SLAC-PUB-2043
NOVEMBER 1977
(T/E)

ELEMENTARY PARTICLE PHYSICS*

S. D. Drell

Stanford Linear Accelerator Center, Stanford University, Stanford, California 94305

Reprinted from DAEDALUS
Summer 1977: Discoveries & Interpretations: Studies in Contemporary Scholarship, Volume 1. Issued as Volume 106, No. 3 of the Proceedings of the American Academy of Arts & Sciences.

*Work supported by the Department of Energy.

THE ELUSIVE QUARK has emerged during the past decade as the newest candidate for the most elementary constituent, or ultimate building block, of Nature. For example the proton, which is the nucleus of the hydrogen atom, is believed to consist of three quarks bound together by a strong force. However, not a single quark has ever been detected in the laboratory in spite of many efforts to find one. In addition to lacking proof of their existence by a direct laboratory observation, we also lack a fundamental theory of quarks. Among the unanswered questions are: How many different kinds of quarks are there? Do they have an internal structure of their own? What forces bind them to one another to form protons and the other observed subnuclear structures? Can these bonds ever be broken so that isolated quarks can be observed?

In spite of this profound and extensive state of uncertainty the quark concept has proved to be very useful in our search for an underlying unity to the diverse phenomena observed in today's world of elementary particle physics as we probe in the laboratory on scales of distance that are tens to hundreds of millions times smaller than the size of what we now call the "atom," and on scales of time that are shorter than one hundredth of one millionth of typical atomic time scales. And along with the quarks, new ideas of elementary particle structure have been proposed. Foremost among these is the possibility that the individual quarks, although viewed as elementary constituents, have not been seen so far because, in principle, they may never be isolated one at a time. By turning to the concept of a quark thus shrouded in mystery in its search for basic building blocks, modern science appears to be turning its clocks back to the time of the Greek philosophers of 2500 years ago, and turning its back on the traditions and methods of more than five centuries, which have rooted nature's laws in observable concepts.

The newest ideas on the frontiers of the submicroscopic unknown and the radical change in our concepts from the early Greeks to today's quarks are the subject of this essay. This is an especially hazardous time for such an undertaking. We have just passed through two highly explosive years of progress in the field of elementary particle physics and the new discoveries not only have not been fully understood; they are still occurring. With detectors of improved sensitivity and with facilities of greater resolving power new frontiers of the unknown are being explored. Perhaps individual quarks will be observed directly, perhaps not. Perhaps new species of quarks will have to be introduced, perhaps not. No matter what the results, the recent momentous experimental discoveries and theoretical successes have generated considerable optimism that quarks represent a major stride forward in our understanding of the basic structure of nature.

If, in the future, individual quarks are observed directly in the laboratory they will assume a respected and more traditional role as the next hierarchy of elementary building blocks. If, on the contrary, individual quarks remain unobservable as basic building blocks of nature, no matter how hard we look (a result predicted by some of the most attractive current theoretical hypotheses) our view of submicroscopic nature will undergo a major revolution. The last such major revolution occurred earlier this century with the formulation of the quantum theory of atomic phenomena in 1925–1926.

"What Are We Made Of?"

Since the beginning of recorded history man has asked, "What are we made of?" Science first flourished twenty-five hundred years ago with the search of the early Greek philosophers for the answer to this question. Their search for an underlying unity to the rich diversity observed in the world around them led beyond appearances, for, as Aristotle wrote a century later: "So we must advance from the concrete whole to the several constituents which it embraces; for it is the concrete whole that is the more readily cognizable by the senses. And by calling the concrete a 'whole' I mean that it embraces in a single complex a diversity of constituent elements, factors, or properties."[1] Once beyond appearances, different Greek philosophers moved in various directions in postulating the invisible constituents or qualities out of which all else is constructed. According to Democritus, who was the most important founder of the atomic theory, the world is built of *indivisible* constituents, or *atoms*. The atoms are simpler than all the rich, varied phenomena we see around us, but by their motion and behavior they, though not directly observable, control all we do see. (In fact Democritus wrote that, apart from free atoms and empty space, all else in the world was opinions![2]) The realization that the search for nature's fundamental building blocks and basic processes carries us beyond the sense world of appearance into an atomic realm was a major advance in man's view of nature.

This search for a unity underlying the diversity observed in nature was to the early Greeks a purely theoretical exercise of the intellect. They did not insist on seeing their atoms. It was a matter of debate whether ultimate atoms existed in the mold of Democritus, or whether matter could be continually subdivided, as envisioned by Anaxagoras, into seeds within seeds within seeds ad infinitum.[3] However, modern science recognizes that logic alone cannot answer which, if indeed either, of these views is consistent with nature. Above and beyond the metaphysician, experiments and data are required; and, beyond the data themselves, laws with predictive power and subject to experimental testing must be constructed by a process of abstraction and generalization. In the words of Niels Bohr, one of the most important architects of modern atomic theory, "Only by experience itself do we come to recognize those laws which grant us a comprehensive view of the diversity of phenomena."[4] This is the scientific process as developed and applied for the past five centuries.

Atomic Measurement

In contrast to the early Greeks and in accord with the words of Bohr, modern science has insisted on detecting the so-called elementary particles which are the building blocks of matter. The particles' traces are identified from the tracks and signals which they form in specifically designed detectors. Early in this century the remarkable art of experimentation developed to the extent that it was possible to study the properties of individual atoms. As a result of this fantastic sensitivity of measurement our entire concept of observation underwent revolutionary changes. On the atomic frontier the most profound and radical change occurred when it was realized that it is necessary to take into account the effect of the observation itself on the physical system being observed. The interaction between observer and object can be determined with very high accuracy or even ignored as negligibly small for measurements on the macroscopic world. This is not so, however, for observations of individual atoms.

A recognition of the measurement process itself and an understanding of its inevitable and not entirely predictable effect on what is observed formed the foundations of the modern quantum theory. As Heisenberg, the father of the quantum theory, has written:

> . . . the interaction between observer and object causes uncontrollable and large changes in the system being observed, because of the discontinuous changes characteristic of atomic processes. The immediate consequence of this circumstance is that in general every experiment performed to determine some numerical quantity renders the knowledge of others illusory, since the uncontrollable perturbation of the observed system alters the values of previously determined quantities. If this perturbation be followed in its quantitative details, it appears that in many cases it is impossible to obtain an exact determination of the simultaneous values of two variables, but rather that there is a lower limit to the accuracy with which they can be known.[5]

Along with this recognition of a fundamental limitation in principle to the accuracy with which measurements of atomic phenomena can be made, it became necessary to introduce probability notions and to resort to wave as well as particle concepts to describe the observed properties of electrons and atoms. Classical particle concepts and causal connections from the macroscopic world of experience proved inadequate to describe the observed properties and interactions of electrons and of atoms. As a result of fundamental limitations of the measurement process, a major revolution occurred in man's concept of the elementary particle which could no longer be described by concepts of classical mechanics alone. But there is no uncertainty in what is meant when we say that we *observe* an electron as an elementary constituent of the atom.

Now, fifty years later, with the vastly increased sensitivity of the new particle detectors and with the vastly increased resolving power of the great accelerators, we have advanced our elementary particle frontier onto a scale of distances a million and more times smaller than the atomic scale, and we have come upon the quarks. And once again we are facing a major revolution in our concepts. This time we encounter very strong forces and seemingly unbreakable bonds acting among the individual quarks, leading us to question the possibility of ever detecting them in the laboratory. Are they objects whose existence can be inferred only from the properties of larger, complex structures, such as a proton, in which they are the constituents? Can we never study quarks and their structure by isolating them and probing them individually, as we do the electrons and nuclei of the atom, or the protons and neutrons that form complex nuclei? If this situation persists, physicists will be forced to question what is meant by the concept of elementary particles or basic building blocks. Unseeable and elusive elementary particles may even have to be replaced by an elementary "something else" that can be more directly related to laboratory measurements: elementary currents or symmetries, perhaps.

Why Quarks

The quark hypothesis was introduced in 1963 independently by Murray Gell-Mann and by George Zweig in an effort to summarize and systematize the great proliferation of nuclear particles that were being produced by accelerators on the high energy frontiers of the 1950s. Regularities were perceived in the masses of these particles as well as in the characteristics of their creation, their interactions, and their decay, for such was the fate of all but the proton itself in the subnuclear zoo that burgeoned to more than one hundred different inmates. Gell-Mann and Zweig showed that these regularities, as well as new ones found later, could be accounted for in terms of the simple motions and interactions of just three different kinds of quarks.

Because the quark hypothesis made correct predictions of new observations as well as provided a systematic organization of a large mass of data, and because it also brought simplicity plus a unifying harmony to our view of nature, the concept of quarks was a crucial step forward more than a decade ago, similar in many ways to the discovery of the nuclear atom by Ernest Rutherford in 1911. The periodic table of the elements, which had been developed around 1870 by Dmitri Mendeleev and shortly thereafter by Lothar Meyer, provided a systematic basis for modern chemistry. It also led to the discovery of several new elements whose existence was predicted on the basis of missing entries in the table. The first to be discovered was gallium in 1875. Furthermore, electrical properties of the atom were also being established around the turn of the century, but there was as yet no hypothesis or model of the constituents or of the structure of the atom itself. It was in the celebrated experiments of Rutherford that we first caught a glimpse of the picture of the atom that we are so familiar with today: a small, compact nucleus containing almost the entire mass of the atom, around which very light electrons circulate in orbits much in the same way as the planets orbit our sun.

Even before Rutherford's discovery the atom of the nineteenth and very early twentieth centuries was no longer viewed as the indivisible building block of all matter, as implied by its Greek name. However, its constituents and structure were still unknown prior to 1911. In particular it had been characterized as "plum pudding" by J. J. Thomson, who discovered the electron.

The complete confirmation of this picture of the nuclear atom was achieved in the decade following Rutherford's discovery, as physicists broke apart individual atoms and studied the fragments (i.e., electrons and ions). Rutherford was able to infer the existence of a very small, hard, and massive nucleus because the projectiles with which he bombarded thin foils of matter deflected through large angles, and some even bounded backward, much as does the cue ball in a break shot in pocket billiards. Knowing the nature of the electrical force between his projectiles and the target atom (in point of fact he assumed that the classical laws of electricity applied to atoms, but experimental confirmation of this came only later), he could quantitatively interpret his observed scattering pattern in terms of the atom's shape and structure. Later when these very same atomic constituents were also identified experimentally in the atomic debris as electrons and nuclei, the atomic model was confirmed. It was both understandable and successful. By 1922 the evidence had all been pieced together and the rules for atomic spectroscopy had been given in terms of the electronic structure of the atom. Quickly thereafter a grand synthesis was achieved with the development of the quantum theory, spearheaded by the works of W. Heisenberg, E. Schrödinger, M. Born, and others, starting in 1925. What emerged was a complete dynamical theory based on simple, elegant, and general physics laws; that is, mathematical equations plus unique, definite rules of physical interpretation including the understanding of the limitations on measurements made on the atomic scale.

Once the implications of a finite quantum of action for the observation and interpretation of atomic phenomena were fully acknowledged and understood, all the mysteries of atomic spectroscopy were stripped away. For atomic phenomena physicists could interpret known experimental results and predict new ones correctly from the same general laws and at the same time describe atomic behavior in a simple conceptual framework.

What we are trying to do now on the subnuclear frontier of elementary particle physics is simply to reproduce this triumph of fifty years ago. Whether or not we will succeed without a major conceptual upheaval is anybody's guess. The picture of electrons and nuclei as the basic constituents for describing all atomic properties and spectroscopic phenomena was built on the radically new conceptual foundations of the quantum theory that were developed in the 1920s. We are still applying these very same concepts today because experience has not yet forced us to abandon or modify them. No new principles were required as physics pushed ahead in the 1930s and 1940s to study what is going on within the nucleus. A rich variety of new phenomena was encountered in the study of nuclear structure on a scale of distances tens of thousands of times smaller than atomic dimensions. The nuclear forces and the spectra of excited nuclear states are very different from the atomic realm. There does not exist as yet a finished dynamical theory of nuclear phenomena at a fundamental level akin to our atomic theory. However, the interpretation of nuclear phenomena and the detection and identification of neutrons and protons as the constituents of nuclei has been accomplished wholly in terms of evolutionary advances within an existing conceptual framework.

There is no guarantee that this framework will still suffice as we push on another factor of a thousand and beyond to a higher scale of energies and a smaller scale of distances onto the quark frontier. But, as its frontiers advance, science can measure its need for new principles and concepts only by testing its existing ones. Therefore our identification of quarks as the basic building blocks of the elementary particle frontier invites comparison and contrast with what we mean when we identify electrons as constituents of the atom and when we identify neutrons and protons as constituents of the nucleus.

Nuclear Constituents

Let us first describe some of the new phenomena encountered when probing within the nucleus and recall how they were interpeted using known concepts. Following the discovery of the neutron in 1932 by James Chadwick, the nucleus was revealed to be a structure built of individual protons and neutrons bound to one another by a new force—a strong nuclear force that is still only partially understood today. This discovery was made originally by observing that individual protons and neutrons, as well as light nuclear structures such as alpha particles formed from them, emerged as the debris from nuclei that were given a hard smash by energetic particle beams. During the early 1930s nuclear physicists went so far as to project a simple world view with all matter, animate or inanimate, built of combinations of just three elementary building blocks, or fundamental constituents—electrons, protons, and neutrons.

Though appealing in its simplicity, this unifying view was nevertheless beset by fundamental problems. First of all it relied on the existence of a mysterious new and very strong nuclear force acting only over very short distances, less than the dimension of the nucleus itself, that binds the nucleus together. It is evident that electrical forces are not responsible for binding nuclei built of protons and neutrons: the protons all have the same electrical charge and therefore electrostatically repel, rather than attract, one another. Furthermore the neutron is electrically neutral and is therefore essentially immune to electrostatic forces. What, however, is the origin of the nuclear force? And what is its relation to the electrical forces responsible for atomic binding?

Another problem was the occurrence of beta nuclear radioactivity. These are nuclear tranisitions leading to the emission of electrons. However, the emerging electrons which form the nuclear radiation were observed to carry only a variable fraction but not all of the energy made available in the nuclear transition. It was soon realized that energy did not balance in the radioactive decay of nuclei to their observed decay products. Rather than abandon on he atomic scale the principle of energy conservation, which is so well verified and which is so important conceptually in macroscopic classical physics and chemistry, an additional unseen particle called the *neutrino* was postulated in 1930 by Wolfgang Pauli. It served as the agent to carry off the missing amounts of energy as well as to provide the momentum balance, as was confirmed by later precise measurements. The neutrino, which is a massless and electrically neutral particle (a little neutron and hence its name) fully earned its status as a respectable member of the particle family only after it was observed directly. "Observation" of a neutrino meant demonstration that it initiated observable reactions when impinging on matter as well as confirmation that, in the radioactive decay process, it carried away definite values of both energy and momentum characteristic of a single particle with zero mass. This story of the neutrino in the 1930s is an example of a fundamental principle being preserved—the principle of energy-momentum conservation—at the price of introducing a new particle, later confirmed by direct detection.

The simple picture of the early 1930s was further eroded by the introduction of a new force together with the new neutrino particle on which it acts. This is the so-called *weak force* responsible for the observed particle decay with beta emission processes of which the neutron decay is the most direct example. Physicists were now faced with very weak as well as very strong forces in addition to the electrical ones that bind the atoms. Subsequently, the remaining vestige of a simple unified view was completely buried in the debris of high energy nuclear collisions which were detected first in cosmic rays and shortly thereafter at the man-made large accelerators. Since the 1940s we have known that when protons and neutrons are hit very hard—as in the collisions of particle beams incident on hydrogen and deuterium targets at high energy accelerators—many fragments are produced. It is evident that both protons and neutrons are themselves also composite systems. Typically the proton changes its state of motion when bombarded by a high energy projectile such as another proton. Its internal excitations are generally very short lived and, as it returns to its normal state, it emits other short lived unstable particles. Some of these emerging fragments are heavier than the proton itself. For very energetic collisions, fragments of antimatter are also produced in abundance. These exotic forms of debris—very heavy particles and antimatter—make their appearance because, the more energy available, the greater the masses that can be created according to the famous Einstein relation $E=MC^2$. They also immensely complicate our quest for the subnuclear constituents of which the particles are built.

Where Are the Quarks?

The quarks were introduced to provide a simple conceptual basis for the observed patterns of subnuclear spectroscopy. Their role is similar to that of the electrons in atomic spectra and of the neutrons and protons in nuclear spectra. Although there is a fundamental dynamical theory, as expressed through Maxwell's equations, of the electromagnetic forces binding electrons in an atom, we do not have a fundamental theory of the strong forces binding either the protons and neutrons together or the quarks to one another. However, a highly successful predictive phenomenology, or scheme-of-things, has been created in the latter two cases for the nuclear and subnuclear spectroscopy. What is new, however, is that in contrast to both our atomic and nuclear experiences we do not detect the individual quarks isolated from one another. When we break apart, or *ionize*, an atom its electrons and its nucleus are clearly evident and are detected in the debris. The same is true when we identify individual protons and neutrons in the debris stripped away in nuclear collisons. How then do we account for our failure to detect quarks in the debris of a shattered proton or of any other subnuclear particle? Must we alter or abandon tested concepts, as Niels Bohr first proposed to abandon conservation of energy in the atomic decays of individual neutrons before Pauli's suggestion of the then undiscovered neutrino? Or is there a more elementary explanation within our established conceptual framework as to why quarks fail to show up in the subnuclear debris of energetic collisions?

Massive Quarks

An elementary and conservative approach to the dilemma of the nonappearance of quarks is to suppose that individual quarks are so massive that they cannot be made by existing accelerators. According to this view the forces of attraction binding three quarks to one another to form a proton, for example, are enormously strong. In order to overcome these attractive bonds and free the quarks one has to add a very large amount of energy to the proton. After this very large amount of energy has been added, the mass of the system formed—in this example the three freed quarks—will be very much heavier than the initial proton as a result of the Einstein relation $E = MC^2$.

This view is an extrapolation to much stronger binding forces of known behavior in nuclear physics. Consider as the simplest nuclear example the nucleus of a deuterium atom, that is, the deuteron, which is composed of a proton and a neutron bound together by nuclear forces. It is a stable substance in nature. If undisturbed externally, the deuteron and its two mutually bound constituents live forever. However, if the deuteron is smashed apart into an isolated proton and neutron we observe that the neutron decays into a proton with the emission of beta radioactivity in about twenty minutes. Evidently the free neutron is sufficiently more massive that it becomes unstable and exhibits a very different behavior than one bound in the deuteron. In fact, however, the binding energy between the proton and the neutron is only a fraction of a percent of the energy of a free proton and neutron, and barely holds them together. In contrast, the subnuclear interaction energies are a factor of a thousand or more greater than those in the deuteron and other typical nuclear systems. Hence their bindings are also comparably stronger, and there is a possibility that the proton itself has but a tiny fraction of the mass that its individual quark constituents would have when they are isolated from one another and free. According to this view there is nothing fundamentally new in the picture of protons made of quark constituents; we just need higher energy accelerators before producing single quarks.

Future experiments may support this picture, but for a number of reasons, some of which are experimental and some theoretical, many physicists are skeptical. For one thing extensive searches for evidence of quarks surviving the initial seconds of the existence of the universe or arriving with the cosmic radiation incident throughout the earth's history have all failed. It is becoming increasingly difficult to explain their nonobservation if they indeed exist. On the other hand, one may assume that very massive quarks are unstable and die so rapidly—in less than trillionths of a second, a typical time unit on the subnuclear scale—that we cannot recognize them as we do the very long-lived neutron constituent of the deuteron with its twenty-minute lifetime. This is again a possibility that cannot be ruled out at present, although it leads to difficulties in the detailed interpretation of existing data, and it lacks any supportive evidence at this time.

One of the most difficult general problems with the very massive quark models is simply the fact that the quarks were introduced originally to provide a simple understanding of subnuclear spectroscopy in terms of relatively light constituents moving approximately as independent particles. The intuitive simplicity of this picture is lost if we must resort to a model with very massive quarks bound together very strongly and moving in tightly bound systems such as the proton. Just as a proton is very much lighter than its constituent quarks, we would have to face a picture of the universe which had lost most of its material mass when matter formed!

Where do we go from here? Nature is telling us on one hand that quarks behave as light, independently moving constituents in a proton. This aspect of quarks is revealed by the general pattern of quark spectroscopy as well as by detailed interpretation of experiments using very high energy particle beams produced at the accelerators. On the other hand, nature has at the same time revealed quarks to be strongly and tightly, if not permanently, imprisoned together in the subnuclear particles. In fact, essentially all of the observed subnuclear particles, both the stable and unstable ones, can be classified into one family, the *hadrons*, which has two major branches. One branch, which includes the proton and all of its many dozens of cousins, is built of three quarks bound to one another in different states of motion. If we assume that there are just three different kinds of quarks, all members of this branch, known as the *baryons*, find a natural classification. The second, or *meson* branch, is built of one quark and one antiquark bound together. The existence of antiquarks along with the quarks themselves is a consequence of Einstein's theory of special relativity and the principles of quantum mechanics, according to which there must be an antiparticle for each kind of particle in nature.

Peculiarly absent from nature are structures built of two quarks, four quarks, or of two quarks and one antiquark, for example. The basic quark combination is always observed to be one quark with one antiquark, or three quarks or three antiquarks together; and no other combination occurs. If one of these baryon or meson particles is shattered in a nuclear collision new particles are formed, each of which is also itself characterized by the same quark content of three quarks, three antiquarks, or one quark and one antiquark. But never have isolated individual quarks been detected!

Analogy to a Bar Magnet

It has been frequently noted that this behavior is analogous to that of magnetic material with which man has been familiar since the early Greek and Roman times when the discovery of naturally occurring magnetic substances such as lodestone was historically first recorded. A magnet always has both a north pole and a south pole. When broken in two, a bar magnet becomes not isolated north and south poles separated from one another, but two magnets each with its own north and south poles. This is very similar to what was being described above: when a meson made of a quark and an antiquark is smashed apart, single quark and antiquark fragments do not emerge. The debris of the shattered particle consists of more mesons, each with its own quark and antiquark; or of more baryons and antibaryons, each made of three quarks and three antiquarks, respectively. From the occurrence of baryons it is evident that the analogy of quarks with magnetic poles is not a literal one. There is no magnetic correspondence to the baryon as there is to the meson. The fact that, in addition to the mesons, baryons made of three quarks exist whereas there are no "magnets" with three north poles is a result of the underlying formal mathematical structure of the quark theory. Nevertheless, the physical analogy of quarks with magnetism is sufficiently close and accurate to be a useful guide.

Our curiosity and present plight with quarks is not very different from that of an inquisitive mariner at sea about ten centuries or so ago. In a moment of calm on a passage he might have viewed a (spare) compass needle, which is just a simple bar magnet, with idle bafflement or scientific curiosity and tried to break it apart in order to separate the north pole from the south pole. But to no avail, for with each breaking of the compass needle he ended up with an additional one having both a north pole and a south pole. The understanding of this impossibility to isolate single magnetic poles came only many centuries later when, in 1820, the French physicist André Ampère first explained magnetism in terms of electrical currents. In fact a fundamental theory of magnetism at the atomic level in terms of the currents of circulating and spinning electrons was achieved only in this century on the basis of the modern quantum theory. When an electrically charged particle circulates it forms a current loop, thereby producing a magnetic field in a pattern similar to that of a compass needle or a bar magnet with both a north pole and a south pole. Many atoms then make a big bar magnet, but never a single pole.

The experimental search for magnetic monopoles continues, as it must also for quarks, for we can never know for certain that they do not occur under exotic or very rare circumstances. However, for all known matter the unbreakable bond between magnetic north and south poles is understood at a fundamental atomic level. Just as a meson built of a quark and an antiquark is not a fundamental particle, a magnet is not the "elementary unit" of magnetism. The elementary unit of magnetism is also not a single north or south pole. However, if we view the quark itself as a fundamental particle we are breaking this analogy with magnetism. On the other hand, if we wish to pursue the analogy of quarks with magnetism, a new "elementary structure," analogous to the atomic current, must be introduced.

A New Elementary Structure?

In either event we face a new point of departure, leading either to a new "elementary structure" or to quarks mysteriously tied together as if by an unbreakable string. If the former is the case, we have neither seen nor theorized creatively as to what these new "elementary structures" are. However, we can hope that, given the expensive and large facilities of high energy physics that are required in order to probe with increasing sensitivity on smaller and smaller scales of distance, it will take less than the nineteen centuries required to advance from lodestone to amperian currents. On the other hand, if the quarks are indeed the elementary particle, then a new kind of dynamics is required, different from that with which we are familiar on the atomic or nuclear scale. In atoms the electrical forces decrease with the inverse square of their separation as the distances between the charged electrons and nuclear protons increase; these forces increase with decreasing separation. For the quark dynamics we seem to find an opposite behavior: when the separation between quarks is small, as deep within a proton or meson, their interaction is relatively weak. However, it grows in strength when they are no longer very close together and remains strong even for large distances between the quarks. This behavior is frequently called *quark confinement* or *quark trapping*. It takes more and more energy to pull a quark and an antiquark further apart against the attraction of this force. Eventually the energy invested increases to the point that it equals the amount required to create a quark-antiquark pair. When those materialize we have the analogue of splitting the compass needle in two; that is, we have two mesons, each built of an original quark and its materialized antiparticle.

There is optimism that the general structure of dynamical theories exhibiting such properties has been specified. The technical problems of solving these so-called asymptotically free gauge theories have not been all worked out, but special calculations have led to encouraging results in some cases. Whether the observed subnuclear particles are built out of elementary currents analogous to the amperian currents underlying magnetism in a future theory of the submicroscopic world, or whether they are built out of elementary quarks permanently tied together as if by strings so that we cannot isolate them in the laboratory, our concept of the elementary particle will be greatly changed. In any event, unless we actually detect the individual constituents, modern physics will find itself closer to the biblical description found in Chapter 11, Verse 3, of the Epistle of St. Paul to the Hebrews: "Through faith we understand that the worlds were framed by the word of God, so that things which are seen were not made of things which do appear."[6]

Werner Heisenberg approached this issue from a different viewpoint in a lecture on the nature of elementary particles delivered in 1975 shortly before his death. He raised the possibility that the root of our present dilemma is that we are asking the wrong question in particle physics when we ask what a proton "consists of":

> I will now discuss that development of theoretical particle physics that, I believe, begins with the wrong questions. First of all there is the thesis that the observed particles such as the proton . . . consist of smaller particles: quarks . . . or whatever else, none of which have been observed. Apparently here the question was asked: What does a proton consist of? But the questioners appear to have forgotten that the phrase "consist of" has a tolerably clear meaning only if the particle can be divided into pieces with a small amount of energy, much smaller than the rest mass of the particle itself.[7]

Heisenberg is referring here to the fact that very high energies are required to overcome the strong binding forces and shatter a proton into its various forms of debris, or pieces, many of which may be as massive or more massive than the proton itself. These energies are hundreds to thousands of times greater than those required to break apart a nucleus and billions of times greater than needed to break up the atom. We are no longer dealing with energies that are but small fractions of the rest masses of the particles themselves. We can no longer rely solely on nonrelativistic ideas of slowly moving particles and weak forces. Rather we must contend with the special theory of relativity because in the presence of strong binding forces and high energies the distinction between energy and mass can no longer be clearly made. Both matter and antimatter are present and the energy can transform back and forth between the various possible forms into which it can materialize.

We must also take into account another aspect of the theory of relativity, and that is the fact that the force between neighboring particles such as between two quarks does not act instantaneously but is transmitted no faster than the speed of light. If we are dealing with weakly bound systems with slowly moving constituents we can neglect the fact that the speed of light, although very great, is finite; or we can take it into account by suitable simplifying approximations. Within baryons and mesons this is no longer possible and the dynamics of how energy and momentum are exchanged among interacting quarks and bind them together is central to the description of the structure of subnuclear particles. We have earlier referred to unbreakable strings that tie the quarks to one another. These strings may be fundamental new dynamical objects, or, viewed more conservatively, no more than an approximate description of the effects of the elementary quanta known as *gluons* that carry the discrete units of energy and momentum between quarks and antiquarks, "gluing" them into mesons and baryons. On the subnuclear level the gluons are the analogues of the photons or quanta of the electromagnetic field for atomic processes. A complete dynamical description of the structure of the proton necessarily prominently involves the gluons, or strings, as well as the quarks themselves. Moreover, as the quarks move around, the pattern and numbers of gluons binding them to one another also change. So do the forms of matter and antimatter into which the gluons can materialize. Hence a complete picture of what the proton consists of is less than clear, as Heisenberg suggested. Most physicists agree, however, that the simple quark picture is a very useful, if not a literal, picture of the structure of hadrons.

How Many Quarks?

The motivation for introducing quarks in the first place was to bring unity out of diversity—to attempt to explain the construction of a vast number of particles and their properties using only a few building blocks. At the atomic level this was achieved by introducing the electron and building up all the known elements in terms of different numbers of identical electrons in different atomic orbits. At the nuclear level a basic classification of all nuclei was achieved in terms of different combinations of just two kinds of particles—the proton and the neutron. We of course also want to know how many different kinds of quarks there are as the subnuclear building blocks. In the original proposals of Gell-Mann and of Zweig there were three different kinds (denoted by u, d, and s for up, down, and sideways, or "strange," with the physicists' usual penchant for simple and wry vocabulary). These three, together with their inevitable antiquarks, were all that was required to explain all observed nuclear particles.

A novel feature of the quarks as originally proposed is that they are assigned electric charges that are fractions of the electron charge. All observed particles in nature have either zero electric charge or an integer number of units of the electron charge. As fractionally charged particles with either one-third of the electron's charge (the d and s quarks) or two-thirds and the opposite sign of the electron's charge (the u quark) quarks were a new invention.

The two quarks with the same one-third unit of charge differ in another physical property, or in the language of physics, in one of their *quantum numbers*. Physicists identify and describe subatomic particles by assigning them quantum numbers; each number designates a property that is conserved, or left unchanged, when particles interact. Some quantum numbers, such as electric charge and spin angular momentum, refer to physical, measurable attributes of the particle. Others are more abstract; they denote family resemblances among particles and provide a valuable bookkeeping system for classifying particles and their interactions in algebraic form.

Some quantum numbers—such as electric charge—are always conserved. Certain others are only "approximately conserved"—that is, for most interactions they are conserved but in some cases, the rare ones, they are not. On the way to a fundamental dynamical theory the search for conserved quantum numbers has been a very productive and valuable one for physicists. For each quantum number there exists a corresponding symmetry of the underlying dynamical laws. With each discovery of a quantum number, or conservation law, the possible forms of the underlying laws of nature—the "holy grail" of natural science—are further delimited and the search more narrowly focused.

The great excitement and frenetic activity in elementary particle physics during the past two years, that I mentioned at the outset, has been caused by the experimental discovery of an additional new species of quark: a fourth quark carrying a new quantum number called *charm*. Like the three quarks before it, the new quark with charm was not directly detected by itself. What was actually observed in the laboratory was a new family of particles with extraordinarily surprising properties: they were very heavy, more than three times as massive as the proton, but still they lived for a very long time, about ten thousand times too long on the subnuclear time scale. In order to accommodate the new particles in a Mendeleev or Gell-Mann/Zweig table for quarks, a new species, the fourth quark, was required.

But it is not quite that simple. As the quark hypothesis was applied in detail during the past decade it became clear that the full body of data to which it was applicable could be accommodated and interpreted in terms of known fundamental particle properties only if each of the three, and now four, different quarks was endowed with yet another quantum number that could assume just three possible values. This new property was dubbed colloquially *color*, though it has nothing to do with the usual meaning of color. In effect each of the four different kinds of quarks was tripled so that with this extension of the quark family there are now altogether twelve quarks as basic building blocks (along with their inevitable antiquarks). In addition more may be required if, as the experiments reach to higher energies, new genera of particles are encountered. To some physicists there are already compelling suggestions based on the detailed analysis of recent experiments that no fewer than six different kinds of quarks, each occurring in three colors for a grand total of eighteen, are required. One wonders whether the family of quarks has not already grown so large that there are too many for them to be legitimate candidates for "fundamental particles"! Perhaps a smaller family of basic currents and a simpler scheme of symmetries will mark the next major advance analogous to the amperian currents underlying magnetism. Whether or not the quark is the "end of the line" as the Greek "atom," the scientist can be sure of one thing: nature's imagination has always proved richer than man's vision of what lies ahead beyond the next frontier of scientific inquiry.

Unity of Forces

It is the fundamental faith of a scientist that simple general laws exist in terms of which the rich diversity of nature can be explained. As noted earlier in quoting Heisenberg, we may question whether the search for basic building blocks or elementary constituents of the subnuclear world has been correctly posed. But there is no doubt about the ultimate goal of a unified theory which reduces the broad variety of experience to simple and general laws. To achieve this goal we must know not only the basic constituents or currents that interact with one another; we must also know the nature and variety of the quanta transmitting the forces, as well as the relation between quanta and the currents or constituents that are their sources.

At present there are four distinctly different kinds of forces that are believed to account for all observed interactions of matter: gravitation, electromagnetism, the strong force, and the weak force. In the everyday world gravitation is the most obvious of the four forces; it influences all matter and acts over very large distances. For the infinitesimal masses involved in subatomic events, however, its effects are vanishingly small and can be ignored. Its quantum is called the *graviton*. The electromagnetic force is also felt over very large distances, but it acts only on matter that carries an electric charge or current. The *photon* is the quantum, or carrier, of the electromagnetic force, and when two particles interact electromagnetically, they can be considered to exchange a photon or photons.

Following the development and confirmation of his general theory of gravitational phenomena, Einstein and his colleagues spent three decades in a futile effort to construct a unified theory including the effects of both gravitation and electromagnetism in a single general force law describing all phenomena of classical physics on a macroscopic level. In the realm of atomic and subnuclear phenomena we may ignore the very weak gravitational effects, but we face the very great challenge of unifying electromagnetic forces with the weak and the strong ones.

The strength referred to in the names of the strong and the weak interactions is related to the rate at which the interactions take place. The strong force is short ranged; that is, its influence extends only over a distance comparable to the sizes of the subnuclear baryons and mesons. When two particles that interact by means of the strong forces approach within this distance there is a high probability that they will deflect one another or that they will produce other baryons or mesons. Compared with the strong force the weak one, which is responsible for the neutron decay, is very feeble indeed. It is weaker by a factor of roughly ten trillion for low energy collisions. It is even shorter in range by a factor of a hundred or so than the strong forces, and it has the unique property of growing in strength as the collision energy of the interacting particles increases.

Apparently the weak, electromagnetic, and strong forces have little in common, and the effort to unify them would seem to be hopeless, or at best discouraging, were it not for the guidance available from symmetry principles and conservation laws.

The first crucial step in the unification of the weak and the electromagnetic forces dates back to 1967 and is largely the work of Steven Weinberg and Abdus Salam. If the weak processes, such as neutron decay with beta emission, are governed by a theory of the weak forces that is constructed on the same symmetry principles as the theory of electromagnetism, we would expect to find common features of the weak and electromagnetic interactions. For example, for the weak processes massless quanta should be radiated and absorbed just like photons, and the weak forces should extend over very long ranges. This is very far from the observed case, however. There are no massless quanta for the weak forces, and the extreme difference in the apparent character of the short-range weak forces from the very much stronger and long-range electromagnetic forces defied attempts at unification for many years prior to 1967. It was realized then that the fundamental equations of the weak and electromagnetic interactions could in fact be unified under a common symmetry principle, but—and this was the key new idea—the symmetry is partially broken in a well-defined and simple way for weak processes, though not for the electromagnetic ones, which honor it. As a result of this symmetry-breaking, the quanta of the weak interactions are predicted to acquire a mass approximately forty or more times heavier than that of a proton. It is this large mass that is responsible for the very different character of the weak and electromagnetic processes as studied with present accelerators.

However, when we probe the forces at very small distances the unification schemes predict that the two interactions become very similar in character. In order to experimentally study such small distances, which must be well within the short range of the weak forces themselves or less than one hundredth the size of the proton itself, interaction energies are required beyond those that can be achieved with existing accelerators. This condition defines the new high energy frontier of elementary particle physics. "High energy" in this context means energies that are large when compared with the mass energy of the weak interaction quantum, or energies hundreds of times larger than the mass energy of the proton itself. On such a high energy and small distance scale the mass of the weak interaction quantum can be neglected and the common features of the weak and electromagnetic forces will be revealed if these very attractive ideas of an underlying symmetry principle are correct.

There are already experimental results in accord with some of the predictions of this approach to a unified theory of weak and electromagnetic processes. In particular, new types of neutrino-scattering processes can be, and were, predicted. These have since been observed and have a strength compatible with the theoretical framework of the unified theories. The predicted properties of the quanta of the weak interaction processes include their large masses. Although the thresholds for producing them lie beyond the ranges of energies accessible to the existing accelerators, they are within the range of the next generation of machines now being planned and designed. Physicists are confident that these quanta will be found!

Strong Forces

That leaves the strong processes—those involving the quarks and gluons—still to be explained. Detailed progress here is more difficult because the mathematics of dealing with very strong forces is technically much more formidable. Even if our equations are correct we have yet to construct their accurate solutions! Yet there is optimism that, by once again modeling our ideas on the symmetries found in the theories of electromagnetic (and now also of the weak) processes, we may be making progress. In this case gluons are the quanta and also acquire large masses because of a broken symmetry. Quarks carry the color quantum number, as noted earlier, and interact with the gluons in much the same manner that electrons interact with photons, the quanta of the electromagnetic field. Color plays an analogous role to electric charge for the electron: only the charged particles interact with the electric forces, and similarly only quarks with the quantum number color, and baryons and mesons built out of quarks, react to the strong forces. Electrons and neutrinos do not carry color and hence do not react to strong forces.

With this approach the color quantum number acquires a fundamental importance as the "charge" of the strong forces. The absence of isolated quarks is then equivalent to the absence of any states of color in nature. According to the mathematical formalism associated with the idea of color, the only structures that can "hide their color" are those formed of the combinations of a quark and antiquark, or of three quarks. These are the color analogues of electrically neutral systems, and they are the only ones to occur in nature. On this basis we can understand why all observed mesons and baryons have this quark structure and none other: they are neutral in color and hence not confined.

Analogies often prove very useful to guide scientific inquiry. If taken too literally and followed uncritically they also can lead one astray. In no way can they substitute for a complete dynamical theory with predictive power. Whether the analogy of quarks with magnetic poles as described earlier, or that of color with electric charge as given above, remains very useful for very long is uncertain. However, the very possibility that, guided by symmetry principles, we are on the right track toward a unified theory of nature's laws has had a very exhilarating effect on elementary particle physics. Predictions of allowed, inhibited, and forbidden processes have enjoyed experimental successes. However, grave hazards remain because the theoretical structure is still in the process of being constructed and is necessarily largely untested. At least we have, for the first time, simple, unified, and even aesthetic principles as guides.

This theory that is built in terms of quarks and gluons that are permanently confined, and in terms of the "hidden" quantum number color that, unlike the electric charge and other such quantum numbers, also can never be seen, carries with it the ghost of the aether. In the late nineteenth century there was prodigious scientific activity devoted to endowing the aether, or the medium thought to transmit light and electric signals, with a physical reality: it drifted in space, it dragged, it compressed, etc. At the same time, however, one never observed the aether, and there were increasingly intricate and artificial excuses for hiding evidence of it from experiment. The web of intrigue surrounding the invisible aether was cleared away by Einstein in 1905 with his theory of special relativity. Some of today's elementary particle physics concepts, such as color, not to mention the quarks themselves, remain hidden from direct experimental observation. It is not beyond the realm of possibility that they may also suffer the fate of the aether, though few today find it possible to conceive of establishing an underlying unity on the subnuclear frontier without the quarks. And what will happen to this theoretical edifice being constructed if isolated quarks are in fact observed in the laboratory?

Conclusion

As its frontiers advance in the exploration of ever smaller distances and systems, the concepts, if not the vocabulary, of elementary particle physics grow more and more abstract. Still there persist strong intellectual bonds with other fields of natural science as a vital source of stimulation. For example, what we learn on the subnuclear frontiers is already shaping the astrophysicists' views of what took place in the first seconds of the "big bang" at the creation of our universe. It contributes to our understanding of the continuing evolution at the far reaches of outer space as new objects of highly concentrated energy are discovered and studied. Also the richly developed theory of critical phenomena and phase transitions in the field of many-body physics concerned with the behavior of macroscopic solids, liquids, and gases has provided seminal ideas for the study of quark confinement.

In order to further advance the subnuclear frontiers, higher energies and hence bigger accelerators and detection devices are also required. There are but few centers in the world where the frontiers can still be experimentally probed with the very sophisticated and complex equipment on which we must rely for the data that are the lifeblood of physics—and of all natural science. As the number of such facilities decreases and the complexity of performing the experiments themselves increases, large research teams and collaboration on an international scale become the norm. The field of elementary particle physics is unavoidably and unquestionably "big science."

References

[1]Aristotle, *The Physics*, translated by Philip H. Wicksteed and Francis M. Cornford (Cambridge, Mass.: Harvard University Press, 1963), Book 1, Chapter 1, p. 11.

[2]See George Sarton, *A History of Science: Ancient Science Through the Golden Age of Greece* (Cambridge, Mass.: Harvard University Press, 1952), p. 253.

[3]See Sarton, *A History of Science*, p. 242. Anaxagoras' infinitely divisible seeds were not embedded within each other like layers of the onion's skin, for each seed was as complex as the whole. This is more akin to the modern "bootstrap theory" of elementary particles developed by Geoffrey Chew of the University of California, Berkeley, and colleagues in recent years.

[4]N. Bohr, *Atomic Theory and the Description of Nature* (New York: Macmillan, and Cambridge, England: Cambridge University Press, 1934), p. 1.

[5]W. Heisenberg, *The Physical Principles of Quantum Theory*, translated by Carl Eckart and F. C. Hoyt (New York: Dover, xxxx), p. 3.

[6]King James Version. I am indebted to V. F. Weisskopf for first directing my attention to this passage.

[7]W. Heisenberg, "The Nature of Elementary Particles." *Physics Today*, 29 (3) (1976):38.