SLAC-PUB-3288
February 1984
(N)

Particle Substructure. A Common Theme of Discovery in this Century†

W. K. H. PANOFSKY*
Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305

SLAC-PUB--3288

DE84 006776

It is not once nor twice but times
without number that the same ideas
make their appearance in the world.

Aristotle, *On the Heavens*

1. Introduction

After exposure to public accounts of scientific work the layman might conclude that new results or 'breakthroughs' in science tend to supersede results that were previously known and accepted. However, in physics this is rarely the case. Rather, the more common situation is that the range of validity of an older concept is found to be less than universal. For instance, Newtonian mechanics remains as valid for velocities small compared to the velocity of light as it did before the advent of relativity.

What is even more remarkable is that, despite the explosion of knowledge about nature on a smaller and smaller scale, some of the basic rules that describe the behaviour of matter remain equally valid today. It is to this latter phenomenon that this talk is addressed. Specifically, I will give some examples from modern developments in particle physics which demonstrate that the fundamental rules of quantum mechanics, applied to all forces in nature as they became understood, have retained their validity. The well-established laws of electricity and magnetism, reformulated in terms of quantum mechanics, have exhibited a truly remarkable numerical agreement between theory and experiment over an enormous range of observation.

As experimental techniques have grown from the top of a laboratory bench to the large accelerators of today, the basic components of experimentation have changed vastly in scale but only little in basic function. More important, the motivation of those engaged in this type of experimentation has hardly changed at all.

* Work supported by the Department of Energy, contract DE-AC03-76SF00515.

† Lecture Presented at the Cherwell-Simon Memorial, Oxford, England, May 8, 1981. Published in Contemporary Physics, Vol. 23 (1982).

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

2. The nuclear atom

Let me begin this stody with the classical experiments of Lord Rutherford around 1911 which demonstrated that the atom is not a continuous blob of matter, but is rather a structure whose mass is concentrated in a small, central nucleus. Let me show the basic arrangement of Geiger and Marsden's experiment in his laboratory.

Rutherford's experiments made use of a beam of α-particles formed simply by collimating the particle flux emitted in all directions by a natural radioactive source. This *beam* of α-particles was then directed at a *target*, which in the case of Rutherford and collaborators was a thin foil of the material under study. The particles were then scattered into a *detector*, which in this case was a fluorescent zinc sulphide screen. This was in turn viewed by a graduate student through a small microscope. Thus we see here a typical scattering experiment—beam (collimated α-particles), target (scattering foil), and detector (fluorescent screen, microscope and graduate student). You will see that all of the more recent scattering experiments to which I will refer still have precisely these same basic components.

Scattering experiments are the most common method by which the physicist analyses the constituents of particles which are not (with the techniques available at a given epoch) accessible to *individual* observation. A scattering experiment simply measures the probability that an incident particle is scattered at a given angle away from its initial direction and in so doing either has or has not suffered an energy loss or change in some other characteristic during the scattering process. If the energy of the incident particle after deflection is diminished only by the recoil energy of the target particle without any other changes, then the scattering is called elastic; otherwise it is called inelastic.

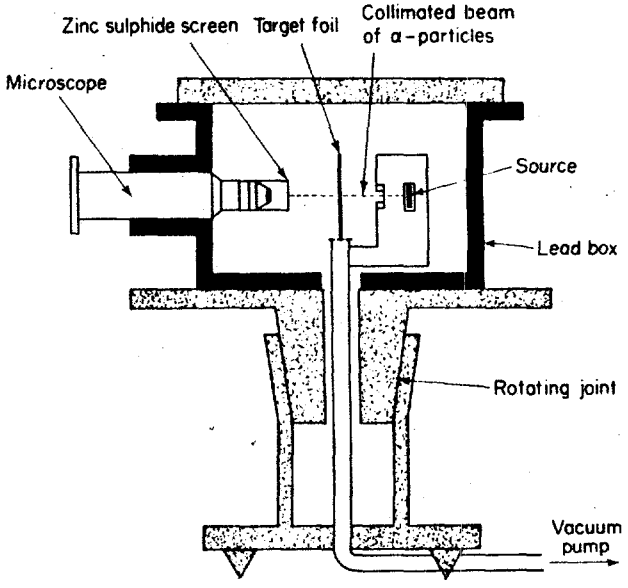

Fig. 1. Schematic diagram of Geiger and Marsden's apparatus which demonstrated the existence of the nucleus within the atom. The apparatus consists of a source of α-particles followed by a collimator which forms a beam of these particles striking a target foil. The particles then strike a zinc sulphide fluorescent screen after scattering in the foil. The location of the particles on the screen is determined by a human observer viewing the screen through a microscope.

Scattering of one particle on another gives a 'kick', or what the physicist calls 'momentum transfer', to both particles. It is the magnitude of this momentum transfer which measures the scale to which the scattering process can give information on the structure of the particles. The relationship between the resolution Δx to which the spatial structure of the particles can be revealed by the scattering process and the momentum transfer Δp is given by Heisenberg's Uncertainty Principle, $\Delta p \Delta x \cong h/2\pi$ where h is the Planck constant. This relation, in turn, sets the practical energy required for the incident beam if structure down to a specified dimension is to be resolved.

It is well known that Rutherford's scattering experiment led to the discovery of the nucleus, that is the central core of the atom which is about 10000 times smaller in diameter than the atom itself, yet contains almost all of its mass. Figure 2 shows schematically how Rutherford reached his conclusion. It was observed that the number of particles scattered with large momentum transfer (that is, at a large angle) exceeded the amount predicted assuming that the atom was a continuous blob of matter.

It is interesting to note that the analytical tools available to Rutherford to predict how much scattering would occur at what angle had to be based on the then accepted laws of classical mechanics and classical electricity and magnetism. In other words, even though Rutherford's experiments probed matter at a scale smaller than had hitherto been accessible to experimental observation, he had to assume that the physical laws derived from large scale observation were still valid at small distances. Thus, using the same rules to analyse atomic collisions as would be used to analyse collisions between charged ping-pong balls, the conclusions about the existence of the atomic nucleus were drawn.

In retrospect, Rutherford's conclusions were right but his methods were wrong; we know that at the magnitude of momentum transfer which was involved in the collisions observed by Rutherford, the laws of classical mechanics would no longer be valid but that instead the laws of quantum mechanics, which were not established until more than a decade later, were to be applied.

It turns out, however, fortuitously, that the classical and quantum mechanical calculations for the probability of scattering give essentially the same answer provided the scattering is controlled by the laws of electromagnetism, and also provided that the

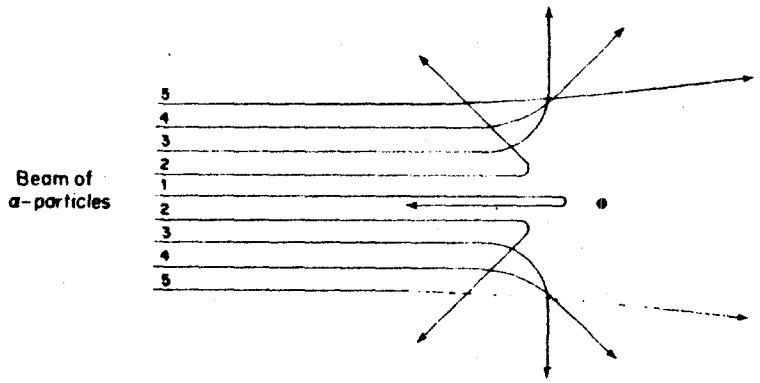

Fig. 2. Schematic representation of orbits of α-particles 'encountering' with a concentrated positively charged object. Note that close 'collisions' result in large angles of deflection, while collisions at a distance produce only a small change in direction of the incident particles.

scattered particles move at speeds well below the velocity of light. Had the forces active in Rutherford's experiment been other than those of electromagnetism, then indeed the quantitative result of Rutherford's calculations would have been incorrect.

3. The scattering of X-rays

Now let me turn to another class of scattering experiment, the scattering of light and X-rays from the atom. The scattering of light or X-rays from the atom is dominated by the extra-nuclear electrons, and not by the nucleus. The reason is that the interaction of light or X-rays, which are electromagnetic radiations, is with an electric *current*; in the scattering process such a current is produced by the recoiling particle from which the scattering takes place. If the mass of the recoiling particle is small, then a larger recoil current is produced. Thus the very light electrons surrounding the heavy nucleus are the principal contributors to the scattering of electromagnetic radiation by the atom.

This type of scattering was observed first by A. H. Compton in experiments beginning in 1922. In these experiments it was shown clearly that the dynamic properties of the incident X-ray beam were described by its photon characteristics, that is, by the particle-like energy and momentum variables of electromagnetic radiation predicted by a quantum mechanism.

It is characteristic of the process of scattering of X-rays by a free particle that the frequency and energy of the scattered quantum is shifted relative to those of the incident radiation. Particularly interesting is the fact that the shift in wavelength of the X-rays in the scattering process depends only on the mass of the particle struck and on the angle

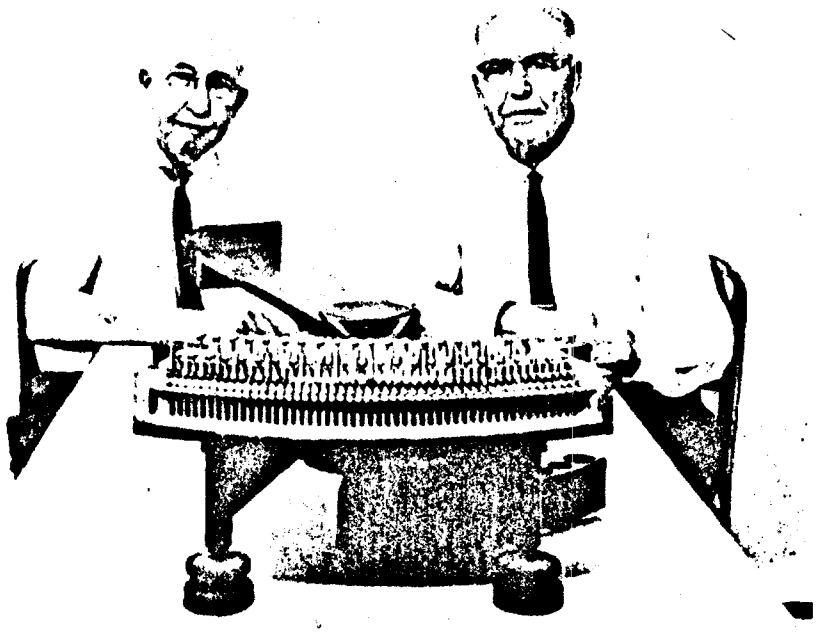

Fig. 3. Photograph of J. W. M. DuMond and his collaborator, Professor H. A. Kirkpatrick, with their multi-crystal spectrometer. This instrument was used to analyse the shape of X-ray spectra which resulted from the scattering of X-ray photons from various materials.

of deflection of the incident radiation; it is independent of the energy of the incident X-rays. The most important result from the observation of Compton scattering was the establishment of the quantum properties of the photon itself, rather than any new insight into the structure of the atom. In fact, in the original Compton scattering experiments it would have made little difference if the electrons on which the scattering took place were free or were bound to the nucleus.

The power of Compton scattering in analysing not only the structure but also the internal dynamics of atoms became evident with the experiments of J. W. M. DuMond and collaborators starting in 1926, which observed the frequency of scattered photons with much higher precision. These experiments revealed not only that scattering of incident photons on atomic electrons indeed took place, but also that these electrons

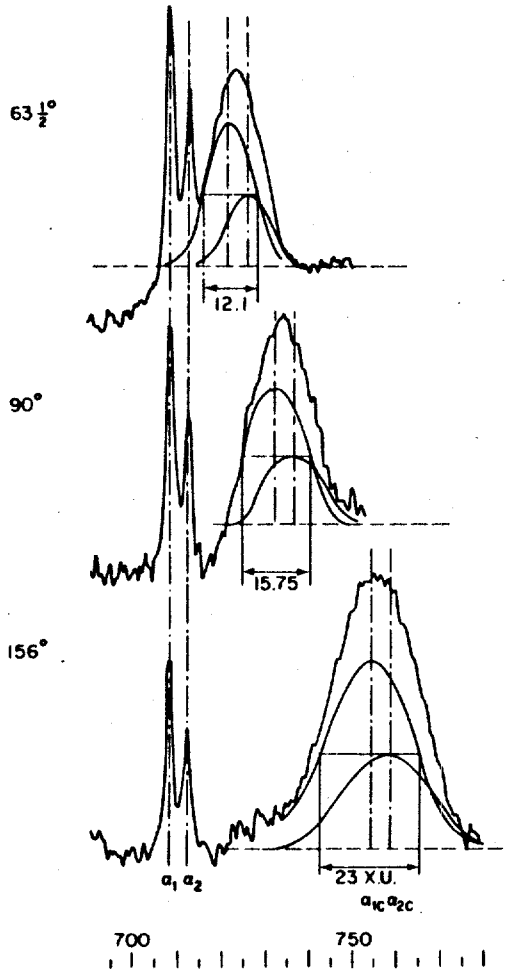

Fig. 4. Microphotometer traces of photographic spectra obtained in the multi-crystal spectrometer of DuMond and Kirkpatrick. The spectra show on the left the unmodified molybdenum K_{α_1} and K_{α_2} lines, together with on the right the Compton-shifted line as scattered from graphite at various angles. Note that the broadening as well as shift itself increases with the scattering angle, as expected by theory. (The scale at the bottom is graduated in ~~x.u.; 1 x.u. ≈ 10^{-13} m.~~)
X.U.; 1 X.U. = 10^{-13} m)

were themselves in a state of motion. Figure 3 shows DuMond with his apparatus. This instrument, composed of a complex assembly of crystals, recorded data on photographic plates. In turn, the exposure of these plates was measured by a microphotometer; figs. 4 and 5 show samples of the resulting spectra. Classically the experiments of DuMond and collaborators could be interpreted to yield the velocity distribution of the electrons in atoms. Figures 4 and 5 show that the dynamic width of the 'scattered' line increases with scattering angle and incident wavelength, respectively, as can easily be computed from the nature of the process.

DuMond initially analysed his scattering experiments by assuming that the electrons moved in classical Bohr orbits around the nucleus. In that case Compton scattering is simply modified by the Doppler shifts produced by the motion of the struck particle. Quantum mechanically, one must analyse modified Compton scattering by describing the momentum distribution of the extranuclear electrons on a probabilisitic basis, without talking about actual orbits. Both types of analysis give

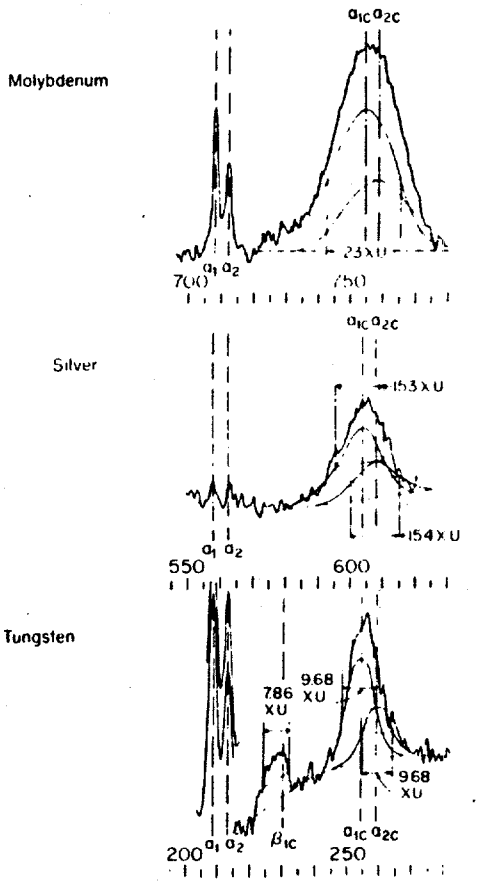

Fig. 5. Microphotometer traces of photographic spectra obtained in the multi-crystal spectrometer of DuMond and Kirkpatrick. The spectra show both the unmodified and Compton-shifted K_{α_1} and K_{α_2} lines of molybdenum, silver and tungsten, respectively, scattered at a fixed angle from graphite. Note that the shorter is the wavelength, the narrower the scattered line. Note also that the central Compton shift at fixed angle is independent of the incident wavelength. (The scales are graduated in ~~x.~~ X.U, defined in Fig. 4)

similar results. In a certain sense these experiments are complementary in that Rutherford scattering is elastic scattering on the nucleus, while Compton/DuMond scattering is *inelastic* scattering on the atom as a whole, but *elastic* scattering on the atomic electrons. The former constitutes the basic discovery of the structure of the atom, while the latter constitutes measurements of a dominant feature of the dynamics of that structure.

The above discussion has dealt with scattering by a *single* electron. Naturally, if more than one electron is involved, the situation becomes more complex. This complication arises if the wavelength of the incident radiation is comparable to the size of the distribution of the electrons. In that case, the scattering of the light by the individual electrons produces interference effects similar to those observed when visible light scatters off the individual elements of a diffraction grating. In general, one can separate the observed scattering into two factors: one is the term that governs the probability of scattering of the X-rays from an individual electron, and the other is the factor which measures the interference effect due to the multiplicity of scattering sources; the latter is known as a 'form factor'. We will meet this type of factorization again when we talk about scattering at much higher energies.

4. High energy electron scattering

Now let us switch from X-rays to electrons for the incident beam and go forward in time by about four decades and up in energy of the particles to be scattered by a factor of about a million. Scattering again can be both elastic and inelastic. Elastic scattering yields information on the radius and general distribution of charge within the proton and neutron. It was elastic electron scattering experiments during the 1950s, for which Robert Hofstadter received the Nobel Prize in 1961, which determined these basic parameters. The fact that the proton has a finite radius indicates in itself that the proton cannot be an ultimate constituent of matter, but rather must have a substructure of some kind. The general nature of this substructure was revealed through experiments at The Stanford Linear Accelerator Center, SLAC, beginning in 1967, which concentrated primarily on inelastic rather than elastic scattering, that is, interactions in which the proton is disintegrated in consequence of the scattering process.

Figures 6 and 7 show, first schematically and then as an actual photograph, the apparatus that was used at SLAC to study both inelastic and elastic scattering of electrons of energy up to 20 GeV on hydrogen targets. Note that the basic components of this apparatus are the same as the ones used by Rutherford. We have an incident *beam* of charged particles; we have a scattering *target*, here consisting of a chamber containing liquid hydrogen; and we have a *detector* composed of magnetic spectrometers which measure precisely the angle of scattering and the energy and nature of the scattered particle. Thus the basic components of a scattering experiment have remained the same throughout this century, as has the spirit of the investigation: Rutherford wished to investigate the substructure of the atom, while the SLAC experiments investigated the substructure of the proton and neutron, which had been established as the fundamental building blocks of the nucleus discovered by Rutherford. While the basic nature and motivation of the experiments have not changed, the scale indeed has. This is a consequence of the uncertainty principle: in order to study matter at smaller dimensions, the transfer of momentum has to be proportionately larger. Roughly speaking, the proton has a diameter one hundred thousand times smaller than the atom, and therefore the energies have to be increased roughly in that proportion.

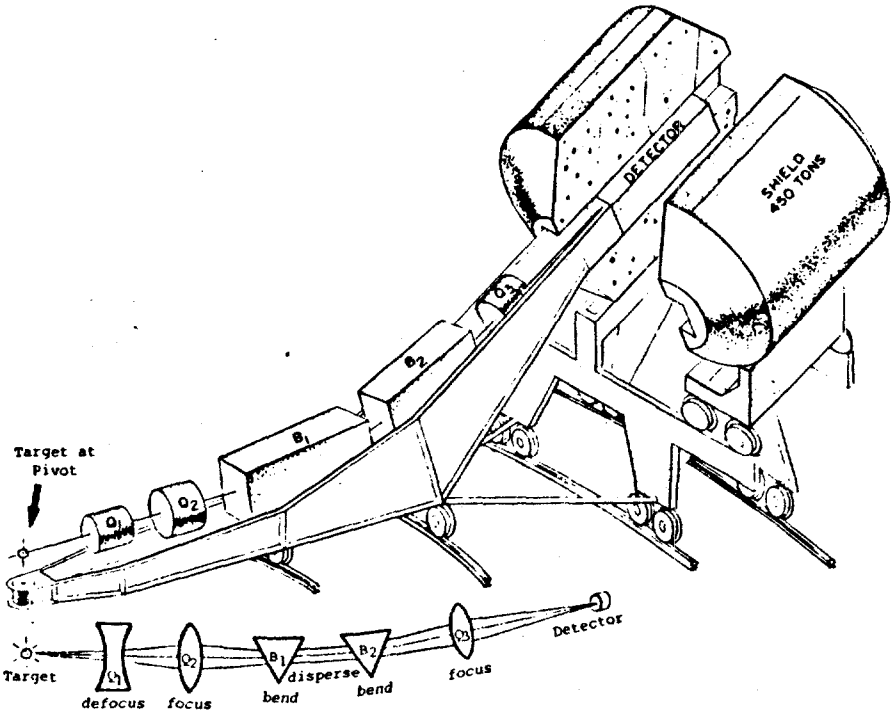

Fig. 6. Diagram of the magnetic spectrometer at SLAC capable of analysing particles of momentum up to 8 GeV/c. The incident beam strikes a target around which the spectrometer rotates. The spectrometer consists of focusing quadrupole lenses 'Q' and deflecting magnets 'B'. The system analyses the particles scattered in the target which are then identified and registered in the detector.

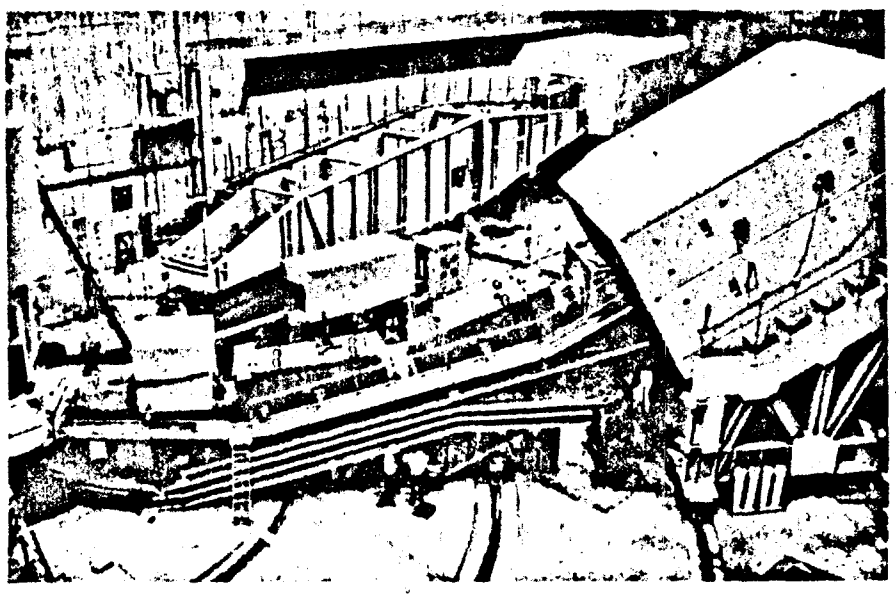

Fig. 7. Photograph of the spectrometer shown in fig. 6, together with a second instrument capable of analysing particles up to 20 GeV/c.

In analysing the inelastic scattering the experimental physicist has several tools at his command. He can examine the fragments that are ejected as a result of the disintegration process, or he can examine the loss of energy of the incident electron during the scattering process; it is this loss in energy which presumably corresponds to the energy of creating and propagating the 'ejecta' from the proton. Most of the relevant information regarding the electron scattering process was obtained by experiments that looked only at the scattered electron.

If we assume that the proton has constituents that are point-like, or at least have a radius very much smaller than that of the proton itself, then an inelastic scattering process can be envisaged as the totality of elastic electron scattering from each of these point-like constituents. This is shown in fig. 8. One can easily derive geometrical relations between the direction and energy of the incident and scattered electron under a number of plausible assumptions. Specifically, by arguments exactly analogous to those applied to X-ray scattering by the atom, one can show by simple mathematics that the energy and angle distribution of the outgoing electron should be a product of two factors. The first factor is the distribution corresponding to scattering from a single, point-like object. This first factor can be calculated with confidence from theory, since the forces governing the interaction between an electron and a point-like charged object are almost purely electromagnetic. In essence, electron scattering is thus "exploring unknown structures with known forces". The second factor, again called the 'form factor', is a function characteristic of the distribution of the point-like objects within the proton. As long as the basic model of the process is correct (i.e., scattering occurs from point-like objects coupled by forces that can absorb transverse momenta up to a specified limit), then the form factors should depend only on a certain dimensionless ratio defined by the kinematics of collision. This ratio can be identified with the fraction of the momentum within the proton which is carried by the struck point-like object. The assumption made here is that binding among these objects is negligible in considering their motion.

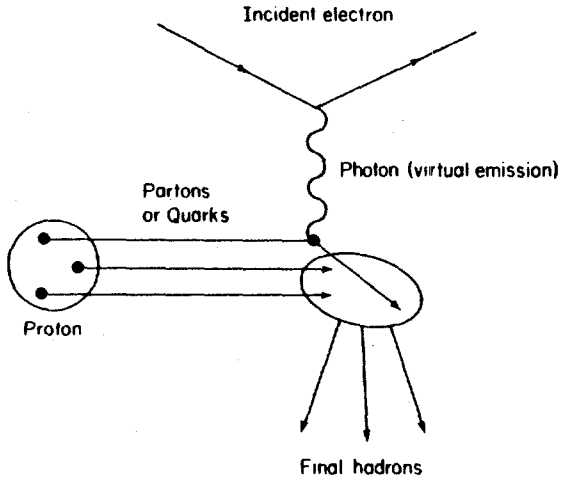

Fig. 8. Diagram of the deep inelastic electron scattering process on the proton. It is assumed that protons are composed of three partons or quarks. The electron 'virtually' emits a photon during the scattering process which interacts with one of the three partons. The three partons interact after the scattering process and recombine, forming a combination of final hadrons.

This simple dependence of the 'form factors' is known as 'scaling' and can be considered to be an indication of the fact that such point-like constituents within the proton might indeed exist. Figures 9 and 10 show how well this simple description agrees with the experimental data. It can be seen that the agreement is good but not

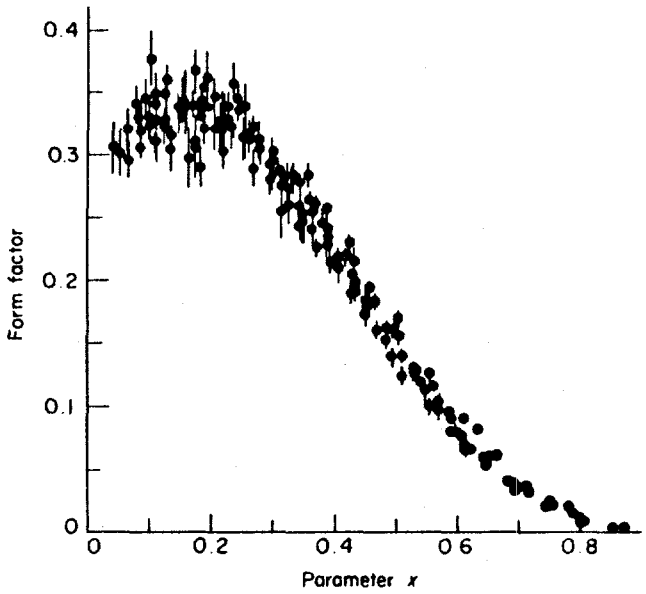

Fig. 9. The experimental evidence for 'scaling' in deep inelastic electron scattering on the proton. The graph shows the form factor plotted as a function of the parameter x which is the fraction of the momentum carried by the 'parton' struck by the incident electron, as measured in a frame in which the proton is in rapid motion.

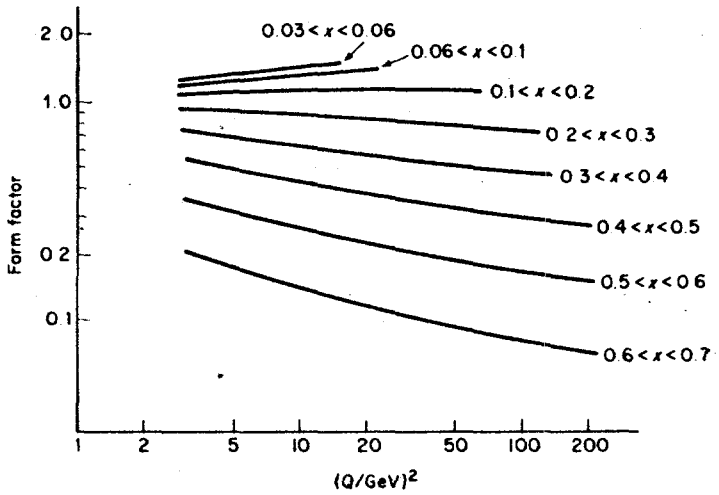

Fig. 10. Deviation from scaling. These graphs show that there is a weak dependence on energy Q of the form factor in deep inelastic scattering where the character of that weak dependence is different for different values of x. Interpretation of this type of information gives valuable information on the interaction between quarks within the proton.

perfect. Naturally the total range of variables over which the correctness of this model can be tested is limited both by the energy of the available beams and by the data rates which can be recorded.

Another, more dramatic demonstration of the existence of point-like particles can be made by comparing elastic and inelastic scattering. This is shown in fig. 11. It is seen that the probability of elastic scattering on the proton as a whole falls off much more rapidly with increasing momentum transfer than does the probability of inelastic scattering which involves presumably only single point-like objects. You will recall that this is precisely what Rutherford observed in regard to the nucleus within the atom. By this analogy we see at least a strong indication that the proton and neutron do indeed have point-like constituents, which were first dubbed 'partons' by Richard Feynman. These are now recognized to be identical to the point-like 'quarks' whose existence had been postulated through a completely different line of reasoning; this was to explain the great number of different particle states that had been discovered by the high energy physicist, the systematic relations among their masses, and the rules that govern their conversions into one another.

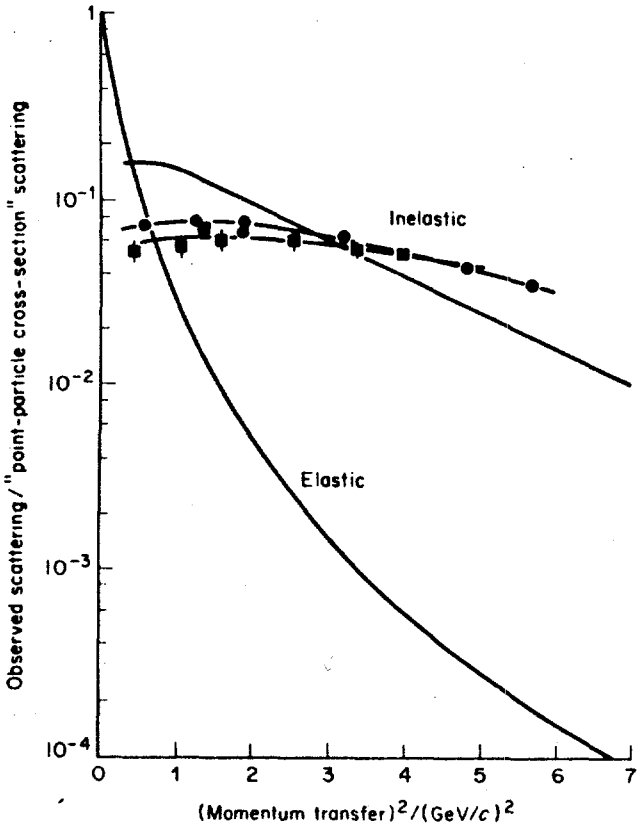

Fig. 11. Comparison of inelastic and elastic electron–proton scattering. This graph shows the ratio of the observed scattering to the point particle cross-section for elastic and inelastic events plotted as a function of the square of the momentum transfer. Note that inelastic scattering falls off much more slowly with momentum transfer than does elastic scattering. This indicates that inelastic scattering appears to take place on point-like objects within the proton.

More careful examination of the inelastic scattering process has shown that scaling is not exact. This conclusion can be drawn both from the highly precise experiments at SLAC using electron energies up to 20 GeV or so, and from the less precise experiments using beams of higher energy muons at Fermilab near Chicago, and at CERN in Geneva.

Experiments that measure the deviation from scaling constitute a valuable tool for examining the nature of the forces among the partons or quarks. According to modern concepts, these quarks interact through the exchange of certain objects called 'gluons'. The interaction between these particles and the quarks is measured through a coupling constant whose strength in turn determines the deviation of the data from the ideal scaling relationship, which is derived assuming the quarks to be 'free'. Thus the deviation from scaling measures the strength and character of this basic force.

Note here the analogy with the X-ray experiment that I described earlier. The first experiments of Compton demonstrated that X-ray photons indeed bounced off the electron component of the atom, which behaved nearly as if they were 'free'. The more refined experiments of DuMond and collaborators showed *deviations* from pure 'free' Compton scattering, and thus gave evidence of the dynamics of the atom, in particular the momentum distribution of electrons, in turn derived from the strength of the interaction between the electron and the nucleus. In the recent experiments, the existence of scaling as such indicates the existence of partons or quarks, but it is the *deviation* from ideal scaling that gives valuable evidence of the dynamics which bind the quarks together to form a proton or neutron. In the case of the atom, the basic forces involved are purely electromagnetic, and they can be described by what is now known as Quantum Electrodynamics. This theory is well understood, and its validity has been demonstrated experimentally over distances ranging from many Earth diameters down to about 10^{-18} m. The analogous theory of the strong forces acting between quarks and carried by gluons is known as Quantum Chromodynamics: it is now being formulated, although many outstanding questions remain to be answered.

However, irrespective of the nature of these forces, the basic 'rules of the game', namely quantum mechanics and relativity, have continued to apply throughout. I will now turn to a similar story in regard to the study of the simplest possible bound-state systems which consist of just two particles: a particle and its antiparticle.

5. Particle–antiparticle bound states

One of the predictions of relativistic quantum mechanics is that for each of the particles found in nature there should also be a corresponding antiparticle, that is, a particle with not only its electric charge, but also certain of its other characteristics reversed. This, in turn, led to the expectation that a charged particle and its antiparticle could combine to form a quasi-stable system, essentially a planetary system in miniature. Specifically, it was predicted soon after the discovery of the anti-electron, or positron, that there should be an entity now known as 'positronium', which is the bound system of an electron and a positron. Similarly, again looking ahead many decades, if indeed quarks are fundamental constituents of nuclear matter, or more accurately of all hadronic matter, then there should also be 'quarkonium' systems of various kinds consisting of a quark and an antiquark. Such objects are the simplest bound systems one can imagine. One would therefore presume that the positronium system constitutes an ideal test object to examine the validity of the theory of electromagnetic forces, while the quarkonium systems might be a similar laboratory for

examination of the forces that govern the behaviour of the constituents of nuclei or other strongly interacting particles.

The first experiments on positronium were done in 1951 by Martin Deutsch of MIT. A very large volume of work on positronium has been done between that time and today. This work has resulted in a complete and accurate tabulation of many energy levels of positronium and extremely precise measurements of the transitions between them. Figure 12 shows a diagram of these energy levels and of the numbers which go with them. These numbers constitute one of the most fundamental tests of the validity of quantum electrodynamics. For instance, the transition energy between the two lowest levels of positronium in theory is given by 203·40 GHz with an uncertainty of $\pm 0{\cdot}01$, while the experimental measurements give 203·3870.

This type of spectacular agreement is characteristic of the many experiments that have demonstrated that, as long as only electromagnetic forces are involved, quantum electrodynamics fully explains quantitatively all the observed phenomena.

The most sensitive tests of the validity of quantum electrodynamics at the smallest distances of interaction come from experiments at the highest energies. If electrons and their antiparticles are not bound together as they are in positronium, but rather collide

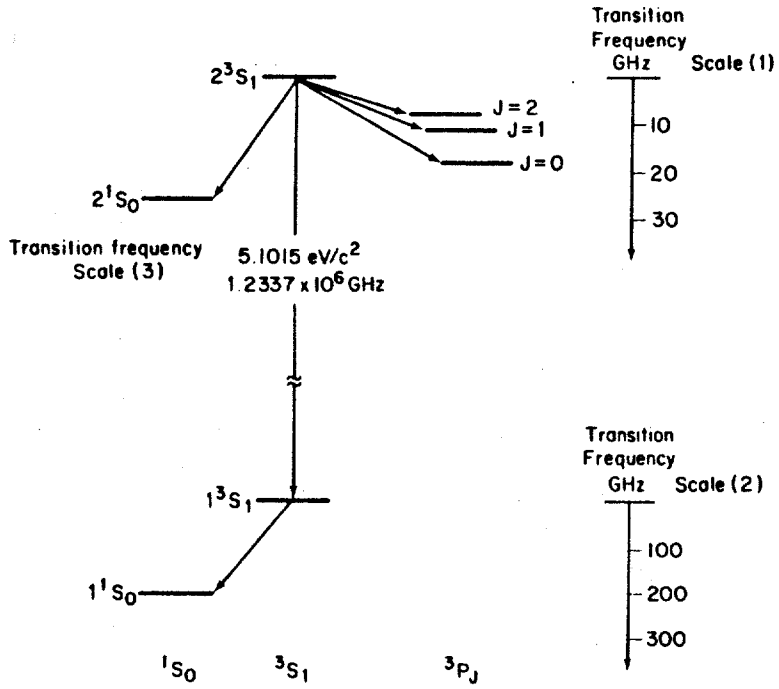

Fig. 12. Energy level spectrum of positronium, using the ordinary spectroscopic term designation giving spin multiplicity, orbital angular momentum and spin angular momentum, respectively. The spectrum is plotted in three different scales: (1) one scale shown in the upper right-hand corner corresponding to the transitions among levels of principal quantum number two; (2) a second scale shown in the lower right-hand corner corresponding to the transition between levels of principal quantum number one; and (3) the transition between the energy levels of principal quantum number 1 and 2 is shown on a still different scale, since it is larger by 5 orders of magnitude than the transitions within each principal quantum number.

at the highest energies available in the laboratory, then a number of things can happen, as indicated in fig. 13. The best way to produce such collisions is in high-energy electron–positron storage rings. Figure 14 shows an example of such an installation at SLAC.

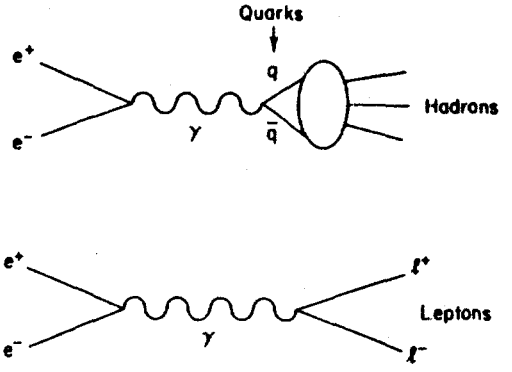

Fig. 13. Diagrammatic representation of electron–positron annihilation resulting in final hadrons (strongly interacting particles) in the upper diagram and final lepton pairs (not subject to the strong interaction) in the lower diagram. Since hadrons are composed of mixtures of quarks, the primary process in the upper diagram is creation of quark pairs via electromagnetic interaction carried by the intermediate photon γ. Hadrons are then produced subsequently through forces between the quarks. In contrast, in the lower diagram it is shown that leptons are produced through a purely electromagnetic process.

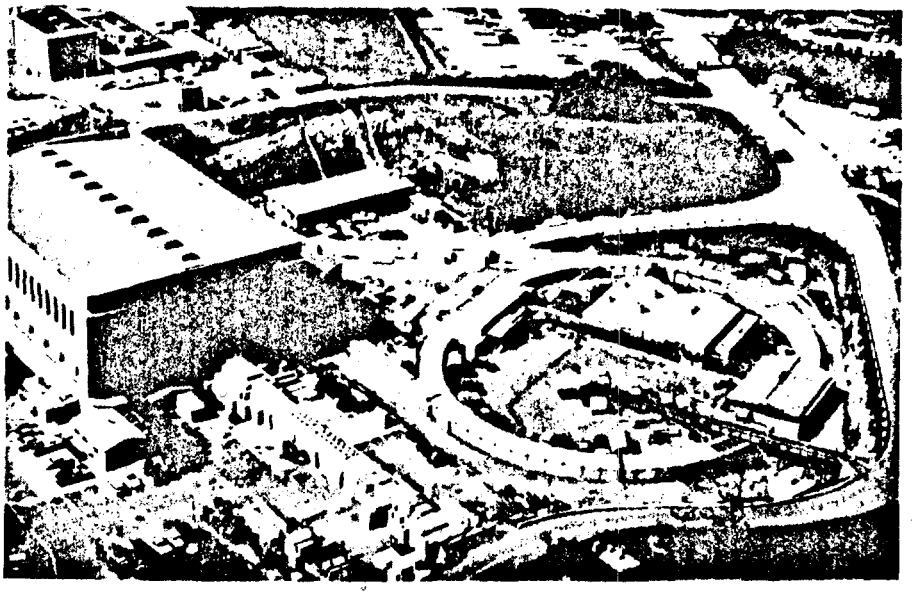

Fig. 14. Aerial photograph of the housing containing the SLAC electron–positron colliding beam storage ring SPEAR. The two buildings lying within the ring house the apparatus for detecting the results of electron–positron annihilation.

6. Digression: the growth of accelerators

Let me digress here to give some indication of how the enormous gap between the early low energy experiments of Rutherford, Compton, DuMond and Deutsch and the recent high energy experiments on electron scattering and electron–positron collisions has been bridged. I am here emphasizing the similarity in concept but the dissimilarity in scale between the early and recent experiments. The actual *pace* of this progression has been defined by the evolution of the technology of high energy accelerators and colliders. Figure 15 is an update of a chart, originally due to S. Livingston, which shows how the energy available through the use of accelerators has evolved over time. The pattern is indeed dramatic: the energy of accelerators has increased by a factor of 10 approximately every 7 years ever since the 1930s. This has been achieved not simply by building larger and larger accelerators of a single type, but rather by a succession of new technologies which were invented whenever an old technology became saturated in its ability to reach higher energy. Thus the growth in size and cost of accelerators has not

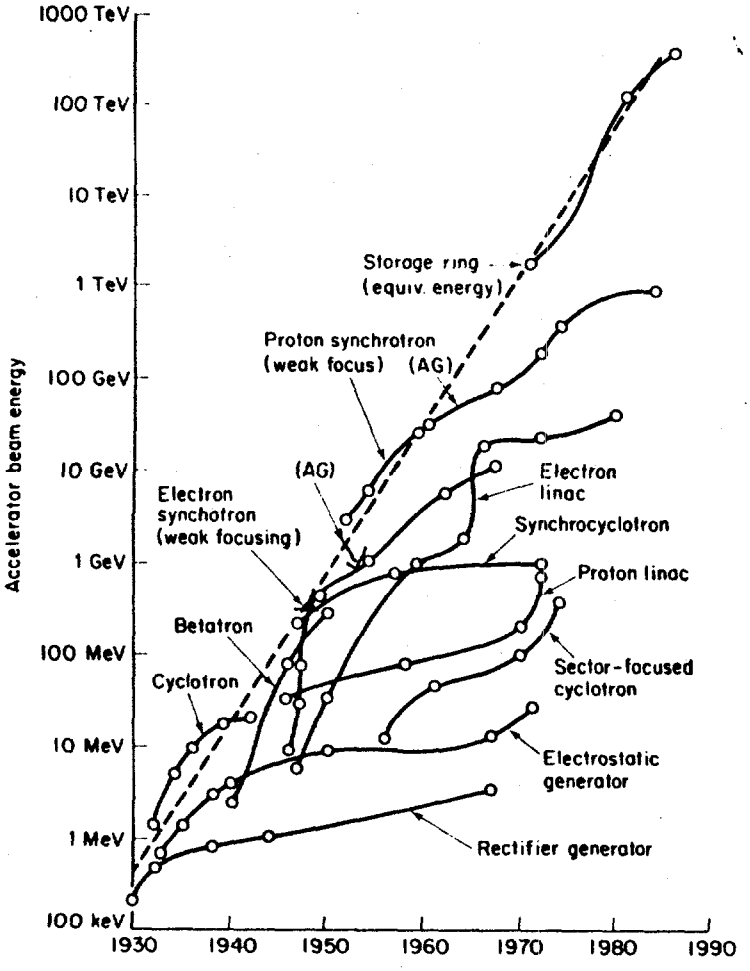

Fig. 15. Progress of the energy attainable through particle accelerators and storage rings. This growth is achieved through a succession of machines employing different technologies.

been anywhere near as large as the increase in scientific potential. I cannot describe here in any great detail the succession of inventions and technological advances which have made this evolution possible. Let me just name a few.

The early accelerators, and the present day machines operating at low energy, are electrostatic. The maximum energy attainable by such methods is thus limited by the technology of producing and maintaining extremely high voltages (some millions of volts). This limitation was removed by the invention of the cyclotron by E. O. Lawrence. Here the accelerating voltage is supplied by a high frequency power source, and the same gap is traversed over and over again by confining the particles in a spiral orbit through the use of a large electromagnet. The geometry is arranged so that synchronism is achieved between the time that the particles transit the gap and the crest of the alternating voltage. A similar idea was also proposed to accelerate particles in a straight line, first by Ising in Sweden in 1926, but it was not applied to very high energy acceleration until after World War II.

The cyclotron had its limits in that synchronism between the particle transit and the alternating voltage could not be maintained once particle speeds approached a fraction of that of light. In that case relativity limits the particle speed, while the radii of the particle orbits in a magnet continue to increase with increasing particle energy. In addition to achieving particle synchronism with the accelerating voltage, it is also necessary to focus the particle orbits so that in the long trajectory from start to target, only a minimum of particles is lost. The struggle to achieve both synchronism with an alternating field and particle focusing at the same time was attacked by several ingenious ideas; the simple cyclotron does not permit this at particle energies beyond about 20 MeV. One idea, that of L. H. Thomas, was proposed in 1938, but was then understood by only few physicists. Thomas proposed an azimuthal variation in the field of the circular magnet of Lawrence in order to achieve focusing and synchronism simultaneously. The Thomas idea was not applied extensively until after the war, and then only to machines of moderate energy.

The second idea, which led to much higher energies, was independently conceived in 1945 by Vexler in the Soviet Union and McMillan in the United States. Vexler and McMillan showed that if particles were not accelerated at the crest of an alternating field, but rather during either the rising or falling part of that field, then dynamic conditions could be achieved where the phase of acceleration of the particles would automatically be maintained. Those particles which are not accelerated by the ideal amount will arrive either early or late. Depending on the geometrical configuration and the energy, stable acceleration is achieved by an appropriate choice of transit time across the accelerating gap during either rising or falling voltages. Particle focusing can then be attained in a manner less tightly related to the conditions of acceleration. This 'phase stability' principle led to a variety of machines, including the synchrocyclotrons, electron synchrotrons and proton synchrotrons of the post-World War II era.

The next essential technological advance was the invention of the 'strong focusing' principle in 1950 by Nicholas Cristofilos in Greece, and then independently in 1952 by Courant, Snyder and Livingston in the United States. This principle made it possible to focus particle beams with magnets of much smaller aperture, and therefore led to a major reduction in the cost of particle accelerators. The combination of this principle with the phase stability principle is the basis of all large circular accelerators of today.

As these techniques approach their limits, attention has begun to shift from beams striking stationary targets, which made possible the type of experiments such as the scattering measurements I described earlier, to colliding beam devices, or colliders as

they are called today. What counts in high energy collisions is the *energy of collision* observed in that frame of reference in which the total centre of mass of the colliding particles is at rest. When a beam of particles strikes a stationary target, then at least half of the energy of the incident particle is used to move the centre of mass of the two colliding particles ahead, while only the remaining fraction of the energy is available for the collision itself. According to the laws of relativistic collisions, this useful fraction decreases continuously as the energy of the incident beam becomes larger. For example, the energy of collision of a 500 GeV proton striking a stationary target is only 31 GeV. If, on the other hand, two 500 GeV particles undergo a head-on collision, then the total collision energy available is 1000 GeV.

This enormous advantage in collision energy is somewhat offset by the fact that the density of practical particle beams is very much lower than that of ordinary matter. Thus while colliding beams produce dramatic gains in collision *energy*, they lead to large losses in collision *rate*. However, one cannot have everything, and the loss in collision rate can to some extent be compensated by surrounding the points of collision with detectors sufficiently large to catch almost all the fragments from the events which do occur.

I have digressed to this brief outline of the evolution of accelerating devices during the past fifty years in order to emphasize that, while our studies of the fundamental nature of matter have focused on structures of smaller and smaller size, the actual rate of progress in this field has been determined by the much more mundane matter of successive accelerator technologies. However, as with all exponential growth patterns, the dramatic evolution shown in fig. 15 must sooner or later slow down.

Among the colliding beam devices, the most productive has been the electron–positron collider. An example of such a device was shown in fig. 14. Here electrons and positrons counter-rotate in a ring, and the orbits of these particles are confined by a group of electromagnets. The more powerful recent electron–positron storage rings all use strong focusing magnets; in addition, large radiofrequency power systems are required to compensate for the loss of energy caused by the electromagnetic radiation emitted by the electrons as they traverse their circular orbits.

The advantage of carrying out high energy physics experiments with electron–positron collisions is the inherent simplicity of the annihilation process. As was indicated in fig. 13, the electrons and positrons indeed annihilate, resulting in what physicists call a virtual photon, which describes a state of pure electromagnetic energy. This electromagnetic energy in turn can rematerialize into any combination of particles which conserves both the energy involved and certain other of the symmetry characteristics of the initial collision. Since, unlike collisions in which protons strike material targets, the initial particles have completely disappeared, the final state can be particularly simple and therefore relatively easy to analyse. In particular, the final state can be pairs of particles and their antiparticles, whatever these particles may be. This, in retrospect simple, situation led to the spectacular discoveries of November 1974, aptly called the November Revolution.

7. The November Revolution

The November Revolution began with the simultaneous publication of two experimental results. The work of Samuel Ting and collaborators at the Brookhaven National Laboratory demonstrated that pairs of electrons produced from a beryllium target bombarded by 25 GeV protons in the Alternating Gradient Synchrotron exhibited a peculiar distribution. The correlation in angle and energy of the electron

pairs can be expressed as an effective mass of a conjectured object which, when disintegrating into pairs of electrons, would give rise to the observed distribution. The distribution in effective mass as observed in the Brookhaven experiments exhibited a sharp peak, as shown in fig. 16, located at an energy of 3·1 GeV. At the same time, experiments carried out by Burton Richter and collaborators with the SPEAR storage ring at SLAC showed that the yield of many kinds of particles, including electrons, muons, pi mesons, neutrons and protons, showed a sharp peak when the total energy of collision of electrons and positrons was near the same energy of 3·1 GeV. Figure 17 shows these observations.

As future developments indicated, the storage ring methods proved to be more powerful for investigating these new phenomena, and I will therefore restrict the following discussion to the results obtained with electron–positron storage rings subsequent to the original discoveries.

What made these discoveries a 'revolution'? There was little expectation that energy distributions of any kind at these high collision energies could exhibit such a narrow peaked structure. The reason for this assumption, in retrospect naive, is quite simple. If new particles or objects are formed at these high energies, then there is plenty of energy available for them to decay rapidly into many of the lighter particles that are already known. The greater the energy available for such disintegrations, the shorter is the life of the parent particle. And the shorter the lifetime of a new particle, the less well-defined its mass or energy would have to be according to the energy-time formulation of the uncertainty principle which is usually written as $\Delta E \Delta t \simeq h/2\pi$. Thus the only way the sharpness of these peaks could be explained is by postulating a new 'selection rule', that is, some new physical principle that inhibits the rate of disintegration. In turn, a new

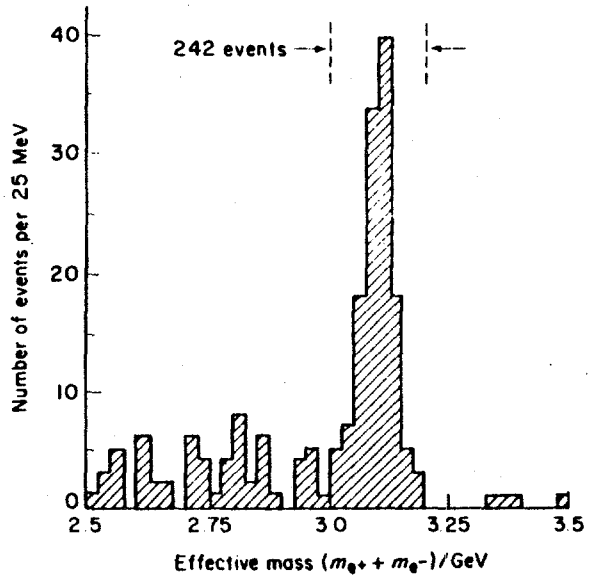

Fig. 16. Distribution in the *effective mass* of electrons and positrons produced at the Alternating Gradient Synchrotron at Brookhaven National Laboratory through impact of a high energy proton beam on a beryllium target as observed by Sam Ting and collaborators.

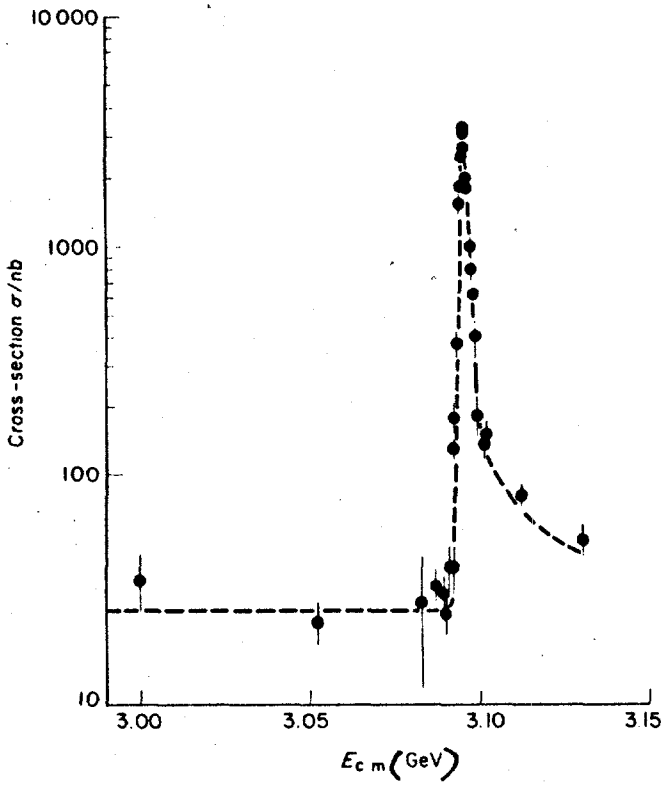

Fig. 17. Cross-section σ for the annihilation of high energy electrons and positrons resulting in the formation of hadrons, plotted as a function of the collision energy $E_{c.m.}$ of the electrons and positrons, as observed by Richter and collaborators at the SPEAR storage ring at SLAC.

selection rule has to be based on the existence of a new 'quantum number' or characteristic property of the particle involved. In short, the November Revolution signalled the discovery of a new kind of fundamental object in Nature.

Whenever a label or quantum number changes in some way, then, depending on the dynamics in question, the rates of processes can be inhibited; if nothing changes, then only the available disintegration energy and the masses of the particles involved determine how fast reactions will 'go'. Starting from this hypothesis, consensus developed rapidly that the new peaks at 3·1 GeV observed at Brookhaven and SLAC must be states of a new kind of quarkonium, that is, objects composed of a new quark and its antiquark. The quark involved is a new example of that species, beyond the three quarks originally postulated by Zweig and Gell-Mann in 1964 to account for the multiplicity of particles then known†. The existence of a fourth quark, dubbed the 'charmed' quark, had been suspected before, but these new discoveries not only confirmed its existence but at the same time, as discussed above, provided a new laboratory in which the interaction between these new quarks could be investigated.

† A discussion of the quark model can be found in the article by F. E. Close in *Contemporary Physics*, 1979, **20**, 293.

8. Charmonium

These discoveries initiated an immediate hunt for the full spectroscopy of this new 'charmonium', in complete analogy with the spectroscopy of positronium that I outlined before. This hunt, using electron–positron storage rings, has proved eminently successful. I shall not describe here the work through which the various spectroscopic states of charmonium were discovered sequentially. There are still some gaps in that spectrum, and work is in progress to fill these. Some higher energy states belong to the same family as those of the 3·1 GeV state originally discovered, and they were found simply by raising the collision energy of electrons and positrons in the storage ring, and observing further peaks. Charmonium states of different quantum numbers were located by observing the products of transitions from the directly produced states to the new states. Figure 18 shows this total spectrum as we see it today, and this figure should be compared with fig. 12 which gave the corresponding spectrum for positronium.

Note that the difference in energy scales between these two spectroscopies is a factor of about one hundred million, yet the fundamental arrangement of the lines in the two spectra is strikingly similar. This surely indicates strongly that the basic rules of quantum mechanics, which determine the existence of such discrete states and their energy, are the same whether one deals with electron volts, millions of electron volts, or thousands of millions of electron volts. Note that the proportions in spacing among the

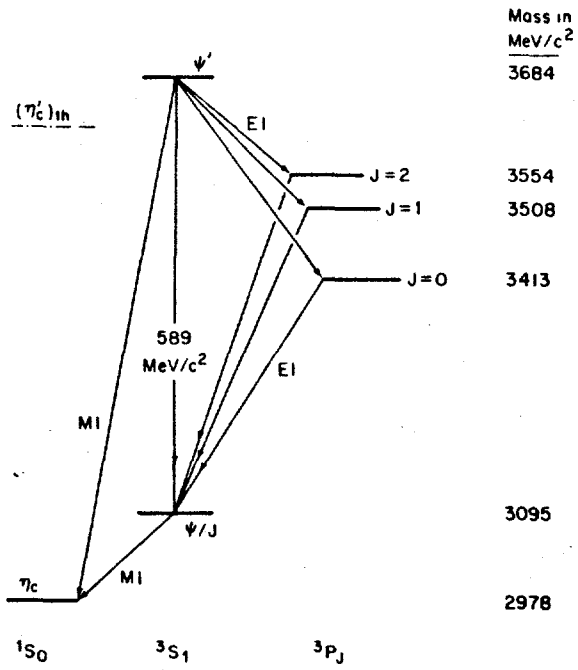

Fig. 18. The energy level spectrum of charmonium. Three classes of spectroscopic terms ($^1S_0$, $^3S_0$, ^3P_j) are shown using the usual spectroscopic notation for spin multiplicity, orbital angular momentum, and spin angular momentum. The observed masses in MeV/c^2 are shown on the right-hand side, and the observed transitions by arrows. The spectral lines in this classification are the same as those shown for positronium in fig. 12, but, the *scale* of the principal transition is different by a factor of about 100 million (589 MeV/c^2 relative to 5·1 eV/c^2). Moreover, the separations among the lines of differing angular momenta are very much larger than those for positronium.

lines are, however, greatly different. This is not unexpected, since the positronium spectrum is solely governed by the laws of electricity and magnetism, while the charmonium spectrum is governed by the not-as-yet fully understood laws of quantum chromodynamics; this is the name given to the rules which govern the forces between quarks as they exchange gluons with one another†. Naturally, many investigations have been carried out by theorists attempting to fit the new spectra under various assumptions drawing on the analogy with positronium. One exploits this analogy by comparing the simple inverse square law of the electromagnetic interaction that governs the positronium spectrum with the appropriate radial dependence of the force that can account for the charmonium spectrum. This quest has thus far been only partially successful.

9. Beyond charmonium

The spectroscopy of charmonium, analogous to that of positronium, is not the only result from electron–positron annihilation at high energies. History was again to repeat itself as accelerators and electron–positron storage rings moved to higher energy. At Fermilab a fifth member of the quark family, the b quark (b for bottom) was discovered. Subsequently, the combination of the b and anti-quarks—'bottomonium'—was observed in the form of peaking in the yield of various particles as a function of electron–positron annihilation energy at the storage ring DORIS in Hamburg, similar to earlier SPEAR observations for charmonium. Some of the more detailed spectroscopy of this object began to unfold at the Cornell storage ring CESR in the United States. Although these spectra are not as yet as complete as those from the charmed quark states, it is clear that we are again seeing the same basic quantum states as those in positronium. Interestingly, the utility of the spectroscopy as a laboratory to explore the forces acting between quarks increases as the masses of the quark pairs under study increase. The reason is, the heavier the quarks are, the slower they move in the 'quarkonium' combination, and therefore the simpler the theoretical description becomes because the effects of relativity need not be included.

Electron–positron annihilation gives a very direct signature on the number of quarks which contribute to hadronic matter as we know it. By 'hadronic matter' we mean the totality of those particles in Nature, including the neutron and proton and many others, that interact through the strong force; that is, the force we believe to be responsible for binding the neutron and proton in the nucleus. If we recall the diagram of fig. 13, we see that the initial interaction at the first vertex, in which the virtual photon first forms a pair of quarks, is purely electromagnetic. In consequence, the theoretical calculations giving the probability of these reactions is simply proportional to the sum of the squares of the electric charges of each possible quark that could contribute to the process being investigated. As the energy of electron–positron annihilation increases, presumably more and more kinds of quark pairs can be produced, depending on the masses of the quarks and the energy of the annihilation.

Figure 19 describes the present picture of the probability of annihilation of electrons and positrons as a function of energy. What we see are essentially two distinct plateaus separated by a transition region. The arithmetic shows that the first plateau corresponds, at least fairly closely, to three 'flavours' of quarks, while the second plateau corresponds to the addition of a fourth, the charmed, quark flavour. Thus this

† The ideas of quantum chromodynamics and the quark–gluon model are also explained in the article by F. E. Close referred to in the previous footnote.

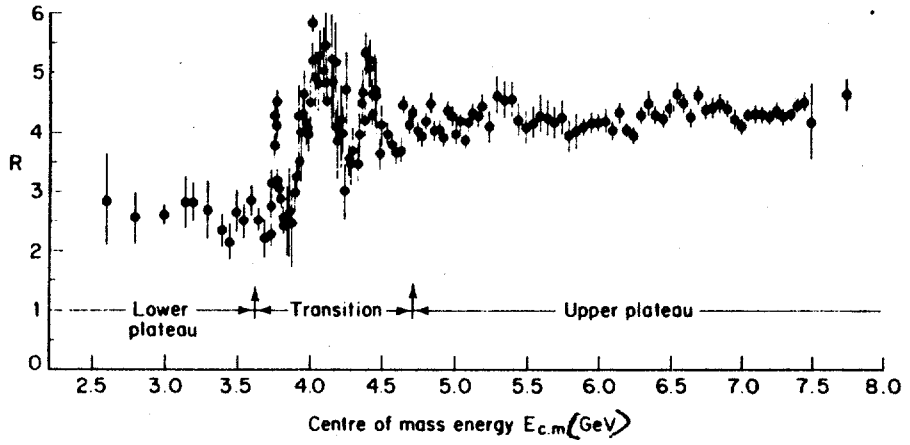

Fig. 19. The ratio R of formation of hadrons to that of lepton pairs resulting from the annihilation of electrons and positrons, plotted as a function of the energy of the colliding electrons and positrons.

general picture confirms the interpretation of the peaks, which gave the initial incentive to the November Revolution, as being pairs of charmed quarks. Again, as in some of the previous examples, detailed quantitative comparison of this simple picture with the data shows some discrepancies; and again we can learn from reconciling these differences more about the detailed dynamics of the processes involved. Thus, just as the deviation from the simple scaling relationship has taught us something about the strength of the interactions among the quarks in the proton, so too can we learn from the discrepancies between the exact numerical predictions concerning annihilation cross-sections and the experimental results how quarks interact with one another once they are formed.

The story I have described is, of course, only a small fragment of the evolution of our understanding of particle substructure during the century. What I hope, however, is that this sketch makes clear that we are witnessing an amazing contrast between the surprising new discoveries paced by the evolution of technology, on the one hand, and a consistency of general approach and motivation and of overriding physical principles, on the other. It is this contrast between sameness and change that has made this period one of the most exciting epochs in physical discovery.

DISCLAIMER