symmetry, it is not necessary that $m_d - m_u$ is small compared to $m_u + m_d$. It suffices that the symmetry breaking parameter is small compared to the intrinsic scale, $m_d - m_u \ll \Lambda$.

In the case of the eightfold way, the symmetry breaking parameters are the differences between the masses of the three light quarks. If they are small compared to the intrinsic scale of QCD, then the Green functions, masses, form factors, cross sections ... are approximately invariant under the group $\mathrm{SU}(3)_{R+L}$. Isospin is an even better symmetry, because the relevant symmetry breaking parameter is smaller, $m_d - m_u \ll m_s - m_u$. The fact that $m^2_{\pi^+}$ is small compared to $m^2_{K^+}$ implies $m_u + m_d \ll m_u + m_s$. Hence all three light quark masses must be small compared to the scale of QCD.

In the framework of QCD, the presence of an approximate chiral symmetry group of the form $\mathrm{SU}(3)_\mathrm{R} \times \mathrm{SU}(3)_\mathrm{L}$ thus has a very simple explanation: it so happens that the masses of u, d and s are small. We do not know why, but there is no doubt that this is so. The quark masses represent a perturbation, which in first approximation can be neglected – in first approximation, the world is the paradise described above.

1.1.17 Ratios of quark masses

The confinement of color implies that the masses of the quarks cannot be identified by means of the four-momentum of a one-particle state – the spectrum of the theory does not contain such states. As parameters occurring in the Lagrangian, they need to be renormalized and the renormalized mass depends on the regularization used to set up the theory. In the $\overline{\mathrm{MS}}$ scheme [60–62], they depend on the running scale – only their ratios represent physical quantities. Among the three lightest quarks, there are two independent mass ratios, which it is convenient to identify with

$$S = \frac{m_s}{m_{ud}}\,, \qquad R = \frac{m_s - m_{ud}}{m_d - m_u}\,, \tag{1.1.10}$$

where $m_{ud} \equiv \frac{1}{2}(m_u + m_d)$.

Since the isospin breaking effects due to the e.m. interaction are not negligible, the physical masses of the Goldstone boson octet must be distinguished from their masses in QCD, i.e. in the absence of the electroweak interactions. I denote the latter by $\hat{m}_P$ and use the symbol $\hat{m}_K$ for the mean square kaon mass in QCD, $\hat{m}^2_K \equiv \frac{1}{2}(\hat{m}^2_{K^+} + \hat{m}^2_{K^0})$. The fact that the expansion of the square of the Goldstone boson masses in powers of m_u, m_d, m_s starts with a linear term implies that, in the chiral limit, their ratios are determined by R and S. In particular, the expansion of the ratios of $\hat{m}^2_{\pi^+}$, $\hat{m}^2_{K^+}$ and $\hat{m}^2_{K^0}$ starts with

$$\frac{2\hat{m}^2_K}{\hat{m}^2_{\pi^+}} = (S+1)\{1 + \Delta_S\}\,, \tag{1.1.11}$$

$$\frac{\hat{m}^2_K - \hat{m}^2_{\pi^+}}{\hat{m}^2_{K^0} - \hat{m}^2_{K^+}} = R\{1 + \Delta_R\}\,, \tag{1.1.12}$$

where Δ_S as well as Δ_R vanish in the chiral limit – they represent corrections of $O(\mathcal{M})$. The left hand sides only involve the masses of π^+, K^+ and K^0. Invariance of QCD under charge conjugation implies that the masses of π^-, K^- and $\bar{K}^0$ coincide with these. There are low energy theorems analogous to (1.1.11), (1.1.12), involving the remaining members of the octet, π^0 and η, but these are more complicated because the states $|\pi^0\rangle$ and $|\eta\rangle$ undergo mixing {at leading order, chiral symmetry implies that the mixing angle is given by $\tan(2\theta) = \sqrt{3}/2R$}. In the isospin limit, $\{m_u = m_d,\ e = 0\}$, the masses of π^0 and π^+ coincide and $\hat{m}_\eta$ obeys the Gell-Mann-Okubo formula, $(\hat{m}^2_\eta - \hat{m}^2_K)/(\hat{m}^2_K - \hat{m}^2_\pi) = \frac{1}{3}\{1 + O(\mathcal{M})\}$.

While the accuracy to which S can be determined on the lattice is amazing, the uncertainty in R is larger by almost an order of magnitude [63]:

$$S = 27.42(12)\,, \qquad R = 38.1(1.5)\,. \tag{1.1.13}$$

The reason is that R concerns isospin breaking effects. The contributions arising from QED are not negligible at this precision and since the e.m. interaction is of long range, it is more difficult to simulate on a lattice.

The difference shows up even more clearly in the corrections. The available lattice results [63] lead to $\Delta_S = 0.057(7)$, indicating that the low energy theorem (1.1.11) picks up remarkably small corrections from higher orders of the quark mass expansion. Those occurring in the Gell-Mann-Okubo formula are also known to be very small. The number $\Delta_R = -0.016(57)$ obtained from the available lattice results is also small, but the uncertainty is so large that even the sign of the correction remains open.

The quantities Δ_S, Δ_R exclusively concern QCD and could be determined to high precision with available methods, in the framework of $N_f = 1+1+1$: three flavours of different mass. For isospin breaking quantities, the available results come with a large error because they do not concern QCD alone but are obtained from a calculation of the physical masses, so that the e.m. interaction cannot be ignored. A precise calculation of $\hat{m}_{\pi^+}$, $\hat{m}_{K^+}$, $\hat{m}_{K^0}$ within lattice QCD would be of considerable interest as it would allow to subject a venerable low energy theorem for the quark mass ratio $Q^2 \equiv (m_s^2 - m_{ud}^2)/(m_d^2 - m_u^2)$ [64] to a stringent test. The theorem implies that the leading contributions to Δ_R and Δ_S are equal in magnitude, but opposite in sign: $\Delta_R = -\Delta_S + O(\mathcal{M}^2)$ [65]. The available numbers are consistent with this relation but far from accurate enough to allow a significant test. There is no doubt that the leading terms dominate if the quark masses are taken small enough, but since the estimates for Δ_R and Δ_S obtained at the physical values of the quark masses turn out to be unusually small, it is conceivable that the corrections of $O(\mathcal{M}^2)$ are of comparable magnitude. For $m_u = m_d$, the masses of the Goldstone bosons have been worked out to NNLO of Chiral Perturbation Theory [66]. An extension of these results to $\hat{m}_{\pi^+}$, $\hat{m}_{K^+}$, $\hat{m}_{K^0}$ for $m_u \neq m_d$ should be within reach and would allow a much more precise lattice determination of Δ_R.

1.1.18 U(1) anomaly, CP-problem

Even before the discovery of QCD, it was known that, in the presence of vector fields, the Ward identities for axial currents contain anomalies [67–69]. In particular, an external e.m. field generates an anomaly in the conservation law for the axial current $\bar{u}\gamma^\mu\gamma_5 u - \bar{d}\gamma^\mu\gamma_5 d$. The anomaly implies a low energy theorem for the decay $\pi^0 \to \gamma + \gamma$, which states that, to leading order in the expansion in powers of the momenta and for $m_u = m_d = 0$, the transition amplitude is determined by F_π, i.e. by the same quantity that determines the rate of the decay $\pi^+ \to \mu + \nu_\mu$.

In QCD, the conservation law for the singlet axial current contains an anomaly,

$$\partial_\mu(\bar{q}\gamma^\mu\gamma_5 q) = 2i\,\bar{q}\mathcal{M}\gamma_5 q + 2N_f\,\omega\,, \tag{1.1.14}$$

where N_f is the number of flavors and ω is specified in (1.1.6). The phenomenon plays a crucial role because it implies that even if the quark mass matrix $\mathcal{M}$ is set equal to zero, the singlet axial charge is not conserved. Hence the symmetry group of QCD with 3 massless flavors is $\mathrm{SU}(3)_{\mathrm{R}}\times\mathrm{SU}(3)_{\mathrm{L}}\times\mathrm{U}(1)_{\mathrm{R+L}}$, not $\mathrm{U}(3)_{\mathrm{R}}\times\mathrm{U}(3)_{\mathrm{L}}$. QCD is not invariant under the chiral transformations generated by the remaining factor, $\mathrm{U}(1)_{\mathrm{R\text{-}L}}$. This is why the paradise described above contains 8 rather than 9 massless Goldstone bosons.

The factor $\mathrm{U}(1)_{\mathrm{R\text{-}L}}$ changes the phase of the right-handed components of all quark fields by the same angle, $q'_{\mathrm{R}} = e^{i\beta}q_{\mathrm{R}}$, while the left-handed components are subject to the opposite transformation: $q'_{\mathrm{L}} = e^{-i\beta}q_{\mathrm{L}}$. This change of basis can be compensated by modifying the quark mass matrix with $\mathcal{M}' = e^{2i\beta}\mathcal{M}$, but in view of the anomaly, the operation does not represent a symmetry of the system. The relation (1.1.14) shows, however, that current conservation is not lost entirely – it only gets modified. In fact, if the above change of the quark mass matrix is accompanied by a simultaneous change of the vacuum angle, $\theta' = \theta - 2\beta$, the physics does remain the same. Note that, starting from an arbitrary mass matrix, a change of basis involving the factor $\mathrm{U}(1)_{\mathrm{R\text{-}L}}$ is needed to arrive at the convention where $\mathcal{M}$ is diagonal with real eigenvalues. In that convention, the vacuum angle does have physical significance – otherwise only the product $e^{i\theta}\mathcal{M}$ counts.

The Lagrangian of QCD is invariant under charge conjugation, but the term $-\theta\,\omega$ has negative parity. Accordingly, unless θ is very small, there is no explanation for the fact that CP-violating quantities such as the electric dipole moment of the neutron are too small to have shown up in experiment. This is referred to as the strong CP-problem.

There is a theoretical solution of this puzzle: if the lightest quark were massless, $m_u = 0$, QCD would conserve CP. The Dirac field of the u-quark can then be subject to the chiral transformation $u'_{\mathrm{R}} = e^{i\beta}u_{\mathrm{R}}$, $u'_{\mathrm{L}} = e^{-i\beta}u_{\mathrm{L}}$ without changing the quark mass matrix. As discussed above, the physics remains the same, provided the vacuum angle is modified accordingly. This shows that if one of the quarks were massless, the vacuum angle would become irrelevant. It would then be legitimate to set $\theta = 0$, so that the Lagrangian becomes manifestly CP-invariant.

This 'solution', however, is fake. If m_u were equal to zero, the ratio R would be related to S by $R = \frac{1}{2}(S-1)$. The very accurate value for S in equation (1.1.13) would imply $R = 13.21(6)$, more than 16 standard deviations away from the result quoted for R.

1.1.19 QCD as part of the Standard Model

In the Standard Model, the vacuum contains a condensate of Higgs bosons. At low energies, the manner in which the various other degrees of freedom interact with these plays the key role. Since they do not have color and are electrically neutral, their condensate is transparent for gluons and photons. The gauge bosons $W^\pm$, Z that mediate the weak interaction, as well as the leptons and quarks do interact with the condensate: photons and gluons remain massless, all other particles occurring in the Standard Model are hindered in moving through the condensate and hence pick up mass. In cold matter only the lightest degrees of freedom survive: photons, gluons, electrons, u- and d-quarks – all other particles are unstable, decay and manifest their presence only indirectly, through quantum fluctuations.

At low energies, the Standard Model boils down to a remarkably simple theory: QCD + QED. The Lagrangian only contains the coupling constants g_s, e, θ and the masses of the quarks and leptons as free parameters, but describes the laws of nature relevant at low energies to breathtaking precision. The gluons and the photons represent the gauge fields that belong to color and electric charge, respectively. Color is confined, but electric charge is not: while electrons can move around freely, quarks and gluons form color neutral bound states – mesons, baryons, nuclei.

The structure of the atoms is governed by QED because the e.m. interaction is of long range. In particular, their size is of the order of the Bohr radius, $a_B = 4\pi/e^2 m_e$, which only involves the mass of the electron and the coupling constant e. The mass of the atoms, on the other hand, is dominated by the energy of the gluons and quarks that are bound in the nucleus. It is of the order of the scale Λ_{QCD}, which characterizes the value of g_s in a renomalization group invariant manner. Evidently, the sum of the charges of the quarks contained in the nucleus also matters, as it determines the number of electrons that can be bound to it. The mass of the quarks, on the other hand, plays an important role only in so far as it makes the proton the lightest baryon – the world would look rather different if the neutron was lighter ...

The properties of the interaction among the quarks and gluons does not significantly affect the structure of the atoms, but from the theoretical point of view, the gauge field theory that describes it, QCD, is the most remarkable part of the Standard Model. In fact, it represents the first non-trivial quantum field theory that is internally consistent in four-space-time dimensions. In contrast to QED or to the Higgs sector, QCD is asymptotically free. The behavior of the quark and gluon fields at very short distances is under control. A cutoff is needed to set the theory up, but it can unambiguously be removed. In principle, all of the physical quantities of interest are determined by the renormalization group invariant quark mass matrix, by the vacuum angle θ and a scale. In the basis where the quark mass matrix is diagonal and real, the vacuum angle is tiny. We do not know why this is so, nor do we understand the bizarre pattern of eigenvalues.

1.2 The origins of QCD

Harald Fritzsch

Murray Gell-Mann and I started to collaborate in October 1970. We considered the results of the experiments on deep inelastic scattering at the Stanford Linear Accelerator Center. James Bjorken had predicted, using current algebra, that the cross sections showed at large values of the virtual photon mass and the energy transfer to the nucleon a scaling behavior, i.e. the cross section is a function of the ratio x, where x is the ratio of the square of the virtual photon mass to the energy transfer to the nucleon, multiplied with the nucleon mass. This ratio x varies from zero to one.

Since in the scaling region the cross sections were determined by the commutator of two electromagnetic currents at nearly lightlike distances, Gell-Mann and I assumed, that this commutator near the light cone is given by the free quark model. Thus the Bjorken scaling followed from this assumption.

The interaction between the quarks was assumed not to be present near the light cone. The cross section in the deep inelastic region determined the distribution functions of the three quarks and antiquarks, which are given by the proton matrix element of the commutator of the electromagnetic current.

In the free quark model the commutator near the light cone is given by a singular function, multiplied by a bilocal function of quark fields [70]. The matrix elements of these bilocal operators determined the quark distribution functions of the nucleon. The integral of the quark distribution functions gives the contribution of all the quark momenta to the nucleon momentum.

Gell-Mann and I expected that this integral would be +1, since inside the nucleon were only the three quarks and three antiquarks. However according to the experiments at SLAC this integral was only about 45%:

$$\int_0^1 x\left[u(x)+\overline{u}(x)+d(x)+\overline{d}(x)+s(x)+\overline{s}(x)\right]\mathrm{d}x \simeq 0.45. \quad (1.2.1)$$

Thus besides the quarks there must exist neutral quanta, which are relevant for the confinement of the quarks and which contribute about 55% to the momentum of a fast moving nucleon. This observation was the first indication that the strong interactions are described by a gauge theory. In such a theory there would be besides the quarks and antiquarks also neutral gluons.

Afterwards Gell-Mann and I considered several problems of the quark theory. The Ω^- particle was a bound state of three strange quarks. The three spin vectors of the quarks were symmetricalls arranged, and the space wave function was symmetric, since the Ω^- is the ground state of three strange quarks. Thus an interchange of two strange quarks was symmetric, but according to the Pauli principle it should be antisymmetric.

Another problem was related to the electromagnetic decay of the neutral pion. The decay rate, calculated in the quark model, is much smaller than the observed decay rate, only about 1/9 of the observed rate.

We also studied the cross section for the reaction electron-positron annihilation into hadrons. The ratio R of the cross section for hadron production and the cross section for the production of a muon pair can be calculated in the quark model. It is given by the sum of the squares of the electric charges of the three quarks, i.e. 2/3. But according to the experiments at the CEA accelerator at Harvard university this ratio was about three times larger: $R \simeq 2$.

To solve theses problems, Murray Gell-Mann, William Bardeen and I introduced for the quarks a new quantum number, which we called "color". Each quark is described by a red, a green and a blue quark. The three colors can be transformed by the color group $\mathrm{SU}(3)_{\mathrm{C}}$, which is assumed to be an exact symmetry. Measurable quantities, e.g. cross sections or the wave functions of hadrons, are color singlets.

The quark wave function ψ_Ω of the Ω^- is also a color singlet:

$$\psi_\Omega \simeq (rgb - grb + brg - rbg + gbr - bgr). \quad (1.2.2)$$

This wave function is antisymmetric under the exchange of two quarks - there is no problem with the Pauli principle. The quark wave functions of mesons are also color singlets:

$$\psi_{\text{meson}} \simeq (\bar{r}r + \bar{g}g + \bar{b}b). \quad (1.2.3)$$

The decay amplitude for the neutral pion decay is three times larger, if the quarks are colored. Thus the decay rate is nine times larger and agrees with the observed decay rate [71]. The ratio R for electron-positron annihilaton, given by the sum of the squares of the quark charges, is now also three times larger: $R \simeq 2$. Thus the introduction of the color quantum number solved the three problems mentioned above.

The color quantum number also explains why mesons are quark-antiquark bound states and baryons are three quark bound states, since they must be color singlets. Thus the mesons and baryons could be considered to be "white" states, since a particular color cannot be seen from the outside - the color quantum number is only relevant inside the mesons and baryons.

In the spring of 1972 Gell-Mann and I tried to understand why a colored quark cannot be observed - it is confined inside a baryon or meson or inside an atomic nucleus. We considered to use the color symmetry group as a gauge group. The gauge bosons of such a gauge theory would be color octets. I proposed to call these gauge bosons "chromons", but Gell-Mann insisted to call them "gluons", mixing the English language and the Greek language.

We called this new gauge theory "Quantum Chromodynamiccs" (QCD). The Lagrangian of QCD is [30, 50]:

$$\mathcal{L} = \overline{q}\left[i\gamma^\mu\left(\partial_\mu + ig_s\frac{\lambda^A}{2}\mathcal{A}^A_\mu\right) - m\right]q - \frac{1}{4}F^A_{\mu\nu}F^{A\,\mu\nu}\,, \quad (1.2.4)$$

where the λ^A are the Gell-Mann matrices, and

$$F^A_{\mu\nu} = \partial_\mu\mathcal{A}^A_\nu - \partial_\nu\mathcal{A}^A_\mu - g_s f_{ABC}\mathcal{A}^B_\mu\mathcal{A}^C_\nu\,. \quad (1.2.5)$$

f_{ABC} are called SU(3) structure constants. This Lagrangian is very similar to the Lagrangian of Quantum Electrodynamics. The electromagnetic field is replaced by the eight gluon fields $\mathcal{A}^A$, the electron mass by the quark mass, and the charge e is replaced by the strong coupling g_s. The strong interaction constant is defined by $\alpha_s = g_s^2/4\pi$.

However, the big difference between Quantum Electrodynamics and Quantum Chromodynamics is the presence of the $\mathcal{A}^2$ term in $F^A_{\mu\nu}$, not present in Quantum Electrodynamics. This term shows that a gluon interacts not only with a quark, but also with another gluon, and gives rise to 3- and 4-gluon couplings.

The quark masses, which appear in the Lagrangian of QCD, are not the masses of free quarks, but the masses, relevant inside the hadrons. The masses of the quarks depend on the energy scale. They are large at small energies and small at high energies. Here are the typical masses for the up-quark, the down-quark and the strange quark at the energy given by the mass of the Z-boson, $M_Z \simeq 91.2\,\text{GeV}$:

$$m_u \simeq 1.2\,\text{MeV},\ m_d \simeq 2.2\,\text{MeV},\ m_s \simeq 53\,\text{MeV}.$$

These masses describe the symmetry breaking of the $SU(3)_{\text{F}}$ flavor group. Interesting is the violation of the isospin symmetry. The down quark is heavier than the up quark. For this reason the neutron is heavier than the proton, and the proton is stable. If there would be no isospin violation, i.e. $m_u = m_d$, the proton would be heavier than the neutron due to the electromagnetic self-energy and it would decay into the neutron - life would not be possible.

Gell-Mann and I assumed that the interaction in QCD is zero at light-like distances. The light cone current algebra, which we had discussed in ref. [70], would not be changed. The confinement of colored states, i.e. the quarks and the gluons, would be due to the interaction at long distances.

Soon we realized that our assumption, that there is no interaction near the light-cone, was not correct. David Gross, Frank Wilczek and, independently, David Politzer calculated this interaction, which is the interaction, given by the Lagrangian, but near the light-cone the relevant coupling constant is not zero, but only very small.

The QCD Lagrangian describes a theory, which is asymptotically free. At small distances the interaction is very small, at large distances the interaction is strong. Thus the coupling constant is not constant, but a function of the energy. The sliding of the coupling constant g_s as a function of the renormalization mass μ is given by the beta-function $\beta(g_s)$:

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu} g_s(\mu) = \beta(g_s). \tag{1.2.6}$$

This beta function is positive for many theories, for example quantum electrodynamics. The fine structure constant α is at the energy of 100 GeV about 10% larger than at low energies.

The beta function can be calculated in perturbation theory. One finds for QCD:

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu} g_s(\mu) \simeq -\frac{1}{16\pi^2}\left(11 - \frac{2}{3}n_f\right) g_s^3(\mu). \tag{1.2.7}$$

Here the coefficient "11" describes the contribution of the gluons to the beta function. The asymptotic freedom of QCD is due to this coefficient - it is related to the self-interactions of the gluons. The number n_f is the number of the different quark flavors. For the three quarks up, down and strange one has $n_f = 3$.

In QCD one can describe the energy dependence of the coupling constant by introducing a scale parameter Λ:

$$\alpha_s(\mu^2) \simeq \frac{4\pi}{\left(11 - \frac{2}{3}n_f\right)\ln\left(\frac{\mu^2}{\Lambda^2}\right)}. \tag{1.2.8}$$

This scale parameter has been measured by many experiments (see Section 3.2):

$$\Lambda = (332 \pm 17)\,\text{MeV}. \tag{1.2.9}$$

In experiments one has measured the scale dependence of the coupling constant. It agrees very well with the theoretical prediction. We also mention the value of the coupling constant at the mass of the Z-boson, where it was possible to measure the coupling constant rather precisely (see Section 3.2):

$$\alpha_s = 0.1181 \pm 0.0011. \tag{1.2.10}$$

In QCD, Bjorken scaling in deep inelastic scattering is not an exact property of the strong interactions. The quark distribution functions change slowly at high energies. This change can be calculated in perturbation theory (see Section 2.3). The results agree rather well with the experimental results. Also the gluon distribution function $g(x)$ has been measured. Since the gluons and the quarks contribute to the momentum of a high energy proton, the following sum rule must be obeyed:

$$\int_0^1 x\left[g(x) + u(x) + \overline{u}(x) + d(x) + \overline{d}(x) + \right. \\ \left. + s(x) + \overline{s}(x)\right] \mathrm{d}x = 1. \tag{1.2.11}$$

Using the scale parameter Λ, one can in principle calculate many properties of the strong interactions, for example the masses of the hadrons like the proton mass: $m_p = \text{const} \times \Lambda$. The proton mass depends also on the quark masses, however the up and down quark masses are very small and can be neglected. The calculations of the hadron masses are complicated and are often carried out by discretizing space and time (see Section 4 on Lattice QCD).

In QCD one can also change the three quark masses. For example we can assume that the three quark masses are zero. In this case the flavor group $SU(3)_{\text{F}} \times SU(3)_{\text{F}}$ would be unbroken. The three pions, the four K-mesons and the η - meson would be massless and the eight vector mesons would have the same mass. There is not a ninth massless pseudoscalar meson, since the singlet axial current has an anomaly:

$$\partial_\mu\left(\bar{u}\gamma^\mu\gamma_5 u + \bar{d}\gamma^\mu\gamma_5 d + \bar{s}\gamma^\mu\gamma_5 s\right) \\ = const \times g_s^2\,\epsilon^{\mu\nu\rho\sigma} F^A_{\mu\nu} F^A_{\rho\sigma}. \tag{1.2.12}$$

where $\epsilon^{\mu\nu\rho\sigma}$ is the totally antisymmetric tensor. In ref. [30] Gell-Mann and I also studied what happens if the quarks are removed from the QCD Lagrangian. In this case only the eight gluons are present. At low energies there would be a discrete spectrum of particles, which consist of gluons - the glue mesons, gluonium particles or glueball (see Section 8.4). If the three quarks are introduced, the glue mesons would mix with the quark-antiquark mesons. The experimentalists have thus far not clearly identified a glue meson. Presumably in nature there are only mixtures of glue mesons and quark-antiquark mesons. But there might be mesons, which are essentially glue mesons, since the mixing is very small for these mesons.

It is useful to consider the theory of QCD with just one heavy quark Q. The ground-state meson in this hypothetical case would be a quark–antiquark bound state (see Sections 8.1, 8.6). The effective potential between the quark and its antiquark at small distances would be a Coulomb potential proportional to $1/r$, where r is the distance between the quark and the antiquark. However, at large distances the self-interaction of the gluons becomes important. The gluonic field lines at large distances do not spread out as in electrodynamics. Instead, they attract each other. Thus the quark and the antiquark are connected by a string of gluonic field lines. The force between the quark and the antiquark is constant, i.e. it does not decrease as in electrodynamics. The heavy quarks are confined.

In the annihilation of electrons and positrons at very high energies it has been possible to test the theory of quantum chromodynamics rather precisely. If an electron and a positron collide, a quark and an antiquark are produced. The two quarks move away from each other almost with the speed of light. Since the two quarks do not exist as free particles, they fragment into two jets of hadrons, mostly pions. These particles form two narrow jets. These jets have been observed since 1979 at the collider at DESY, later at the LEP-collider at CERN. Sometimes a quark emits a high energy gluon, which also fragments into hadrons. Thus three jets are produced, two quark jets and one gluon jet. Such three jet events have been observed since 1979 at DESY, later at CERN (see Section 2.2).

Now we consider high energy collisions of atomic nuclei, for example collisions of lead nuclei. Such collisions are studied at the Relativistic Heavy Ion Collider (RHIC) in Brookhaven, at Fermilab and at the LHC in CERN. In such collisions a new state of matter is produced for a short time, a quark-gluon-plasma. Astrophysicists assume that such a plasma exists also for a long time near the center of a large neutron star (see Section 7.1).

Right after the Big Bang the matter was a quark-gluon-plasma. During the expansion of the universe the plasma changed later into a gas of protons and neutrons (see Section 7.2).

In the fall of 1973 I was convinced, that Gell-Mann and I had discovered the correct theory of the strong interactions: Quantum Chromodynamics. Almost every day I discussed this theory with Richard Feynman, and he also thought that it was correct. In 1974 Feynman gave lectures on QCD. But Gell-Mann still thought that the true theory of the strong interactions should be a theory based on strings.

In the years after 1973 it became clear that QCD is the correct theory of the strong interactions. I was proud that I had contributed to the birth of this theory, which is now a major part of the Standard Theory of particle physics.

2 Experimental Foundations

Conveners:
Franz Gross and Eberhard Klempt

Quantum Chromodynamics or QCD: What a gorgeous theory! You start with free colored quarks. You request invariance with respect to the exchange of colors at any time and any space point, and the quarks interact. That is all what QCD requires. It is a remarkable simple concept. But: is this the true theory of strong interactions? In this Section, the milestones are discussed which convinced even sceptical physicists of the quark model and of the new theory.

A breakthrough was achieved in the *November revolution:* Charmonium was discovered at SLAC, the c-quark was shown to exist, the GIM meachanism (proposed by Sheldon Glashow, John Iliopoulos and Luciano Maiani in 1964) explaining the absence of neutral currents in weak interactions found experimental confirmation.

John B. Kogut's contribution remembers the excitement in these days. A new spectroscopy came into life with a convincing interpretation based on the famous Cornell potential. The mediators of the strong interaction, called gluons, carry – unlike the electrically neutral photons – themselves the charge of the strong interaction and are confined. San Lau Wu recalls her personal contributions to the discovery of gluons at DESY where events were found in which e^+e^- annihilate into three bunches of particles, three jets. The three jets were interpreted as processes in which the two quarks – observed as jets – radiate off a gluon which manifests itself as the third jet.

The evidence for the correctness of QCD grew rapidly. Yuri Dokshitzer reminds us of the most important steps.