
QUANTUM FIELD THEORY

From QED to the Standard Model

Silvan S. Schweber

Until the 1980s, it was usual to tell the story of the developments in physics during the twentieth century as “inward bound” – from atoms, to nuclei and electrons, to nucleons and mesons, and then to quarks – and to focus on conceptual advances. The typical exposition was a narrative beginning with Max Planck (1858–1947) and the quantum hypothesis and Albert Einstein (1879–1955) and the special theory of relativity, and culminating with the formulation of the standard model of the electroweak and strong interactions during the 1970s. Theoretical understanding took pride of place, and commitment to reductionism and unification was seen as the most important factor in explaining the success of the program. The Kuhnian model of the growth of scientific *knowledge*, with its revolutionary paradigm shifts, buttressed the primacy of theory and the view that experimentation and instrumentation were subordinate to and entrained by theory.¹

The situation changed after Ian Hacking, Peter Galison, Bruno Latour, Simon Schaffer, and other historians, philosophers, and sociologists of science reanalyzed and reassessed the practices and roles of experimentation. It has become clear that accounting for the growth of knowledge in the physical sciences during the twentieth century is a complex story. Advances in physics were driven and secured by a host of factors, including contingent ones. Furthermore, it is often difficult to separate the social, sociological, and political factors from the technical and intellectual ones.

In an important and influential book, *Image and Logic*, published in 1997, Peter Galison offered a framework for understanding what physics was about in the twentieth century. Galison makes a convincing case for regarding experimentation, instrumentation, computational modeling, and theory as

¹ See A. Pais, *Inward Bound: Of Matter and Forces in the Physical World* (Oxford: Clarendon Press, 1986); T. Yu Cao, *Conceptual Developments of 20th-Century Field Theories* (Cambridge: Cambridge University Press, 1997); Paul Davies, ed., *The New Physics* (Cambridge: Cambridge University Press, 1989); and R. E. Marshak, *Conceptual Foundations of Modern Particle Physics* (Singapore: World Scientific, 1993).

quasi-autonomous subcultures with languages and practices that are distinct, yet linked and coordinated. Experimental, theoretical, and instrumental practices do not all change of a piece – each has its own periodization; and their relation to one another varies with the specific historical situation in which each is embedded. There is, in fact, continuity of experimental practices across theoretical and instrumental breaks.²

Image and Logic is a brief for “mesoscopic history” – for history written at a level between macroscopic, universalizing history and microscopic, nominalistic history. Galison proposes treating the movement of ideas, objects, and practices as one of *local* coordination – both social and epistemic – and their interconnections and linkages are made possible through the establishment of pidgin and creole languages. He sees the separate, but correlated, subcultures of physics as bound and stabilized by such interlanguages. These suggestions are attractive and valuable. However, limited as I am in this chapter to choices for presentation, most of the following account lies squarely within the history of ideas. The reader is referred to recent books by Andrew Pickering, Gerard 't Hooft, Lillian Hoddeson, and others for more mesoscopic accounts.³

I have not tried to fit my presentation of the history of quantum field theory (QFT) from QED (quantum electrodynamics) to QCD (quantum chromodynamics) into a preconceived pattern – whether that of Thomas Kuhn or that of Imre Lakatos. My concern has been with the telling of the story. One could easily cast the history into a Lakatosian mold of research programs – with S-matrix and field theory the two competing modes.⁴ Similarly, one could pick from that history examples that would instantiate both of Kuhn's notions of paradigm; namely, paradigm as achievement – the body of work that emerges from a scientific crisis and sets the standard for addressing problems in the subsequent period of normal science – and paradigm as a set of shared values – the methods and standards shared by the core of workers who decide what are interesting problems and what counts as solutions, and determine who shall be admitted to the discipline and what shall be taught to them.

Furthermore, one could readily give examples of Kuhnian revolutions. Renormalization theory as formulated in the period from 1947 to 1949,

² Peter Galison, *Image and Logic: A Material Culture of Microphysics* (Chicago: University of Chicago Press, 1997); see also Galison, *How Experiments End* (Chicago: University of Chicago Press, 1987).

³ L. M. Brown and Lillian Hoddeson, *The Birth of Particle Physics* (Cambridge: Cambridge University Press, 1983); Laurie M. Brown, Max Dresden, and Lillian Hoddeson, eds., *Pions to Quarks: Particle Physics in the 1950s* (Cambridge: Cambridge University Press, 1989); L. M. Brown and H. Rechenberg, “Quantum Field Theories, Nuclear Forces, and the Cosmic Rays (1934–1938),” *American Journal of Physics*, 59 (1991), 595–605; Gerard 't Hooft, *In Search of the Ultimate Building Blocks* (Cambridge: Cambridge University Press, 1997); Andrew Pickering, *Constructing Quarks: A Sociological History of Particle Physics* (Edinburgh: Edinburgh University Press, 1984); Michael Riordan, *The Hunting of the Quark: A True Story of Modern Physics* (New York: Simon and Schuster, 1987).

⁴ Steve Weinberg, *Dreams of a Final Theory: The Search for the Fundamental Laws of Nature* (New York: Pantheon, 1992); J. T. Cushing, *Theory Construction and Selection in Modern Physics: The S Matrix* (Cambridge: Cambridge University Press, 1990).

culminating with the work of Freeman Dyson (b. 1923), is surely one such revolution; broken symmetry, as formulated by Jeffrey Goldstone (b. 1933) and Yoichiro Nambu (b. 1921), in the early 1960s another. One probably could constrain the history of quantum field theory into a Kuhnian mold. But I believe that much would be lost in so doing, in particular, a perspective on the cumulative and continuous, yet novel, components of the developments. It seems to me that the later Kuhn's emphasis on "lexicons" – the learnable languages, algorithms, laws, and facts of a given tribe of scientific workers – constitutes a more useful approach to the growth of our knowledge in high energy physics.

Equally helpful, I believe, is Ian Hacking's notion of a style of scientific reasoning: "A style of reasoning makes it possible to reason toward certain kinds of propositions, but does not of itself determine their truth value."⁵ A style determines what may be true or false. Similarly, it indicates what has the status of evidence. Styles of reasoning tend to be slow in evolution and are vastly more widespread than paradigms. Furthermore, they are not the exclusive property of a single disciplinary matrix. Thus, Feynman's space-time approach to nonrelativistic quantum mechanics encapsulates a new style of reasoning: All physical measurements and interactions can be considered as scattering processes. I believe Hacking's notion of a style of reasoning captures something right about the history of quantum field theory.

The use of symmetry is another example of a style of reasoning. The fact that such styles of reasoning are useful in both particle physics and condensed matter physics – and, in point of fact, cross-fertilize these fields – illustrates the (nonlinear) additive properties of styles of reasoning. Since a style of reasoning can accommodate many different paradigms, it is not surprising that one should discern Kuhnian revolutionary episodes within its evolution. The delineations of such revolutions are helpful guidelines and periodizations of the history of the field. But it is the identification of the different styles of reasoning that is, I believe, the important task for the intellectual historian attempting to relate that history.⁶

QUANTUM FIELD THEORY IN THE 1930s

The history of theoretical elementary particle physics from the 1930s until the mid-1970s can be narrated in terms of oscillations between the particle and field viewpoints epitomized by Paul Dirac (1902–1984) and by Pascual

⁵ Ian Hacking, "Styles of Scientific Reasoning," in *Post-Analytic Philosophy*, ed. John Rajchman and Cornell West (New York: Columbia University Press, 1986), pp. 145–65.

⁶ For other accounts, and in greater detail, Silvan S. Schweber, "From 'Elementary' to 'Fundamental' Particles," in *Science in the Twentieth Century*, ed. John Krige and Dominique Pestre (Amsterdam: Harwood, 1997), pp. 599–616, and Schweber, *QED and the Men Who Made It* (Princeton, N.J.: Princeton University Press, 1994).

Jordan (1902–1980), as noted by Olivier Darrigol in Chapter 17.⁷ That the *field* approach was richer in potentialities and possibilities than the *particle* one is made evident by the quantum field theoretic developments of the 1930s (QFT). All these advances took as their point of departure insights gained from the quantum theory of the electromagnetic field and, in particular, from the centrality of the concept of emission and absorption of quanta.

Enrico Fermi's (1901–1954) theory of beta decay was an important landmark in the field theoretic developments of the 1930s. It had been recognized since 1915 that the nucleus was the site of all radioactive processes, including β -radioactivity in which a nucleus ejects an electron. It was, therefore, natural to believe that electrons existed in the nucleus. Already in 1914, Ernest Rutherford (1871–1937) had assumed that the hydrogen nucleus is the positive electron – he called it the H-particle – and he conjectured that nuclei were made of H-particles and electrons. During the 1920s, the generally accepted model of a nucleus was that it consisted of the two elementary particles then known: protons and electrons. Rutherford in his Bakerian Lecture of 1920 had suggested that a proton and an electron could bind and create a neutral particle, which he believed was necessary for the building up of the heavy elements. However, if nuclei were assumed to be composed of protons and electrons, the Pauli principle made it difficult to understand the spin of certain nuclei, such as N^{14} . Similarly, should there be electrons in the nucleus, their magnetic moment – as determined by the hyperfine structure of atoms – ought to be much larger than the values determined experimentally, which are three orders of magnitude smaller than atomic moments. Confusion reigned, compounded by the difficulty in understanding β -decay.

The process of β -decay – wherein a radioactive nucleus emits an electron (β -ray) and increases its electric charge from Z to $Z + 1$ – had been studied extensively during the first decade of the century. If the process is assumed to be a two-body decay, that is, if the decay consists in a nucleus undergoing the process $A^Z \rightarrow A^{Z+1} + e^-$, then energy and momentum conservation requires the electron to have a definite energy. However, in 1914, James Chadwick had found that the energy of the emitted electrons had a continuous energy spectrum – up to some maximum energy. At the maximum electron energy, energy conservation was found to hold – to the accuracy of the measurements in the experiment.

By the end of the 1920s, no explanation of the continuous β -spectrum had proven satisfactory, and some physicists, in particular Niels Bohr, were ready to give up energy conservation in β -decay processes. In December 1930, Wolfgang Pauli (1900–1958), in a letter addressed to the participants of

⁷ See Helge Kragh, *Dirac: A Scientific Biography* (Cambridge: Cambridge University Press, 1990); Steven Weinberg, *The Quantum Theory of Fields*, 2 vols. (Cambridge: Cambridge University Press, 1995–6).

a conference on radioactivity, proposed saving energy conservation with “a desperate remedy,” suggesting that

there could exist in the nuclei electrically neutral particles that I wish to call neutrons [later renamed neutrinos by Fermi], which have spin $\frac{1}{2}$ and obey the exclusion principle. . . . The continuous β -spectrum would then become understandable by the assumption that in β -decay a [neutrino] is emitted together with the electron, in such a way that the sum of the energies of the [neutrino] and electron is constant.⁸

Fermi took Pauli’s hypothesis seriously and in 1933 formulated his theory of β -decay. It marked a change in the conceptualization of “elementary” processes. In the introduction to his paper, Fermi indicated that the simplest model of a theory of β -decay assumes that electrons *do not* exist as such in nuclei before β -emission occurs

but that they, so to say, acquire their existence at the very moment when they are emitted; in the same manner as a quantum of light, emitted by an atom in a quantum jump, can in no way be considered as pre-existing in the atom prior to the emission process. In this theory, then, the total number of the electrons and of the neutrinos (like the total number of light quanta in the theory of radiation) will not necessarily be constant, since there might be processes of creation or destruction of these light particles.⁹

Fermi’s theory made clear the power of a quantum field theoretical description.

For nuclear physics, 1932 was the *annus mirabilis*. The discovery of the neutron by James Chadwick (1891–1974) working at the Cavendish Laboratory led quickly to the view that a nucleus of mass number A is a composite system built out of Z protons and $(A - Z)$ neutrons. The neutron, which was assumed to be an electrically neutral, spin $\frac{1}{2}$ particle with a mass roughly equal to that of the proton, made possible the application of quantum mechanics to the elucidation of the structure of the nucleus, as was shortly done in a series of papers by Heisenberg, based on short-range (static) two-body nucleon–nucleon interactions.

After the discovery of the neutron and of the positron, matter was thought to consist of two sets of entities: electrons and neutrinos (and their antiparticles) and neutrons and protons (and their antiparticles). The charged members of the two groups could interact with one another electromagnetically. Electrons and neutrinos interacted with neutrons and protons through the Fermi interaction; neutrons and protons interacted “strongly” through nuclear forces. The neutron and the protons were recognized as being very

⁸ Wolfgang Pauli, letter to the “Dear Radioactive Ladies and Gentlemen,” 4 December 1930, in K. von Meyenn, *Wolfgang Pauli: Wissenschaftlicher Briefwechsel*, vol. 2 (New York: Springer-Verlag, 1979).

⁹ Enrico Fermi, “Versuch einer Theorie der β -Strahlen. I,” *Zeitschrift für Physik*, 88 (1934), 161–71.

similar, yet also different. They have different electric charges and electromagnetic interactions, but interact very similarly in their “strong” (nuclear) interactions.

The indifference of the nuclear force to the nucleons involved became expressed formally by considering the neutron and the proton as having a new “internal” quantum property, called isotopic spin. Neutrons and protons differ merely in the value of the z-component of their “isotopic” spin. This attribution of an isotopic spin to nucleons by Heisenberg was the first example of the two kinds of internal quantum numbers eventually used to classify particles, namely: (1) (conserved or approximately conserved) additive quantum numbers, like electric charge, strangeness, baryon, and lepton numbers; and (2) “non-abelian” quantum numbers, such as isotopic spin, that label families of particles.¹⁰

In 1935, Hideki Yukawa (1907–1981) published a paper in which he proposed a field theoretical model to account for the nuclear forces. In Yukawa’s theory, the neutron-proton force was mediated by the exchange of a scalar particle between the neutron and proton, with the mass of the scalar particle – called a meson – so adjusted as to yield a reasonable range for the nuclear forces. Yukawa had writ large what had been known in QED, namely that the electromagnetic force between charged particles could be conceptualized as arising from the exchange of “virtual” photons – called virtual because these photons did not obey the relation $E = hv$, which is valid for free photons. The masslessness of photons implies that the range of electromagnetic forces is infinite. In Yukawa’s theory, the exchanged quanta are massive, and the range, R , of the resulting interaction is related to the mass, μ , of the quanta by $R = h/\mu c$. This association of interactions with exchanges of quanta is a general feature of all quantum field theories.

Shortly after the Caltech cosmic ray physicists Carl Anderson (1905–1991) and Seth Neddermeyer (1907–1989) had given evidence for the existence of a new type of particle in the penetrating component of cosmic rays, Robert Oppenheimer (1904–1967) and Robert Serber (1909–1996) in 1937 published a short note in the *Physical Review* in which they pointed out that the mass of the newly discovered particle specified a length that they connected with the range of the nuclear forces, as had been suggested by Yukawa. Oppenheimer and Serber’s note was responsible for drawing the attention of American physicists to the meson theories of nuclear forces that Yukawa, Ernest Stückelberg (1905–1984), and Gregor Wentzel (1898–1978) had advanced. The existence of this “heavy electron” – which existed in both a positive and a negative variety – was authenticated by its direct observation in a cloud chamber by Curry Street (1906–1981) and Edward C. Stevenson

¹⁰ In 1953 Gell-Mann, and independently Nakano and Nijishima, proposed the property of matter called “strangeness.” The quantum numbers which are associated with operators that do not commute with the electric charge operators are called “non-abelian.” See M. Gell-Mann and Y. Ne’eman, *The Eightfold Way* (New York: Benjamin, 1964).

(b. 1907), who also determined its mass (150–220 electron masses) from measurements of the ionization it produced and from the curvature of its track in a magnetic field. Its lifetime was estimated to be about 10^{-6} sec. By 1939 Hans Bethe (b. 1906) could assert that “it was natural to identify these cosmic ray particles with the particles in Yukawa’s theory of nuclear forces.”¹¹

QED, Fermi’s theory of β -decay and Yukawa’s theory of nuclear forces established the model upon which all subsequent developments were based.¹² The model postulated new “impermanent” particles to account for interactions and assumed that relativistic QFT was the natural framework in which to attempt the representation of phenomena at ever smaller distances, that is, at higher and higher energies. It led to a description of nature in terms of a sequence of families of elementary constituents of matter with fewer and fewer members.

By the late 1930s, the formalism of quantum field theory was fairly well understood. However, it was recognized that all relativistic QFTs are beset by divergence difficulties that manifest themselves in perturbative calculations beyond the lowest order. These problems impeded progress throughout the 1930s, and most of the workers in the field doubted the correctness of QFT in view of these divergence difficulties. Numerous proposals to overcome these problems were advanced during the 1930s, but all ended in failure.¹³

The pessimism of the leaders of the discipline – Bohr, Pauli, Heisenberg, Dirac, Oppenheimer – was partly responsible for the lack of progress. They had witnessed the overthrow of the classical concepts of space-time and were responsible for the rejection of the classical concept of determinism in the description of atomic phenomena. They had brought about the quantum mechanical revolution, and they were convinced that only further conceptual revolutions would solve the divergence problem in quantum field theory.

Heisenberg in 1938 noted that the revolutions of special relativity and of quantum mechanics were associated with fundamental *dimensional* parameters: the speed of light, c , and Planck’s constant, h . These delineated the domain of classical physics. He proposed that the next revolution be associated with the introduction of a fundamental unit of length, which would delineate the domain in which the concept of fields and local interactions would be applicable.

The S-matrix theory, which Heisenberg developed in the early 1940s, was an attempt to make this approach concrete. He observed that all experiments can be viewed as scattering experiments. In the initial configuration, the system is prepared in a definite state. The system then evolves and the final

¹¹ Hans Bethe, “The Meson Theory of Nuclear Forces,” *Physical Review*, 57 (1940), 260–72.

¹² L. M. Brown, “How Yukawa Arrived at the Meson Theory,” *Progress of Theoretical Physics*, suppl. 85 (1985), 13–19; Olivier Darrigol, “The Origin of Quantized Matter Waves,” *Historical Studies in the Physical Sciences*, 16:2 (1986), 198–253.

¹³ Steven Weinberg, “The Search for Unity: Notes for a History of Quantum Field Theory,” *Daedalus* (Fall, 1977); Pais, *Inward Bound*.

configuration is observed after a time that is long compared with the characteristic times pertaining in the interactions. The S matrix is the operator that relates initial and final states. Its knowledge allows the computation of scattering cross-sections and other observable quantities. By again suggesting that only variables referring to experimentally ascertainable quantities should enter theoretical descriptions, Heisenberg opened a new chapter in the development of quantum field theories.¹⁴

FROM PIONS TO THE STANDARD MODEL: CONCEPTUAL DEVELOPMENTS IN PARTICLE PHYSICS

Modern particle physics can be said to have begun with the end of World War II. Peace and the Cold War ushered in an era of new accelerators of ever-increasing energy and intensity that were able artificially to produce the particles that populate the subnuclear world. Simultaneously, there developed the expertise to construct particle detectors of ever-increasing complexity and sensitivity that allowed the recording of the imprints of high energy subnuclear collisions. Challenges, opportunities, and resources attracted practitioners: The number of “high energy” physicists worldwide was to grow from a few hundred after World War II to some 8,000 in the early 1990s.

John Archibald Wheeler (b. 1911) summarized the state of affairs in elementary particle physics in the fall of 1945 by observing that the experimental and theoretical researches of the 1930s had made it possible to identify four fundamental interactions: (a) gravitation, (b) electromagnetism, (c) nuclear (strong) forces, and (d) weak-decay interactions. Wheeler believed that the interesting and exciting areas of research were the investigations of the electromagnetic, the strong, and the weak interactions, and these, indeed, became the traditional domain of high energy physics.¹⁵

Two important developments in 1947 shaped the further evolution of particle physics. Both were the result of intense discussions that followed experimental findings presented to the Shelter Island Conference. This was the first of three meetings sponsored by the National Academy of Sciences, which assembled the young American theorists who had made important contributions to the wartime weapons research in order to discuss foundational problems in physics. These conferences were the precursors of the (international) Rochester conferences begun in 1950 that brought together high energy physicists – experimentalists and theorists – biannually.

¹⁴ See Cushing, *Theory Construction*.

¹⁵ John A. Wheeler, “Problems and Prospects in Elementary Particle Research,” *Proceedings of the American Philosophical Society*, 90 (1946), 36–52.

At the 1947 Shelter Island Conference, the curious results that Marcello Conversi (b. 1917), Ettore Pancini (1915–1981), and Oreste Piccioni (b. 1915) had obtained regarding the decay of mesons observed at sea level led Robert Marshak (1916–1992) to formulate the “two-meson” hypothesis. He suggested that there existed two kinds of mesons. The heavier one, the π -meson, which was identified with the Yukawa meson responsible for the nuclear forces, is produced copiously in the upper atmosphere in nuclear collisions of cosmic ray particles with atmospheric atoms. The lighter one, the μ -meson observed at sea level, is the decay product of a π -meson and interacts but weakly with matter. A similar suggestion had been made earlier by Shoichi Sakata (1911–1970) in Japan.

Within a year, Cecil Powell (1903–1969), using nuclear emulsions sent aloft in high altitude balloons, corroborated the two-meson hypothesis by exhibiting $\pi \rightarrow \mu$ decays. During the early 1950s, the data pouring out of the plethora of π -meson-producing accelerators led to the rapid determination of the characteristic properties of the pion or π -meson, which occurs in three varieties: positively charged, negatively charged, and neutral.

The two-meson hypothesis also suggested that the list of the particles comprising the two distinct kinds of matter had to be amended. There were particles like the electron, the muon, and the neutrino that do not experience the strong nuclear forces; these were called leptons. Then there were the particles like the neutron, proton, and π -mesons that do interact strongly with one another and were named hadrons.

In January 1949, Jack Steinberger (b. 1921) gave evidence that the μ -meson decays into three light particles

$$\mu^+ \rightarrow e^+ + \nu + \bar{\nu}$$

and shortly thereafter, Gianpietro Puppi (b. 1917), Oskar Klein (1894–1977), Jaime Tiomno (b. 1924) and Wheeler, and Tsung Dao Lee (b. 1926), Marshall Rosenbluth (b. 1930), and Ning Yang (b. 1922) indicated that this process could be described by a Fermi-like interaction, as in the case of ordinary β -decay. Moreover, they pointed out that the coupling constant describing this interaction was of the same magnitude as the one occurring in nuclear beta decay. Thus, the pre-1947 period can be characterized as that of *classical beta decay*, while the postwar period initiated the modern period of *universal Fermi interactions*.

The second important development in the immediate post–World War II period was a theoretical advance. It stemmed from the attempt to explain quantitatively the discrepancies between the empirical data and the predictions of the relativistic Dirac equation for the level structure of the hydrogen atom and the value it ascribed to the magnetic moment of the electron. These deviations had been observed in reliable and precise molecular beam experiments carried out by Willis Lamb (b. 1913), and by Isidor Rabi (1898–1988)

and co-workers at Columbia, and were reported at the Shelter Island Conference. Shortly after the conference, it was shown by Bethe that the Lamb shift (the deviation of the $2s$ and $2p$ levels of hydrogen from the values given by the Dirac equation) was of quantum electrodynamical origin, and that the effect could be computed by making use of what became known as “mass renormalization,” an idea that had been put forward by Hendrik Kramers (1894–1952).

The parameters for the mass, m_0 , and for the charge, e_0 , that appear in the Lagrangian-defining QED, are not the observed charge and mass of an electron. The observed mass, m , of the electron is to be introduced in the theory by the requirement that the energy of the physical state corresponding to an electron moving with momentum p be equal to $(p^2 + m^2)^{1/2}$. Similarly, the observed charge should be defined by the requirement that the force between two electrons (at rest), separated by a large distance r , be described by Coulomb’s law, e^2/r^2 , with e the observed charge of an electron.

It was shown by Julian Schwinger (1918–1994) and by Richard Feynman (1918–1988) that the divergences encountered in the low orders of perturbation theory could be eliminated by reexpressing the parameters m_0 and e_0 in terms of the observed values m and e , a procedure that became known as mass and charge renormalization. In 1948, Freeman Dyson (b. 1923), working at the Institute for Advanced Study in Princeton, was able to show that charge and mass renormalization were sufficient to absorb all the divergences of the scattering matrix (S-matrix) in QED to all orders of perturbation theory. More generally, Dyson demonstrated that only for certain kinds of quantum field theories is it possible to absorb *all* the infinities by redefinition of a *finite* number of parameters. He called such theories renormalizable. Renormalizability thereafter became a criterion for theory selection.¹⁶

The ideas of mass and charge renormalization, implemented through a judicious exploitation of the symmetry properties of QED – that is, the Lorentz invariance and the gauge invariance of the theory – made it possible to formulate and to give physical justifications for algorithmic rules to eliminate all the ultraviolet divergences that had plagued the theory and to secure unique finite answers. The success of renormalized QED in accounting for the Lamb shift, the anomalous magnetic moment of the electron and of the muon, and the radiative corrections to the scattering of photons by electrons, to pair production, and to bremsstrahlung, was spectacular.

Perhaps the most important theoretical accomplishment of the 1947–52 period was providing a firm foundation for believing that local quantum field theory was the framework best suited for the unification of quantum theory and special relativity. The most perspicacious theorists, for example, Murray Gell-Mann (b. 1929), also noted the ease with which symmetries – both space-time and internal symmetries – could be incorporated into the

¹⁶ Schweber, *QED*.

framework of local quantum field theory. Local QFTs were thus advanced for the description of the “elementary particles” and their internal symmetries. Photons, pions, nucleons, electrons, muons, and neutrinos (the elementary particles as perceived in the early fifties) corresponded to localized excitations of the underlying, “fundamental” local fields.

Although experiments with cosmic rays during the 1940s and 1950s had indicated the presence of new “strange” particles, high energy physics during most of the 1950s was dominated by pion physics. The success of QED rested on the validity of perturbative expansions in powers of the coupling constant, $e^2/\hbar c$, which is small, $\approx 1/137$. However, the pseudoscalar meson theory of the pion-nucleon interaction required the coupling constant to be large – of the order of 15 – for the theory to yield nuclear potentials that would bind the deuteron. No valid method was found to deal with such strong couplings. It also became clear that meson theories were woefully inadequate to account for the properties of all the new hadrons being discovered. The importance of the *tempo* of new experimental findings by the particle accelerators that were coming on-line cannot be overemphasized. The plethora of new experimental discoveries quelled any hope for a rapid, neat, and systematic transition from QED to the formulation of a dynamics for the strong interaction.

To Sam Treiman (1925–1999), an important contributor to the development in particle physics from the 1950s to the 1980s and the teacher and mentor at Princeton University of many of the best young theorists coming of age during that period, “the prospect of finding the *right* quantum field theory, if indeed there were a right one, or even recognizing it if it were presented seemed remote [from 1955 to 1965].”¹⁷

Thus, at the end of the 1950s, QFT faced a crisis because of its inability to describe the strong interactions and the impossibility of solving any of the realistic models that had been proposed to explain the dynamics of hadrons. Efforts to develop a theory of the strong interactions along the model of QED were generally abandoned, although a local gauge theory of isotopic spin symmetry, advanced by Yang and Robert Mills (b. 1927) in 1954, was to prove influential later on.

There were several responses to the crisis that developed in theoretical particle physics at the end of the 1950s. For some theorists, the failure of quantum field theory and the superabundance of experimental results was, in fact, emancipating. It led to explorations of the generic properties of QFT when only such general principles as causality, the conservation of probability (unitarity), and relativistic invariance are invoked and no specific assumptions are made regarding the form of the interactions.

Geoffrey Chew’s (b. 1924) S-matrix program, which rejected QFT and attempted to formulate a theory that made use only of observables embodied

¹⁷ S. Treiman, “A Life in Particle Physics,” *Annual Reviews of Nuclear and Particle Science*, 46 (1996), 1–30, at p. 6.

in the S-matrix, was more radical. Physical consequences were to be extracted without recourse to any dynamical field equations, by making use of general properties of the S-matrix, such as unitarity and Lorentz invariance, and certain assumptions (analyticity) regarding its dependence on the variables describing the initial and final energies and momenta of the particles involved.¹⁸

Another response to the crisis was to make symmetry concepts central. Symmetry considerations were first applied to the weak and the electromagnetic interactions of the hadrons, and they were later extended to encompass low energy strong interactions. Phenomenologically, the strong interactions seemed to be well modeled by (effective) Hamiltonians, the physical variables of which were hadron *current* operators. No dynamical assumptions were made on how these hadron current operators were to be constructed from hadron field operators, but commutation relations were imposed on them, reflecting the underlying symmetries that were assumed to be independent of dynamic details and to be universally valid. These symmetries and their group structure were derived from the exact or approximate regularities that manifested themselves in the experimental data. This research program, known as current algebra, took shape during the late 1950s and early 1960s. The program reached its limit around 1967 because some of its predictions were in direct conflict with experiments.

In fact, during the 1950s and 1960s, progress in classifying and understanding the phenomenology of the ever-increasing number of hadrons was not made by virtue of a fundamental theory. It was accomplished by shunning dynamical assumptions and, instead, making use of symmetry principles (and their associated group theoretical methods) and exploiting kinematical principles that embodied the essential features of a relativistic quantum mechanical description.

Symmetry thus became one of the fundamental concepts of modern particle physics. It is used both as a classificatory and organizing tool and as a foundational principle to describe dynamics. The notion of symmetry was enriched by two developments in the second half of the 1950s: (1) the realization by Lee and Yang that parity is not conserved in the weak interactions; and (2) the extension by Yang and Mills in 1955 of the global isotopic spin symmetry of nucleons to a local symmetry, in analogy with gauge invariance in QED.

This local symmetry, or local gauge invariance, demands that the photon be massless. The requirements of relativistic invariance, gauge invariance, and the absence of dimensionality in the coupling constant scaling the strength of the interaction determine the form of the Lagrangian describing the interaction between charged particle fields and the electromagnetic field.

¹⁸ Cushing, *Theory Construction*.

A Lagrangian that is invariant under some global transformation of the form

$$\psi(x) \rightarrow e^{ie\Lambda} \psi(x)$$

with Λ constant, can be made locally invariant under such a transformation, that is, with $\Lambda = \Lambda(x)$, by introducing appropriate gauge fields. Yang and Mills made use of this observation to introduce a gauge theory of the strong interactions, by extending to a local symmetry the invariance of the nucleon fields under global isotopic rotations:

$$\psi(x) \rightarrow e^{ig\tau\cdot\phi} \psi(x)$$

Local gauge invariance, however, implies that the gauge bosons are massless. This is not the case for the pion, and thus Yang and Mills's theory was considered an interesting model but without relevance for understanding the strong interactions.

Interest in field theories, and in particular in gauge theories, was revived after the notion of spontaneous symmetry breaking (SSB) became fully appreciated in the early 1960s. Jeffrey Goldstone (b. 1933) and Yoichiro Nambu (b. 1921) noted that in quantum field theories, symmetries could be realized differently: It was possible to have the Lagrangian invariant under some symmetry, yet have this symmetry not respected by the vacuum (that is, by the ground state of the theory). Such symmetries are known as spontaneously broken (SBS). It turns out that if the symmetry that is spontaneously broken is a global one, there will be massless (Goldstone) spin zero bosons in the theory. If the (broken) symmetry is a local gauge symmetry, then the Goldstone bosons disappear from the theory, but each of the gauge bosons associated with broken symmetries acquires a mass. This is the Higgs mechanism.¹⁹

In 1967 Steven Weinberg (b. 1933), and somewhat later in 1968, Abdus Salam (1926–1996) independently proposed a gauge theory of the weak interactions that unified the electromagnetic and the weak interactions and made use of the Higgs mechanism. Their model incorporated previous suggestions that Sheldon Glashow had advanced in 1961 on how to formulate a gauge theory of the weak interactions, in which the weak forces were mediated by gauge bosons. The original Glashow theory had been set aside because the consistency of gauge theories with *massive* gauge bosons was doubted, and by the fact that such theories were nonrenormalizable.

SBS offered the possibility of giving masses to the gauge bosons, but whether such theories with spontaneously broken symmetries via a Higgs mechanism were renormalizable was not known. The renormalizability of

¹⁹ For an overview of the mechanisms which implement the broken symmetry, see the presentation by L. M. Brown and the subsequent discussion in Hoddeson et al., *Birth of Particle Physics*, pp. 478–522.

such theories was proved by Gerard 't Hooft (b. 1946) in his dissertation in 1972 at Utrecht University under the supervision of Martinus Veltman (b. 1931). The status of the Glashow-Weinberg-Salam theory changed dramatically thereafter. As Sidney Coleman noted, "'t Hooft's kiss transformed Weinberg's frog into an enchanted prince."²⁰

Gauge theory, the mathematical framework for generating dynamics incorporating symmetries into a QFT, has played a crucial role in the further development of QFT. It can rightly be said that symmetry, gauge theories, and spontaneous symmetry breaking have been the three pegs upon which modern particle physics rests.

QUARKS

All the phenomenological theorizing of the 1960s led to the view that the elementary constituents of matter at the smallest distances, or equivalently at the highest energies, are quarks and leptons. In 1961, Gell-Mann and Yuval Ne'eman (b. 1925) independently proposed classifying the hadrons into families on the basis of a symmetry that became known as the "eightfold way." They realized that the mesons grouped naturally into octets, the baryons into octets, and decuplets. The mathematical expression of the "eightfold way" symmetry was the group of (unitary) transformations $SU(3)$, the generalization to hadrons of the symmetry group $SU(2)$ that had been used to express mathematically the charge independence of the nuclear forces between neutrons and protons. A fundamental representation of $SU(3)$ is three-dimensional, which led Gell-Mann, and independently George Zweig (b. 1937), to suggest that hadrons were composed of three elementary constituents, which Gell-Mann named quarks (from a passage in James Joyce's *Finnigan's Wake*: "Three quarks for Master Mark!") and Zweig called aces.

To account for the observed spectrum of hadrons, Gell-Mann and Zweig assumed that there were three "flavors" of quarks (generically indicated by q), called up (u), down (d), and strange (s), that had spin $1/2$, isotopic spin $1/2$ for the u and d and isotopic spin 0 for the s , and strangeness 0 for the u and d and -1 for the s quark. Ordinary matter contains only u and d quarks; "strange" hadrons contain strange quarks or antiquarks. The three quarks were to carry baryonic charge of $1/3$ and an electrical charge that is $2/3$ for the u , and $-1/3$ for the d and s , that of the proton's charge.²¹

This was a rather startling assumption since there is no experimental evidence for any macroscopic object carrying a positive charge smaller than that of a proton or a negative charge smaller than an electron. Since a relativistic quantum mechanical description implies that for every charged particle

²⁰ See Weinberg, *Quantum Theory of Fields*.

²¹ M. Gell-Mann and Y. Ne'eman, *The Eightfold Way*; and Gell-Mann, "Quarks, Color and QCD," in P. M. Zerwas and H. A. Kastrup, *QCD 20 Years Later* (Singapore: World Scientific, 1993), pp. 3–15.

there exists an “antiparticle” with the opposite charge, it was assumed that there are likewise antiquarks (generically denoted by \bar{q}), having the opposite electric charge and opposite sign of strangeness. Quarks were assumed to interact with one another and to form bound states, giving rise to the observed hadrons. Thus a π^+ meson was assumed to be a bound state of an up and antidown quark. Similarly, a proton was “made up” of two up quarks (that contributed $4/3 e$ to the electrical charge) and a down quark (of electrical charge $-1/3 e$), giving rise to an entity with an electrical charge of $+1 e$. In fact, all baryons could be made up of three quarks, all mesons with one quark and one antiquark.

However, in order to satisfy the Pauli principle in a structure like the Ω^- , which is presumably constituted of three *identical* spin $1/2$ strange quarks, all in s states, quarks had to be given a new attribute, a new form of charge, called color, in order to distinguish otherwise identical quarks. Color is a “three-dimensional” analog of electric charge: It occurs in three varieties (sometimes taken to be red, yellow, and blue). Thus, there are positive and negative red, yellow, and blue colors. Quarks carry positive color charges and antiquarks carry the corresponding negative charge. The observed hadrons are required to be color singlets, that is, to have zero net color charge.²²

If the $SU(3)$ symmetry were exact, all the quarks, and all the baryons in a given octet or decuplet, would have the same mass. Since they do not, the symmetry must be broken; this comes about by virtue of the three flavors of quarks having different masses, with the s quark assumed to have a greater mass than the u and the d quarks.

An entire phenomenology grew out of this classificatory scheme. In the early 1960s, the flavor $SU(3)$ quark model, in which the u, d, and s quarks are considered the building blocks, could classify all the then-known hadrons into three families: an octet of spin 0 mesons (that included the ρ and K mesons); an octet of spin $1/2$ baryons (that included the neutron and the proton, the Λ and the Σ), and a decuplet of spin $3/2$ baryons.²³

In the late 1960s, experiments in which high energy electrons were inelastically scattered off protons were carried out at the Stanford Linear Accelerator (SLAC).²⁴ Since the early 1950s, it had been known that protons had an internal structure. By 1968 electrons were being accelerated at SLAC to 20 GeV, at which energy their wavelength was such that they could resolve entities appreciably smaller than the size of the proton. Such electrons were thus ideal probes for investigating the internal structure of the proton. If charge were uniformly distributed within the proton, high energy electrons would

²² See O. W. Greenberg, “Color: From Baryon Spectroscopy to QCD,” in M. Gai, ed., *Baryons '92: International Conference on the Structure of Baryons and Related Mesons* (Singapore: World Scientific, 1993), pp. 130–9.

²³ J. J. J. Kokkedee, *The Quark Model* (New York: Benjamin, 1969); K. Gottfried and V. Weisskopf, *Concepts of Particle Physics* (New York: Oxford University Press, 1986).

²⁴ See the presentations by Jerome Friedman and James Bjorken in Hoddeson et al., *Birth of Particle Physics*, pp. 566–600.

tend to go through the proton without being appreciably deflected. If on the other hand – in analogy with Rutherford’s interpretation of Geiger and Marsden’s experiment on the scattering of α -particles by gold atoms – the charge within the proton were localized on internal constituents, then an electron, if it were to pass close to one of these concentrations of charge, would be strongly deflected.

Just such large-angle scatterings were observed at SLAC. Upon hearing these experimental findings, Feynman suggested that the proton is composed of pointlike particles that he called “partons,” and from the angular distribution of the scattered electrons he inferred that partons had spin $1/2$. The partons were soon recognized as identical with the quarks of Gell-Mann’s and Zweig’s model. There were, however, paradoxical aspects with this identification of the proton constituents. First, the partons/quarks appeared to be very light, much less than one-third of the mass of the proton; and second, they appeared to move almost freely inside the proton – difficulties that were addressed and resolved only later.

The discovery in November 1974 of the J/ψ , a spin 1 meson, gave further evidence for the correctness of the quark picture and gave credence to the existence of a fourth quark with a new flavor, called charm. (The charmed quark was denoted by c). The existence of such a quark had been suggested by James Bjorken and Glashow in 1964, and the proposal had been elaborated further by Glashow, John Iliopoulos (b. 1940) and Luciano Maiani (b. 1941) in 1970. It was immediately conjectured that the J/ψ was a bound state of a c and \bar{c} . The subsequent detection of the ψ' , a “particle” related to the J/ψ by its decay, made the notion of quarks in general, and of charmed quarks in particular, compelling. The discoveries of November 1974 revolutionized high energy physics. With the November revolution, the conceptualization of hadrons as

quark composites was put beyond dispute, and gauge theory received a tremendous boost – the Weinberg-Salam plus Glashow-Iliopoulos-Miaini Model became the basis of a new hadron spectroscopy. At the heart of these developments was charm. . . . The triumph of charm was simultaneously a triumph for gauge theory.²⁵

With the discovery of the charmed quark, and subsequently of the third family of particles – the τ lepton and its neutrino – and of the “bottom” (or “beauty”) b quark in 1977, and of the “top” (t) quark in 1994, six different “flavors” of quarks were needed to account for the observed hadron spectroscopy. Each successively discovered quark is more massive than its predecessors: The u and d quarks have (effective) masses of 5 and 10 Mev/c^2 , respectively; the s an (effective) mass of 180 Mev/c^2 ; the c has a mass 1.6 Gev/c^2 ; the b of 4.8 Gev/c^2 and the t of 174 Gev/c^2 . All are spin $1/2$ particles

²⁵ Pickering, *Constructing Quarks*, p. 254.

that partake in the strong, electromagnetic, and weak interactions, and all come in pairs: up and down (u, d), charm and strange (c, s), and top and bottom (t, b). The first member of each pair has electric charge $2/3$ and the second $-1/3$. Each flavor comes in three colors.

From the time they were introduced as “hypothetical” particles, an important problem connected with quarks loomed large: If indeed all hadrons are made up of fractionally charged quarks, why is it that one does not eventually reach an energy high enough to liberate the constituent quarks in a collision process and thus allow a fractionally charged hadron to be observed? This is the so-called confinement problem. And even were one able to provide a mechanism that accounts for the confinement of quarks, what meaning is to be attached to the reality of quarks as constituents of hadrons if they can never be observed empirically?

GAUGE THEORIES AND THE STANDARD MODEL

As currently described, a common mechanism underlies the strong, weak, and electromagnetic interactions. Each is mediated by the exchange of a spin 1 gauge boson. In the case of the strong interactions, the gauge bosons are called gluons; in the case of the weak interactions, W^\pm and Z bosons; and in the electromagnetic case, photons. A general chromatic terminology has become popular, and one often refers to the charges as “colors.” Thus, one speaks of QED – the paradigmatic gauge theory – as a theory of a single gauge boson, the photon, coupled to a single “color,” namely, the electric charge. The gauge bosons of the strong interactions carry a 3-valued color; those mediating the weak interactions carry a “two-dimensional” weak color charge. Weak gauge bosons interact with quarks and leptons, and in the act of being emitted or absorbed, some of them can transform one kind of quark or lepton into another. When these gauge bosons are exchanged between leptons and quarks, they are responsible for the force between them. They can also be emitted as radiation when the quarks or leptons are accelerated.

Quantum chromodynamics (QCD) describes the strong interactions between the six quarks. Quarks carry electrical charge and, in addition, carry a (“three-dimensional”) strong color charge. Each of the six quarks carries this color charge and can be in any of three colors states. QCD is a gauge theory with three colors and involves eight massless gluons, the color-carrying gauge bosons, six that alter color, and two that merely react to them. QCD possesses a gauge invariance: The theory is invariant under the addition to the gluon field potentials of a set of gradients and a simultaneous change of the phases of the quark fields. A quark’s color is changed when it absorbs or emits a color-changing gluon. However, a quark’s flavor is not changed by the absorption or emission of a gluon – nor by the emission or absorption of a photon.

The GWS (Glashow-Weinberg-Salam) gauge theory of the weak interactions is a gauge theory involving two colors. Each of the quarks thus carries an additional weak color (or weak charge). There are four gauge bosons that mediate the weak interactions between the quarks. Three of them (the W^+ , W^- , and W^0) change the flavor of the quark when absorbed or emitted; the fourth, the B^0 boson, reacts but does not alter the weak color charges.

As just described, the standard model, although aesthetically beautiful, does not accord with the known characteristics of the weak interactions nor with the properties of quarks as envisaged in their phenomenological descriptions. Local gauge invariance requires that the gauge bosons be massless and, therefore, that the range of the forces they generate be long range. Yet it is known that the weak force is of very short range (less than 10^{-16} cm), and that the mass of the W boson is 80 GeV and that of the Z , 91 GeV. Nor can it accommodate the masses of the quarks. A Higgs mechanism for spontaneously breaking symmetries – accomplished by introducing a (complex) doublet of scalar fields – is the mechanism most commonly invoked to overcome these difficulties. Establishing the reality of such Higgs particles became an important reason for justifying the building of the Superconducting Super Collider (SSC).²⁶

The past two decades have seen a large number of successful explanations of high energy phenomena using QCD. The substantiation in 1973 at the Centre Européen pour la Recherche Nucléaire (CERN) of the process $\nu_\mu + e^- \rightarrow \nu_\mu + e^-$ corroborated the existence of weak neutral currents as embodied in the Glashow-Weinberg-Salam electroweak theory. The detection and identification of the W^\pm and of the Z_0 in 1983 by Carlo Rubbia and co-workers at CERN gave further important confirmation of that theory. Similarly, the empirical data obtained in lepton and photon deep inelastic scattering, and in the study of jets in high energy collisions can be accounted for quantitatively by QCD. Furthermore, computer simulations have presented convincing evidence that QCD does produce quark and gluon confinement inside hadrons.²⁷ Frank Wilczek, one of the important contributors to the field, in his opening remarks at a conference in 1992 devoted to an assessment of QCD since its initial formulation, could assert that “QCD is now a mature theory, and it is now possible to begin to view its place in the conceptual universe with appropriate perspective.”²⁸

The empirical data that can be accounted for quantitatively are indeed impressive.²⁹ As Guido Altarelli remarked in his review of “QCD and

²⁶ Daniel J. Kevles, “Preface 1995: The Death of the Superconducting Super Collider in the Life of American Physics,” in *The Physicists: The History of a Scientific Community in Modern America*, 2d ed. (Cambridge, Mass.: Harvard University Press, 1995).

²⁷ M. Creutz, *Quarks, Gluons and Lattices* (Cambridge: Cambridge University Press, 1983).

²⁸ Wilczek in P. M. Zerwas and H. A. Kastrup, eds., *QCD 20 Years Later* (Singapore: World Scientific, 1993), p. 16.

²⁹ See Frank Wilczek, “The Future of Particle Physics as a Natural Science,” in *Critical Problems in Physics*, ed. V. A. Fitch, D. R. Marlow, and M. A. E. Dementi (Princeton, N.J.: Princeton University Press, 1997), pp. 281–308.

Experiment” at that same conference:

[Since the late 80s] [m]any relevant calculations, often of unprecedented complexity, have been performed. As a result of 3 years of really remarkable progress, our confidence in QCD has been further consolidated, . . . and a lot of additional checks from many different processes have become possible.³⁰

QCD is the accepted framework for describing the interactions of leptons, quarks and gluons below 1 TeV.

The standard model is one of the great achievements of the human intellect. It will be remembered – together with general relativity, quantum mechanics, and the unraveling of the genetic code – as one of the outstanding intellectual advances of the twentieth century. But the standard model is not the “final theory,” for too many parameters that have to be empirically determined enter the description, for example, the masses of the quarks and the various coupling constants.

Very shortly after the realization that the strong and the electroweak interactions could be described by gauge theories that had similar mathematical structures, a new phase in the unification of the different forces of nature began. The similarity between the transformation properties of gluon and quark fields under the (three) color gauge transformations and those of the quark and lepton fields under the (two) weak color gauge transformations immediately suggested the possibility that a larger (five-dimensional) gauge group, $SU(5)$, might encompass both the strong and the electroweak interactions. Howard Georgi (b. 1947) and Glashow advanced such a grand unified (gauge) theory (GUT) as soon as QCD was recognized as the likely theory of the strong interactions.³¹

The greatest immediate impact of GUTs has been in cosmology and in the description of the physics of the early universe. Asymptotic freedom implies that matter at extreme temperatures and densities becomes weakly interacting and, therefore, that its equation of state is rather simply calculable. GUTs made it possible to calculate the consequences of various unification scenarios for cosmology with some confidence. It also offered an explanation for how the observed asymmetry between matter and antimatter could have developed from a symmetric starting condition. In fact, probably the most consequential unification during the past twenty years has been the “unification” of particle physics and astrophysics. The early universe, the immediate aftermath of the big bang, has become the laboratory in which to explore the implications of foundational theories (such as GUTs and string theory) at temperatures and energies that are and will remain inaccessible in terrestrial laboratories.

³⁰ Altarelli in Zerwas and Kastrup, *QCD 20 Years Later*.

³¹ H. Georgi and S. L. Glashow, “Unified Theory of Elementary Particle Forces,” *Physics Today*, 33, no. 9 (1980), 30–9; and S. Dimopoulos, S. A. Raby, and F. Wilczek, “Unification of Couplings,” *Physics Today*, 44, no. 10 (1991), 25–33.