Structure of matter, 3

The particle zoo

Prior to the 1930s the fundamental structure of matter was believed to be extremely simple: there were electrons (each with mass about 0.5 MeV), e^-, photons (no mass), γ, and protons (mass about 938 MeV), p^+. Starting in 1932 the world began to get a lot more complicated. First came Dirac's positron (e^+, with same mass as the electron), postulated in 1928 but mostly ignored until Anderson's accidental discovery (see SM 1). Soon after, the neutron (n) was identified (mass about 940 MeV). In beta decay, the neutron transforms into a proton and an electron. The energy of the electron in beta decay has a maximum cutoff and is otherwise "never" observed to be the same–as it would be if there were only two products. It is almost as if energy is not conserved in beta decay. In 1930, Wolfgang Pauli proposed that a third, unseen, particle was also emitted and that the three products conserved energy and momentum by sharing them in a variety of unpredictable ways. The new particle would have to have spin-1/2 (because the neutron, proton, and electron all have spin-1/2 and not even the crazy rules of addition in quantum mechanics allow $1/2+1/2=1/2$) and be electrically neutral (because the neutron is neutral and the proton plus electron is also neutral). Eventually, Pauli's particle–the neutrino, ν–was directly detected in the 1950s. This set of particles was all that was needed to make sense of nuclei and their properties.

But wait! There's more. With the development of higher and higher energy accelerators and better and better detectors, the list of "particles" grew at breakneck pace in the 1950s and 60s. Many of the newly discovered particles were much more massive than the nucleons. All were unstable—some lasting only 10^{-24} s or so. Today, the Particle Data Group (an international body of over 200 individuals charged with keeping and organizing all particle data) publishes a 1800-plus page book on known particle measurements every two years and a roughly 300-page summary every year, and maintains an amazingly useful searchable website, http://pdg.lbl.gov/, containing not only data but also tutorial reviews on particle physics and cosmology. There are many hundreds of entries in the particle database, each detailing the mass, charge, spin, lifetime, and lots of other properties for what are amusingly called the "elementary particles of nature." Note that the mechanisms by which particles transform into other particles are classified by the particles' lifetimes. Lifetimes on the order of 10^{-24} or 10^{-23} s result from "strong interactions." Lifetimes on the order of 10^{-16} s result from electromagnetic interactions. Lifetimes much longer than 10^{-16} s result from "weak interactions."

Example: The particle Δ^{++} decays into a proton and a positively charged pion, π^+, with a mean lifetime of about 5x10^{-24} s; this is a "strong decay" process. The π^+ in turn decays into a positive muon, μ^+, plus a neutrino with a mean lifetime of about 3x10^{-8} s; that's a "weak decay" process. The neutral cousin of the π^+, the π^0, decays into two photons with a mean lifetime of about 8x10^{-17} s; that's an electromagnetic decay process.

As shown in SM1, a small number of elementary particles appear to have no substructure (at least at the energy scales this has been tested). The six leptons and the five bosons listed in SM1 have all been "observed" in nature or in laboratory experiments. Quarks and gluons have not been observed independently; their existence has only been indirectly inferred. All of the strongly interacting particles produced in accelerator collisions are thought to be composite entities, consisting of collections of quarks, anti-quarks, and gluons. These composite “particles” are categorized by their spins. Particles with integer spin (0, 1, 2, …) are “mesons.” The lightest mesons consist of a primary quark and a primary anti-quark bound together by the exchange of virtual gluons. (“Primary” because, in the quantum field theory picture, the virtual gluons make virtual quark–anti-quark pairs; in this picture, the actual make-up of strongly interacting particles is really complicated.) Particles with half odd-integer spin (1/2, 3/2, 5/2, …) are “baryons.” The lightest baryons consist of three primary quarks also bound together by the exchange of virtual gluons. The proton (two u s and a d) and neutron (two d s and a u) are the lightest of the light baryons. There is some evidence that very heavy, very short-lived mesons and baryons might have more exotic combinations of quarks and anti-quarks.

(Historical comment: In 1963, **Murray Gell-Mann** recognized that approximate regularities of the then known baryons and mesons could be rationalized if these entities were composed of more elementary particles, each being one of three possible types. Thus, in this scheme, if baryons were made of three of these particles then there would 3x3x3 possible baryons, and (according to Gell-Mann’s “group-theoretic” arithmetic) 3x3x3 = 10+8+8+1 different family sizes. If mesons were made of two of these particles then there would be a family size of 3x3 = 8+1. Mass differences within the families could be explained by the differences in quark masses. On the other hand, since such sub-particles had never been seen, Gell-Mann at first thought they were merely mathematical artifacts of some kind. Experiments done in 1967, however, eventually convinced everyone that they might actually exist. The experiments consisted of bombarding protons and neutrons with high-energy electrons and looking at the energies of the electrons emerging from the collisions. Such processes are called “deep *inelastic* scattering” because typically the outgoing electron has less energy than when it entered—which implies that something within the nucleon has taken up the missing energy. Originally, Gell-Mann had called these building blocks “kworks” for some reason, but, according to popular legend, in reading James Joyce’s impenetrable novel, *Finnegan’s Wake*, he found the phrase “three quarks for Muster Mark” on page 383. The close similarity of kwork and quark and the fact that “three quarks” appeared on a page in which 3’s and 8’s were in such exquisite juxtaposition, convinced Gell-Mann to call his particles quarks (though he always insisted that they should still be pronounced kworks).)

Color

The Δ^{++} is a relatively low mass, spin-3/2 baryon consisting of three primary u quarks. But to get spin = 3/2 means that at least two (if orbital angular momentum = 1 or 2), and maybe all three (if orbital angular momentum = 0), of the quarks have their spin angular momenta in the same direction. Quarks are supposed to be spin-1/2 particles, that is, fermions. The Pauli Exclusion Principle states that no two identical fermions can have all the same quantum numbers, so the Δ^{++} (and others) presents a challenge to the validity of the quark model.

This problem can be solved, however, by introducing a new property of matter that can distinguish one u from another. In fact, this property has to have three values if the state uuu is to have all identical spins. This property, proposed in 1964 by O.W. Greenberg, is defined to have the values R,G,B . Thus the Δ^{++} baryon can be thought of as the quark combination $u_R u_G u_B$, all with aligned spins. The properties R,G,B are called “colors” in analogy with the three primary colors of light. Of course, this identification is purely abstract and has nothing to do with visible light. Following in this fanciful naming scheme, the quark types u,d,c,s,t,b are called “flavors.” Thus, each quark has an electric charge (with magnitude equal to 1/3 or 2/3 the electronic charge), one of two possible spin directions, a flavor, and one of three possible color charges. As we don't observe color charge in nature, there is an extra assumption: the only allowed combinations of quarks and anti-quarks are "colorless."

Color combinations follow the rules of adding colors of light. If equal amounts of red, green, and blue light are superposed, the result is "white," that is, no color. Formally, $R+G+B=0$. In this relation, colors can be subtracted from both sides; for example, $G+B=-R$. The meaning of $-R$ is the color that has to be added to red to get white. This is the combination $G+B$, which is "cyan," the "complement" of red. Other complements are $R+B$ = "magenta," the complement of green, and $R+G$ = "yellow," the complement of blue. Because complements "cancel out" colors they can be viewed as "anti-colors." Because baryons are colorless combinations of three primary quarks, at any instant the quark colors have to be R,G,B. Mesons are quark, anti-quark pairs, so the anti-quark has to have the anti-color of the quark. In any quark interactions, total color has to be conserved. This property of color produces a "color force" that determines how quarks interact.

Color force and gauge freedom

Quantum electrodynamics is such a successful theory in part because we know from Maxwell's Equations how to put electric and magnetic effects into the Dirac Equation for electrons: that is, replace E_{op} by $E_{op}-q_E\phi_E$ and replace $\vec{p}_{op}$ by $\vec{p}_{op}-q_E\vec{A}_E$. Quarks have spin-1/2, so should obey the Dirac Equation also, but there is no pre-existing theoretical structure similar to Maxwell's Equations to guide how to insert the desired interactions. Deeper thinking about the structure of the Dirac Equation when electromagnetism is included paves the way.

First, an essential fact about electromagnetic potentials: they aren't unique. Because the physical electric and magnetic fields are derived by differentiating the potentials, it is possible to add "stuff" to the potentials and still get the same $\vec{\mathcal{E}}$ and $\vec{B}$. Trivially, the "stuff" might be just a constant number added to ϕ_E and/or a constant vector added to $\vec{A}_E$. Because $\vec{\mathcal{E}}=-grad(\phi_E)-\partial\vec{A}_E/\partial t$ and $\vec{B}=curl(\vec{A}_E)$, more interesting "stuff" can be added, such as $\phi'_E=\phi_E-\partial\lambda/\partial t$ and $\vec{A}'_E=\vec{A}_E+grad(\lambda)$, where $\lambda(\vec{r},t)$ is *any* (smoothly continuous) function of position and time. It should be clear how this leads to $\vec{\mathcal{E}}'=\vec{\mathcal{E}}$; the vector-calculus identity $curl(grad(\lambda))=\vec{0}$ is needed to make $\vec{B}'=\vec{B}$. Such additions do *not* alter the Maxwell equations, so electromagnetic physics doesn't care about such additions. Such additions are "gauge transformations" (presumably deriving from "gauge pressure," where the pressure reading on the gauge only tells the difference between internal and external pressures not the absolute pressure anywhere); formally we say that *electromagnetism is invariant under gauge transformations*. But, if the potentials were transformed as above then the energy and momentum operators in the Dirac Equation would acquire extra terms. Thus, the transformations would cause extra energy and momentum to show up in the electron field and, because Maxwell relates electromagnetic fields to the electron field, they would also alter the electromagnetic fields! But that's not supposed to happen. This sounds like a logical absurdity.

But requiring the electron field, ψ, to transform at the same time the potentials do saves the day. The correct transformation for this purpose is $\psi'=e^{i\theta(\vec{r},t)}\psi$, where $\theta=q\lambda/\hbar$; this transformation changes the *phase* of the complex field ψ differently at every point in space-time. This works because $E_{op} = i\hbar\partial/\partial t$ and $\vec{p}_{op} = -i\hbar grad()$, which when operating on ψ' yield $e^{iq\lambda/\hbar}E_{op}\psi - (\partial\lambda/\partial t)\psi'$ and $e^{iq\lambda/\hbar}\vec{p}_{op}\psi + grad(\lambda)\psi'$, respectively. The derivative-of-λ terms from the ψ transformation exactly "kill off" the derivative-of-λ terms from the potential gauge transformations. Consequently, there are no unwanted contributions to the dynamics of the electron, and because the ψ transformation leads to $|\psi'|^2 = |\psi|^2$, the probability of finding the electron anywhere doesn't change either. *The latter relation is equivalent to stating that electric charge is conserved*. Physics doesn't know that $\psi' = e^{i\theta(\vec{r},t)}\psi$ has occurred; it is a second kind of gauge transformation. So all is happy. The potential fields can be gauge transformed *in their way* leaving electromagnetism invariant as long as ψ is simultaneously gauge transformed *in its way* leaving electron physics invariant; physics doesn't care. Note that it is not possible to separate electromagnetism from charge; if the Dirac field has electric charge it automatically has to be interacting with electromagnetic potential fields–and these in turn obey Maxwell's Equations. This interweaving of the gauge freedom of fields demands that electric charge be conserved whenever charges and photons interact.

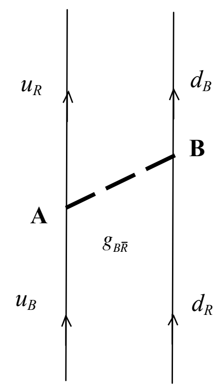
Importantly for the goal of finding a color interaction description, the story outlined above can be run in reverse. Start with a charged Dirac field, but no electromagnetic interactions. Stipulate that charge is conserved when ψ is transformed by $\psi' = e^{i\theta(\vec{r},t)}\psi$. (Such a local phase transformation is more in keeping with special relativity than a global phase change where θ takes the same value everywhere at the same time.) Now, unwanted derivatives occur in the Dirac Equation. They can be removed by introducing ϕ_E and $\vec{A}_E$ fields that multiply ψ and transform simultaneously with ψ in just the right way to kill off the unwanted derivatives. The new fields obey Maxwell's Equations: $grad^2(\phi_E) - \frac{1}{c^2}\frac{\partial^2\phi_E}{\partial t^2} = -\frac{\rho_E}{\varepsilon_0}$ and $grad^2(\vec{A}_E) - \frac{1}{c^2}\frac{\partial^2\vec{A}_E}{\partial t^2} = -\mu_0\vec{J}_E$, where ρ_E is the electric charge density and $\vec{J}_E$ is the electric current density associated with ψ. Thus, the form of the electromagnetic interactions is "derived" from the freedom to arbitrarily change the phase of ψ locally at every point in space and at every moment in time. Note that the fields $\{\psi, \phi_E, \vec{A}_E\}$ are irreducibly intertwined. This is yet another example—arguably, because all of the SMPP depends on it, *the most important*—of how symmetry and dynamics are intertwined. Cool. (See the Appendix below.)

Back to color. We extend the "gauge theory" idea that works for QED to construct a *quantum chromodynamics* (QCD), a theory of color force interactions. The argument starts with the inference that because color is not observed in the macroscopic world, baryons and mesons must have no net color and color must be conserved in any of their interactions. Quarks are spin-1/2 particles, so they obey the Dirac Equation. The quark field, ψ, like the electron field in QED has four major components, two spin components for particles and two spin components for anti-particles. In addition, for quarks each of these four components has three additional color sub-components. The analogy of a QED phase transformation is a "rotation" in 3-D color space: $\psi' = R\psi$, where R is now a 3x3 matrix, instead of just $e^{i\theta}$. (The magnitude of $e^{i\theta}$ is 1; similarly the "magnitude" of R is 1 also; in mathy words, R is a "unitary transformation matrix.") The components of a 3x3 matrix have a row index and a column index. For color transformations, it is useful to think of the row labels as colors and the column labels as anti-colors. If–again in analogy with QED and consistent with special relativity– R is made a function of space and time, then when the quark Dirac Equation is transformed by a color rotation derivatives of R appear, which, if unattended to, would change the quark physics by introducing unwanted contributions to the energy and momentum. But, quark physics is supposed to be invariant under color transformation, so, just as in QED, ϕ and $\vec{A}$ "potential" fields are inserted into the Dirac Equations that transform in the correct way when $\psi' = R\psi$ to kill off the unwanted derivative terms. Thus, the QCD Dirac Equation is

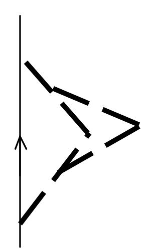
$$(E_{op} - q_C\phi_C)\psi = c\vec{\alpha}\cdot(\vec{p}_{op} - q_C\vec{A}_C)\psi + \eta mc^2\psi$$

where ψ represents the quark field, $\phi_C, \vec{A}_C$ are the "color potential" fields, and q_C is the "color charge." In QED (at "low" energies and "large" separations) the strength of the interaction between charges and photons is measured by the dimensionless quantity, $\alpha_E = k_E q_E^2/\hbar c$, where q_E is electric charge. In other words, $q_E \propto \sqrt{\alpha_E}$. Similarly, the color charge is defined as $q_C = \sqrt{\alpha_C}$, where α_C is the dimensionless strength of the quark-gluon interaction. The color potentials then have the dimensions of energy and momentum.

The particles associated with the color potential fields are "gluons," and in QCD quarks interact via gluon exchange. But there's something different from QED. The unwanted derivative-of-R terms are 3x3 color matrices, so each component of the potential fields, $(\phi, A_x, A_y, A_z)_C$, must be a 3x3 matrix as well. In other words, like the matrix R **the color force potential fields carry color and anti-color**. Similar to photons, gluons carry spin-1 and are massless and have zero electric charge. On the other hand, gluons carry color and anti-color, the values of which ensure that color is conserved in any process involving quarks. See the figure to the right, for example. *This is part of a more elaborate system in which total color is white*. The in-state (at the bottom of the figure) consists of a blue u quark and a red d quark while the out-state (at the top) consists of a red u quark and a blue d quark. That is, the diagram has color blue-red in and out. The gluon emitted at **A** and absorbed at **B** does not change the quark flavors but does change their colors. At one instant between **A** and **B** there is a red u quark on the left and a red d quark on the right. As color has to be conserved, the gluon must carry anti-red to cancel one of the reds out, and blue to make the whole diagram blue-red.



Because gluons carry color, the "Maxwell Equations" for them are more complicated than for photons. The sources for gluons are not only color-charged quarks, but also color-charged gluons themselves. The only source for photons in QED is electrical charge, so photons cannot make photons. But gluons *can* make gluons, and that causes QCD to be *very* different from QED. In the diagram to the right, for example, a quark emits a virtual gluon, but before it is reabsorbed, that gluon emits another gluon. The first gluon might have colors red/anti-blue, for example, then change to red/anti-green by emitting a green/anti-blue gluon. In addition, gluon-gluon interactions produce both three-prong (as in QED) and four-prong vertices, so QCD calculations are structurally more complicated than those in QED.



Another aspect of the color force makes the mathematics of QCD much more difficult than that of QED. In QED, each vertex in a Feynman diagram reduces the probability of the process in the diagram by a factor of $\alpha_E \approx 1/137$. Thus, the simplest diagrams in QED are the most important for doing (most) calculations. On the other hand, an equivalent measure of strength of interaction for the color force, α_C, is larger, so complicated diagrams, like the one above, cannot be ignored in QCD.

Appendix

Emmy Noether first formalized the interplay between conservation laws and the symmetry of physical dynamics in the late 1910s. Noether inherited a gift for mathematics from her mathematician father. Though it wasn't easy (women weren't supposed to do real math in Germany at the beginning of the 20^{th} century), she earned a PhD in 1907 and eventually became a professor at the University of Göttingen (where she mentored 16 of her own PhD students). Because she was Jewish, Noether left Germany in 1933. She obtained a teaching position at Bryn Mawr College in Philadelphia. Two years later, she died while recovering from surgery. She was only 53.

In her professional life, Noether produced numerous important results in the then new field of abstract algebra. For this work she is a kind of giant in modern mathematics. Her early studies of symmetry, conservation laws, and dynamics, inspired by Einstein's general relativity, can be summarized as follows: " if a dynamical system is symmetric under a continuous transformation, there will be an associated conservation law" and conversely, "if there is a conservation law, then hunt for an associated dynamical system that might be symmetric under a continuous transformation." This is roughly the much-discussed *Noether's Theorem*. Long after it was first developed, Noether's theorem has become the central building block for the Standard Model of Particle Physics. The importance of Noether's results in dynamical systems surely would have merited a Nobel Prize in Physics if she had lived long enough to witness the emergence of the modern world of particles.