SLAC-PUB-8640
September 2000

The Discovery of the Point-Like Structure of Matter

Professor R.E. Taylor
Stanford Linear Accelerator Center,
Stanford University, Stanford, CA 94309

Invited talk presented at the
Discussion Meeting on the Quark Structure of Matter,
Royal Society of London, England, May 24 – 25, 2000

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

The organizers of this workshop have invited me here to reminisce. The assigned subject is the proton and how it lost its identity as an elementary particle. In its youth, the proton was very much neglected. It was overweight and introverted, and all the attention went to its lighter and more gregarious companion, the electron. The electron was noticed first and was accepted as a constituent of all matter almost immediately. As a result, the chemical "elements" lost *their* elementary status. With Rutherford's discovery of the nuclear atom it became clear that there was something rather small inside the hydrogen atom with nearly 2000 times the mass of the electron, and equal but opposite charge. That something was called the "positive electron" or "H-particle" until 1930 or so. The Standard Model in those days had only two elementary particles with mass (whether light quanta might also be a particle was a subject of debate) and the only known forces were electromagnetic and gravitational.

In the early days it was assumed that there were some extra positive electrons (each paired with a negative electron) inside nuclei other than hydrogen, to account for the observation that the atomic weight is equal to or greater than twice the atomic number.

In 1914, Rutherford's group at Manchester turned its attention to alpha-particle scattering experiments on light nuclei. The group was intrigued by a calculation predicting that forward-scattered H-particles would have a much greater range than the incoming alpha particles. An experiment, the very first on the proton, verified the prediction experimentally and Marsden and Lantsberry concluded that the Coulomb field of the H particle could account for their results (at distances of closest approach that approximated 10^{-13} cms.) World War I stopped most of the research in Rutherford's laboratory when many of the young scientists left to serve in the armed forces. Rutherford himself continued to do some research in parallel with his war work and in his spare time he discovered the first nuclear reaction on a nitrogen target along with anomalies in the scattering of alpha particles from hydrogen.

Much improved measurements on hydrogen came after the war when Chadwick and Bieler, (now with Rutherford at the Cavendish) redid the earlier experiments, finding that there were too many H particles at large angles when the distance of closest approach was less than 3.5×10^{-13} cms. In their 1921 paper, Chadwick and Bieler stated that there must be "forces of very great intensity" acting at small distances. Great significance was attached to the fact that such distances are about the same as the classical electron radius.

Compare the modest activity on the proton with the intense effort (both experimental and theoretical) on electrons after the war. Progress was swift and by 1929, the basics needed for understanding the atom were in place, although the nucleus was still not understood at all. Only the charge, mass and spin (but not the magnetic moment) of the proton were known. In 1920 Rutherford had suggested that combinations of positive and negative electrons in the nucleus formed a neutral entity where the "ordinary properties of the electrons are suppressed". By the end of the decade there was growing recognition of the problems inherent in assuming the presence of electrons in the nucleus though it still seemed obvious that they had to be in there somewhere.

With the great success of the Dirac equation there was hope that the proton might be connected with the "negative energy" solutions that carried positive charge but it soon became clear that whatever those solutions represented, it was not the proton. Heisenberg suggested that a new force was needed to explain the binding of protons inside the nucleus. Stern's measurement of the proton's magnetic moment was disturbing, showing that it would not be enough to give the proton a straightforward Dirac equation of its own.

With the discovery of the neutron in 1932, the proton was no longer the only heavy elementary particle and with the discovery of deuterium, studies of the force between a proton and a neutron began in earnest. With the availability of proton accelerators, proton-proton scattering became practical, and in 1935 experiments at Berkeley showed that p-p scattering did not follow the Mott formula. That experiment was the first of many hadron-hadron scattering experiments that became a major industry when the accelerator population exploded in the 1950s. Similar experiments have continued right up to the present day. There is a direct path from the studies of scattering to the work on dispersion relations and eventually to the bootstrap mechanism, Regge poles, etc.

The discovery of the neutron, followed by the neutrino hypothesis and the discovery of mesons in the cosmic rays, increased the number of particles that could be considered as elementary. When accelerators reached energies where new particles could be produced artificially, the catalogue of elementary particles grew rapidly. The spectrum of elementary particles discovered in the 1950s led to the SU_3 classification scheme and eventually to the quark model for the hadrons. Most of the elementary particle experiments in that decade were done at proton accelerators. Electron accelerators were poor cousins, few in number and usually far from the energy frontier. Electron synchrotrons did contribute photoproduction data in the early 1950s.

In 1949, Leonard Schiff wrote a technical report about possible experiments for the projected Mark III linac, an accelerator based on microwave technology developed at Stanford. Elastic scattering of electrons from the proton was at the top of his list of experiments. He says:

"The proton is now believed to possess a structure in the sense that it exists part of the time as a proton and part of the time as a neutron surrounded by a positive meson cloud."

and also:

"…it is to be expected that the magnetic field associated with the magnetic moment of the proton will influence the scattering"

Schiff mentions that the electron-proton scattering could perhaps be calculated in a consistently relativistic way, but that such a calculation would be very tedious (Unknown to him, Rosenbluth was making this laborious calculation at the time, and shortly thereafter, Feynman diagrams made such calculations much easier).

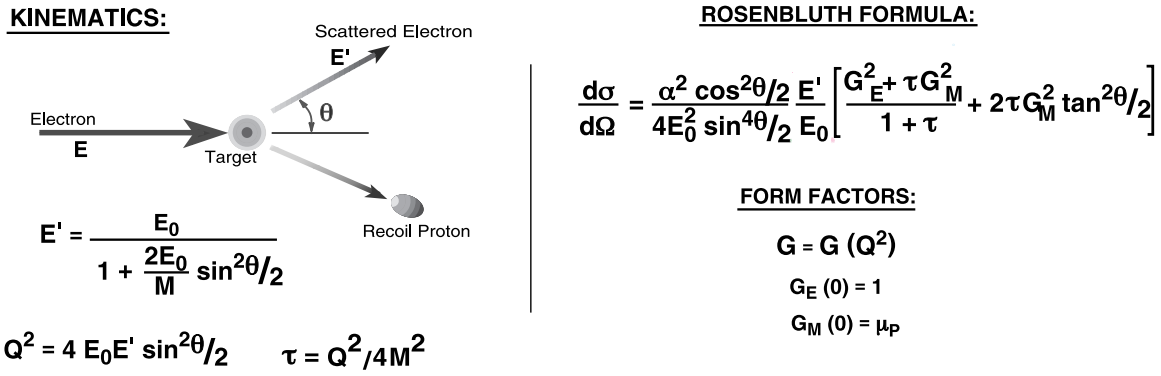

Figure 1 Kinematics and cross-section for elastic electron scattering.

It was also in 1949, that Schiff recruited Robert Hofstadter (his former colleague from the University of Pennsylvania) to Stanford. Hofstadter was already interested in using high-energy electrons as probes of nuclear charge distributions. Upon his arrival at Stanford, he and his group began to assemble apparatus at the halfway point of the new linac for the measurement of electron scattering from various targets. The setup was simple but effective:

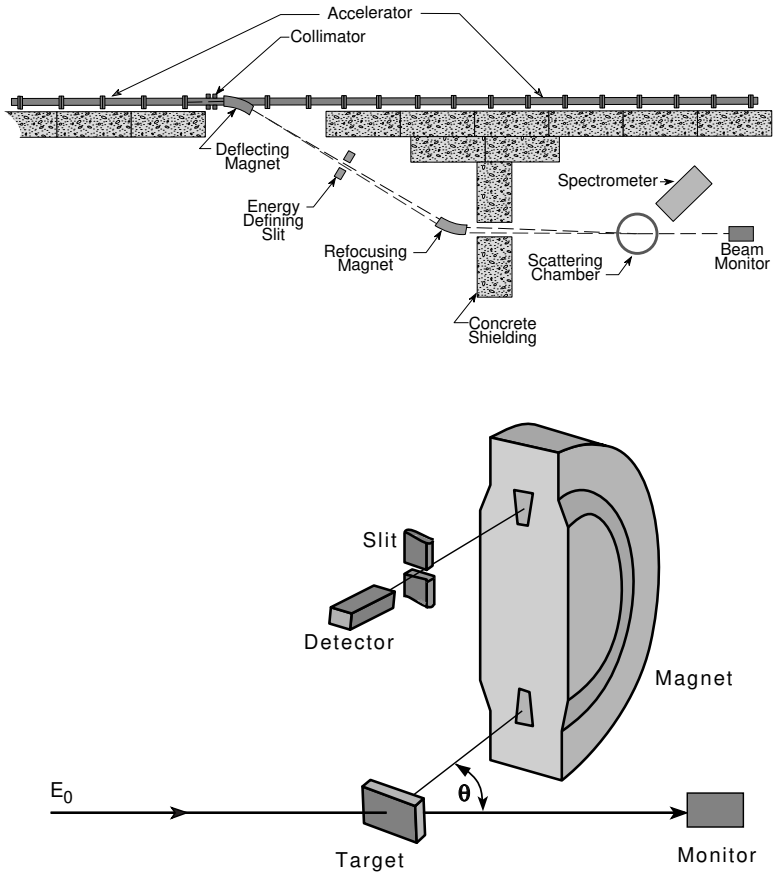

Figure 2 Experimental set-up for electron scattering experiments at the half way point of the Mark III accelerator at Stanford in the early 1950s.

The first electron scattering measurements at Stanford were made using carbon targets. Scattering from hydrogen was first observed using a CH2 target, although the first published data on hydrogen came from measurements using high-pressure gas targets. The early data indicated that the "proton was not a point", that there were contributions from magnetic scattering, and also that the magnetic and electrical sizes were comparable.

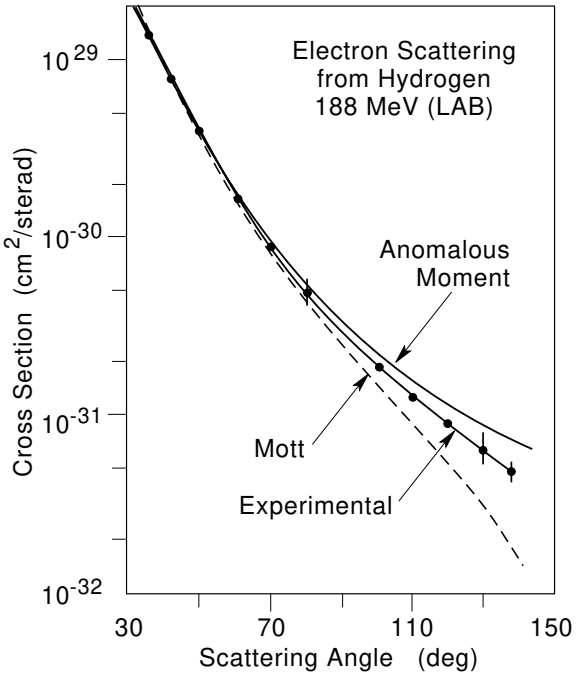

Figure 3 First published results on electron-proton elastic scattering measured at the Mark III accelerator at Stanford.

These very direct measurements of the proton's extended charge and magnetic moment distributions were a major event in high energy physics in the mid-1950s. Measurements on hydrogen (and deuterium) targets continued at energies up to 1 Gev. The electron community began to consider electron accelerators with even higher electron energies. By this time, it had been demonstrated (at Cornell and elsewhere) that scattering experiments could be performed at electron synchrotrons using internal targets or external beams. CEA, DESY and SLAC were soon under construction.

In the original proposal for SLAC (1957), electron scattering was mentioned as an extension of the successful experiments at Stanford's Mark III linac, and as experiments where violations of QED might be observed. In a 1960 summer study at SLAC, J. Cassels produced a more sophisticated analysis of electron scattering, finding that there might still be lots to learn from elastic scattering at SLAC energies.

In 1963 a collaboration of physicists from MIT, Caltech and SLAC began to think seriously about scattering experiments at SLAC, and the equipment that would be needed to make such measurements.

Many of those involved in the plans for experiments with electron and photon beams were alumni of the Mark III Laboratory. Panofsky, the director of SLAC, was very active, and led the SLAC contingent involved in the scattering experiments through the early experimental operations. Richter and his group were planning photoproduction experiments, and would share experimental facilities with the scattering group. The scattering group was small – about a dozen of us in the early years, and fewer than twenty physicists until well after the early inelastic experiments. Seven of the eight senior physicists in the group (senior = over 33 years old!) had worked at the Mark III accelerator on the Stanford campus.

The conventional picture of the proton was still that of an elementary particle surrounded by a cloud of mesons. Elastic scattering at SLAC might be able to demonstrate the existence of a proton "core", an idea that was still current at the time.

Excited states of the proton had been observed at both proton and electron accelerators and the production of such states by inelastic electron scattering would be interesting at SLAC. The 3,3 resonance had been observed in inelastic scattering at the Stanford Mark III linac, and Panofsky and Alton had investigated the radiative corrections that were more complex than those for elastic scattering.

For the electron scattering experiments we chose to concentrate on the detection of scattered electrons arising from interactions of the electron beam in the target.

The single arm inelastic kinematics are similar to the elastic kinematics, with the addition of one extra variable that depends on the energy needed to produce the final hadrons. The two inelastic structure functions can be separated using different angles and energies, just as in the elastic case.

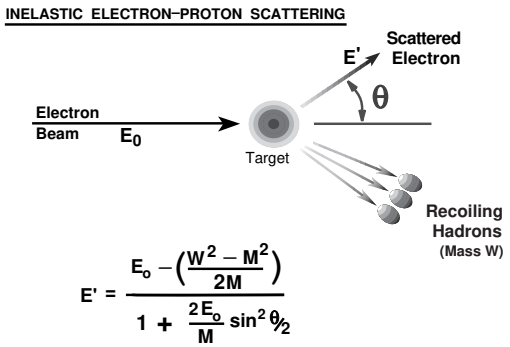

Figure 4 Kinematics and cross-sections for inelastic electron-hadron scattering.

There were two common ways of defining the inelastic form factors in the 1950s, and that history has left us using a mixture of parameters from the two expressions for the cross-section. In the end the virtual photon approach did not simplify the physics, but we still talk about R, the ratio of longitudinal to transverse virtual photons, along with the structure functions W_1 and W_2 .

In the initial planning of the SLAC experimental facilities, the kinematics and the estimates of cross-sections for elastic scattering and for photoproduction of pions provided the main design guidelines for the equipment. It was important to have sufficient resolution to cleanly separate states that differed by a single pion mass in photoproduction or to separate the excited states in inelastic scattering.

I will not describe the spectrometer facility we built in End Station A, since that has been done many times and details are available in the literature. We built three spectrometers capable of analyzing singly charged particles with moment of 20, 8, and 1.6 Gev/c. The solid angle acceptances were 0.1, 1, and 5 milli-steradians respectively. The scale of the devices was quite impressive for its day.

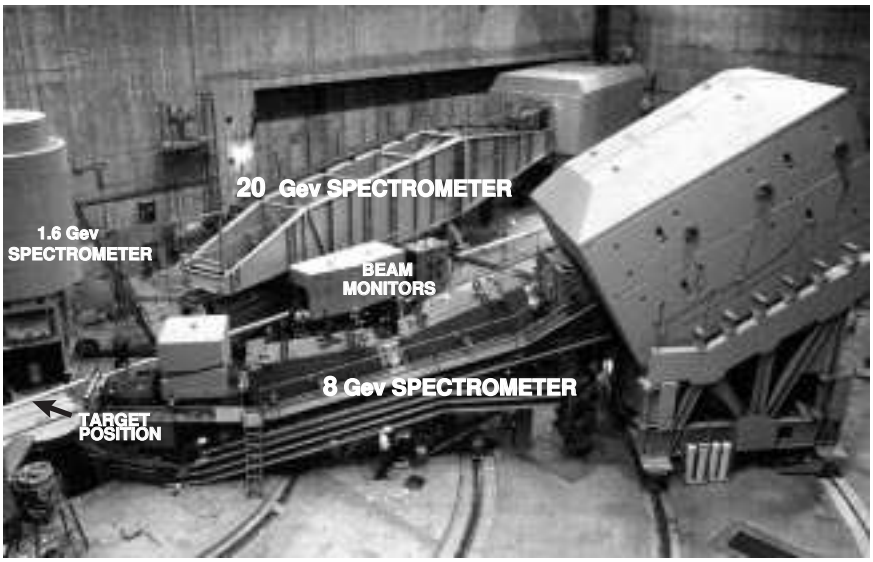

Figure 5 Spectrometer facility at the Stanford Linear Accelerator. Each of the Spectrometers can be rotated about the target position to vary the angle of scattering.

The basic design philosophy was conservative – SLAC was a very visible project, and it was important to the laboratory that reliable results be generated in the early running. It seemed unlikely that the experiments could be reproduced soon at another accelerator, so any wrong answers were likely to mislead for a long time. Beam time would be very costly at SLAC, so we tried hard to make the spectrometer complex efficient. This led us to incorporate a mid-sized computer dedicated to our data acquisition and on-line analysis. Our computer system became a model for those who could afford such things.

By 1966 SLAC and the spectrometer facility were nearly complete. Our first request for experimental beam time was for a study of elastic scattering. By this time CEA and DESY had extended the early measurements, finding that "the peach has no pit". The dipole form factor formula found by Hofstadter's group was still a very good approximation at the highest energies available at CEA and DESY.

Still, there was no dearth of theoretical predictions for the high energy behavior of the elastic form factors that might be tested at SLAC energies. In our first request for experimental beam time we listed several of these predictions:

(1) Existence of a nucleon core.

(2) Validity of the pole description of nucleon form factors.

(3) Validity of the Wu and Yang form factor, that $G \sim e^{-|q|/0.6}$.

(4) Hypothesis of Drell that $G > e^{-|q|/2M}$.

(5) Hypothesis of Sachs that $G_E = G_M$ at large q^2.

(6) Hypothesis that $F_2(q^2)$ and $F_1(q^2)/q^2$ asymptotically approach zero as q^2 becomes large.

Our first measurements on elastic scattering using the 8 Gev spectrometer confirmed the earlier CEA and DESY results – G_m followed the "dipole formula" quite closely out to the highest Q^2s available to us. It was disappointing but we had also proposed to study inelastic scattering to excited states and hoped for better luck there. Three or four such resonances had shown up in CEA and DESY inelastic experiments. We wanted to extend the measurements on those resonances and also search for the excitation of new resonances having higher invariant masses. Inelastic excitation to non-resonant states would be a background in such searches and was estimated to be quite small, following an early suggestion of Bjorken (well before his work on scaling, etc.). Measurements of excitation to the non-resonant states were mentioned in our proposal, but did not figure prominently in our plans. By this time the Caltech physicists had dropped out of the collaboration.

Things did not go smoothly on our first inelastic scattering run on the 20 Gev spectrometer that began in September of 1967. The online computer programs, which had worked quite well on the elastic runs, were now inadequate. To scan the inelastic spectrum we lowered the energy settings of the spectrometer in one percent steps but online calculations of the cross-sections measured on different steps didn't match. The analysis programs had to be rewritten as we continued to test the workings of the spectrometer hardware. After a couple of weeks of total chaos, things began to settle down. We took a full set of data at 6° with several initial energies and with scattered electron energies down to 2 – 3 Gev. The 6° data taking went fairly well, although there was a lot to do off-line to obtain a clean data set. Even on-line, it was clear that we were seeing at least three resonant peaks along with unexpectedly high yields in the deeply inelastic region.

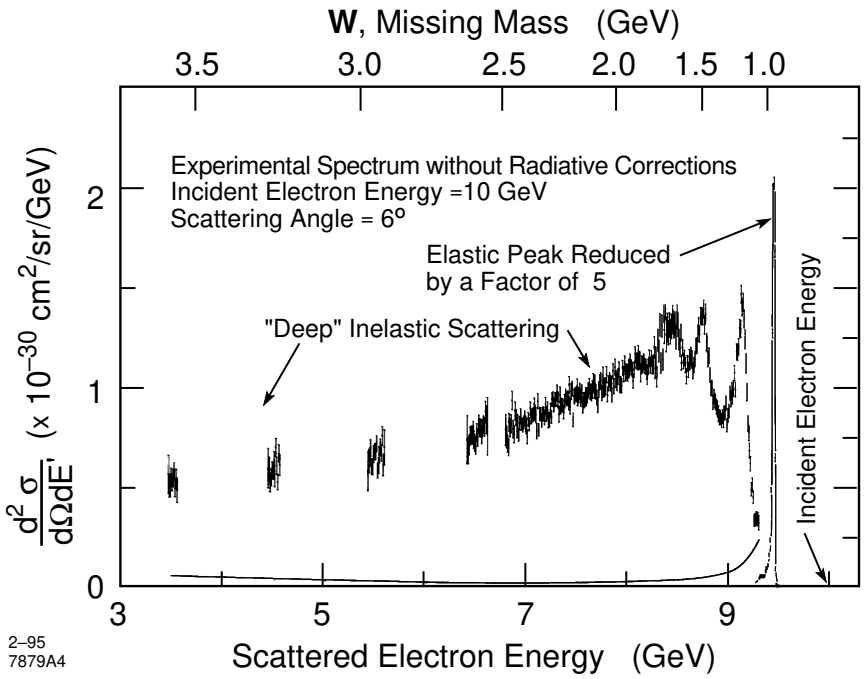

Figure 6 Early electron scattering cross-sections from the first inelastic scattering experiments performed at SLAC by the SLAC-MIT collaboration.

The raw cross-sections were corrected to take account of radiative processes that can occur before, during and after the scattering of the electrons. The corrections are quite complex, and were a significant concern in the analysis. Rough estimates had shown that such processes could account for as much as half of the raw cross-sections, so errors in the corrections could have big effects on our corrected cross-sections. We made two independent studies of the corrections, one at MIT and the other at SLAC (with help from Paul Tsai). The two calculations employed rather different techniques but gave very similar results and we became confident that there were indeed large deep inelastic cross-sections.

The year before this 6° data was taken, Bjorken's studies of electron-nucleon scattering using the tools of current algebra had already predicted that the total cross-section would be of the same order as the Mott cross-section, in the limit of infinite energy. We had asked if 16 Gev was likely to be near that limit, but had not been encouraged. In addition to being large, the measured cross-sections had a very different kind of Q^2 dependence than elastic scattering, as is shown in this early plot.

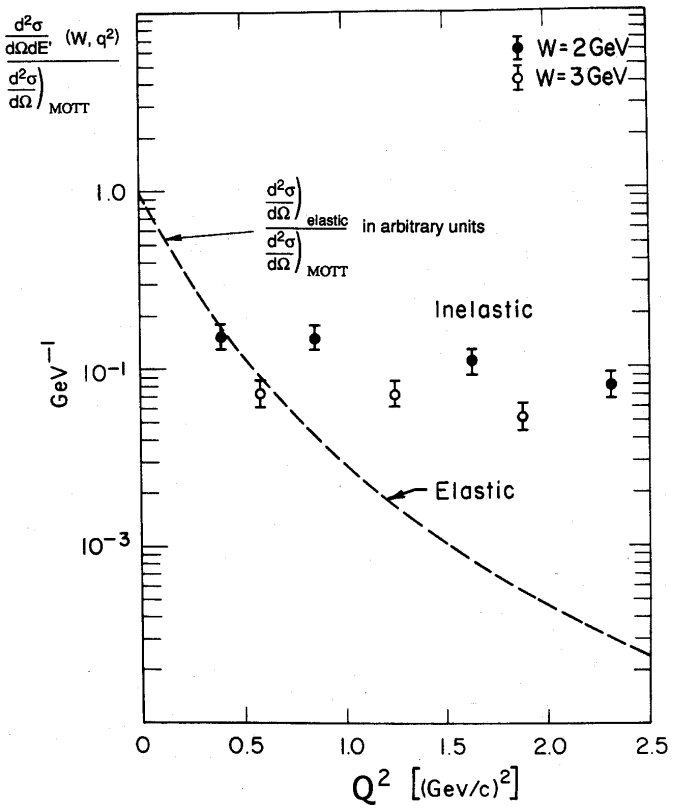

Figure 7 Early results on electron proton scattering, showing the striking difference in Q^2 dependence between elastic and inelastic cross-sections.

Bj had also suggested "scaling"– relating the cross-sections at different values of Q^2 and ν, but having the same value of the ratio, ν/Q^2.

After talking to Bjorken, Henry Kendall plotted the early 6° values of W_2 (assuming no contributions from longitudinal photons) against ν for different values of Q^2.

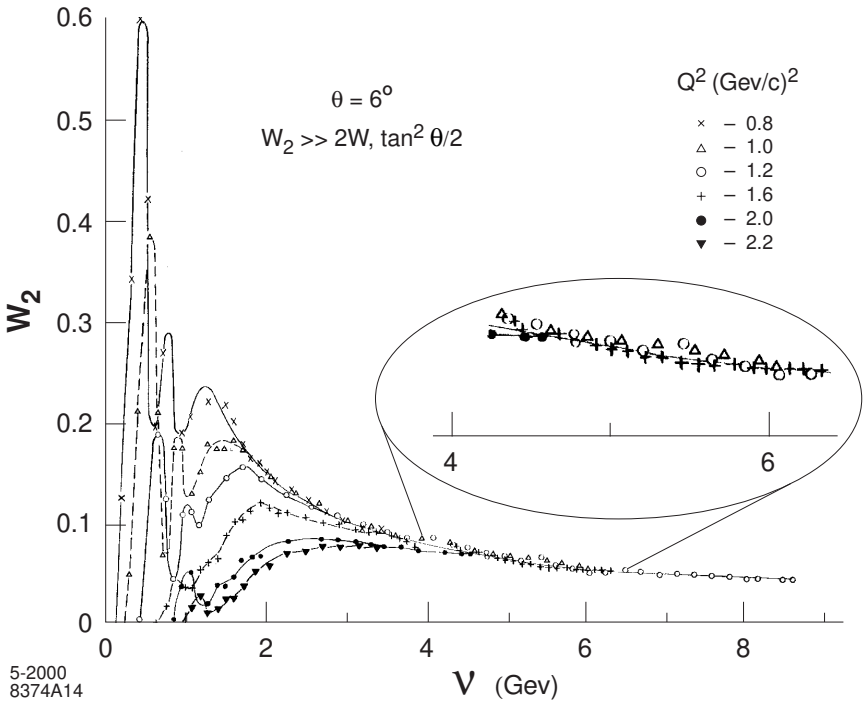

Figure 8 The measured form factor W_2 plotted against the energy loss ν. The magnified portion shows how W_2 depends on Q^2 over a range in ν.

For a small range of values of ν, the value of W_2 doesn't seem to vary with Q^2 (an even stronger constraint than scaling). Kendall prepared a second plot, this time of νW_2 for two cases, corresponding to the two extreme cases of the value of R.

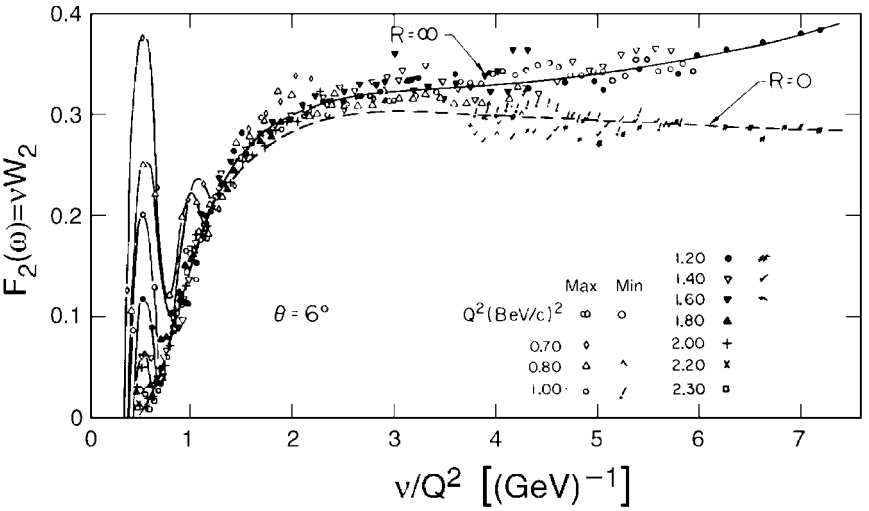

Figure 9 The measured proton structure function, F_2, plotted against ν/Q^2, showing approximate scaling beyond the resonance region (as predicted by Bjorken).

In this plot, the use of "ν/Q^2 "as the abscissa stretches out the interesting regions and makes the resonant region less prominent. The use of ν/Q^2 was quickly superseded by the closely related dimensionless variable $\omega = 2M\nu/Q^2$. The experimenters were somewhat divided over the strength of the "confirmation" of scaling. With such a small amount of data and the limited range of Q^2, one could easily fit the data with some sort of "generalized" vector dominance. Data at larger angles were needed to measure the value of R.

The fairly strong evidence for large cross-sections, and the rather weak but somehow impressive indications of scaling had an immediate effect. Feynman was introduced to the data in the summer of 1968 and quickly produced the "parton" model, a model that assumed that the electrons were elastically scattered from point-like bits of charge in the protons.

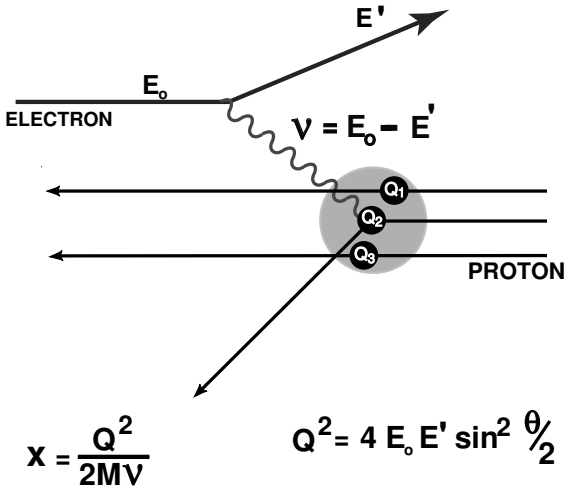

Figure 10 The parton picture of inelastic scattering. The inelastic scattering is treated as *elastic* scattering from small "parts" of the proton.

His model gave a simple physical picture of scaling and a physical interpretation for x ($=1/\omega$) as the fraction of the proton momentum carried by the struck parton.

Bjorken was perfectly aware that scattering from point-like objects in the proton would exhibit scaling, as can be seen in his paper at the 1967 Electron Photon Conference. Feynman once said to me (apropos of scaling, etc.) "I have done nothing that was not already in Bj's notebooks." Bjorken does not agree. For the experimenters, Feynman's sudden enthusiasm for deep inelastic scattering was both gratifying and helpful.

Jerry Friedman presented our early results at the 1968 conference in Vienna, at a parallel session attended by me and two or three others. W.K.H. Panofsky was a plenary rapporteur, and he presented the two scaling curves drawn up by Kendall and mentioned the possibility of "point-like" objects in the proton. Various groups began to share Bj's enthusiasm for our data. Theory groups at Oxford, Harvard, MIT, Moscow, Princeton, and Caltech joined in with the theorists at SLAC and all went to work in earnest. Most of the members of these groups believed that

Feynman's partons were the long sought quarks, and eagerly awaited confirming data. We were already scheduled for more complete measurements at 6° and 10° and had started serious planning for the experiments at larger angles that would allow the separation of W_1 and W_2.

By the time of the 1969 Electron-Photon Conference in Liverpool, the scaling behavior was much better established. The new 6° and 10° data were more accurate and more complete than in the previous runs. A qualitative view of the 10° cross-sections shows the behavior as Q^2 increases.

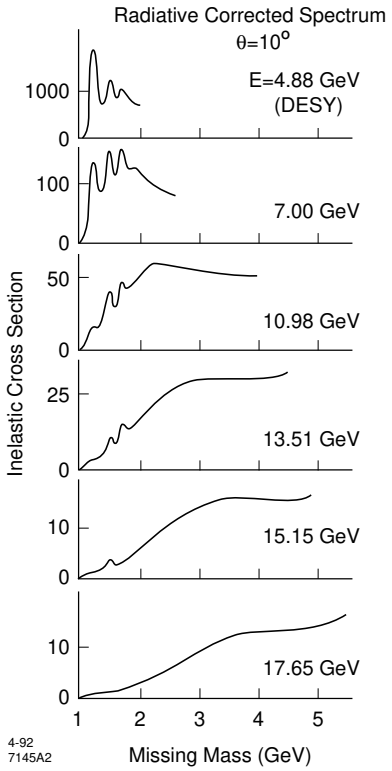

Figure 11 Behavior of the electron-proton inelastic cross-section at a scattering angle of 10° for various initial electron energies, illustrating the shift from a photoproduction-like spectrum to quasi-elastic scattering from constituents as the incoming energy increases.

These graphs still assumed that effects due to longitudinal photons could be neglected. In the spring of 1969, DESY combined some of their backward scattering "single arm" data with our 6° and 10° data to show that σ_L/σ_T was small. At the Liverpool conference that summer we presented preliminary data from 18°, 26°, 34° which, taken together with the 6° and 10° data, also indicated small values of σ_L/σ_T. After this time, I didn't see any calculations based on Generalized Vector Meson Dominance that could fit the whole range of our inelastic data. This did not mean that there was any sort of mass conversion to belief in the quark model in the theoretical community.

Comparing e_p and e_n scattering constituted a further test of the quark model because of the different quark charges in the two nucleons. In the most naive quark model the scattering should depend directly on the sum of the squares of the quark changes in each nucleon, so neutron scattering would be 2/3 of that from a proton.

We repeated our experiments at 6° and 10°, this time measuring both hydrogen and deuterium cross-sections. The deuterium structure functions exhibited "scaling", but the ratio of deuterium to hydrogen cross-sections varied with ω. We submitted plots of D/H -1 (an approximation for the neutron cross-section) to the 1970 ICHEP Conference in Kiev, showing that the neutron cross-sections were different from those of the proton.

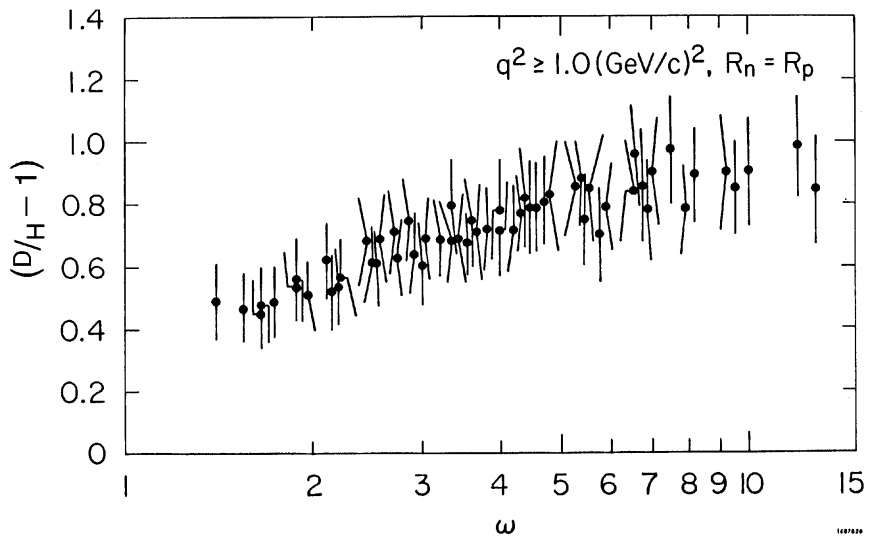

Figure 12 The ratio of deuteron to proton cross-sections (-1), showing that the deuteron scatters less than twice as many electrons as the proton for small values of ω.

The ratio of D/H-1 was close to 1 at very high ω (x = 0), supporting the suggestions of scattering from a quark-anti-quark sea as well as from the three "valence" quarks.

That both neutron and proton cross-sections were smaller than in a simple three-quark model (and also less than expected in the more general Callen - Gross sum rule) was addressed by Kuti and Weisskopf, who suggested that gluons in the nucleus were carrying some of the proton's momentum.

By this time there was much more data and scaling could be tested over a much greater range in Q^2. To include the largest range of Q^2 in the plot we chose data at ω = 4.

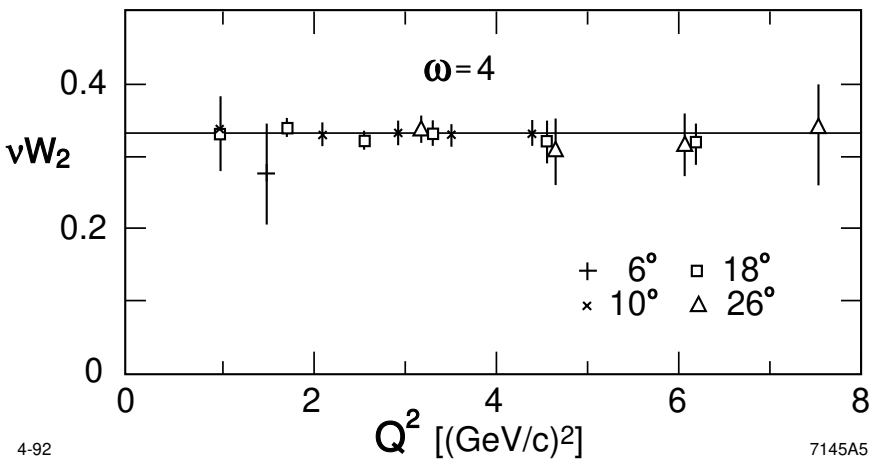

Figure 13 Values of νW_2 vs. Q^2 at $\omega = 4$, showing that νW_2 does not vary with Q^2, (ie. νW_2 "scales")

So things were in pretty good shape, but nothing is ever perfect. As the data improved, it became clear that scaling was not working over the full range of our data (it turns out that $\omega = 4$ is not a good place to look for scale breaking). At first, the scale breaking was observed only below W = 2.6 Gev, and we wondered if the "resonant region" was more extensive than we had assumed, even though we were seeing no visible peaks between 2 and 2.6 Gev. So for a while we made "scaling plots" including only data having W >2.6 Gev. Another solution was to use a slightly different scaling variable, $\omega' = W_2 / Q^2$ – (this variable is equivalent to $2M\nu/Q^2$ in the limit as ν, Q^2, and W go to infinity, so Bjorken's hypothesis was still valid). We could detect only minor deviations from scaling in ω', and used ω' in our presentations for a couple of years.

By the summer of 1971, most people were at least aware of our results and the quark-parton interpretation of our data. A growing number of theorists were hard at work, and soon the concepts of "asymptotic freedom" and "confinement" (sometimes called "infra-red slavery") led on to Quantum Chromodynamics. These advances actually predicted scale breaking, so we could go back to ω (or x) as the scaling variable.

It was in 1971 that the original SLAC-MIT collaboration split into two independent groups. There was, at that time, some disagreement about what to do next. Most of the SLAC contingent worked on an experiment at 4º (and at 58 – 60°), while all the MIT scientists and a couple of people from SLAC repeated the hydrogen measurements at 18º, 26º and 34º along with new measurements on deuterium. The breakup was so friendly that many people don't realize that it ever happened and make no distinction between SLAC-MIT and MIT-SLAC.

Late in 1969 we had heard rumors that an analysis of CERN neutrino data was indicating that the neutrino scattering cross-sections were proportional to the neutrino energy, as expected in the quark-parton picture. By 1972 a major independent confirmation of the quark model was announced by Don Perkins at the ICHEP conference in Batavia. The Gargamelle data showed

cross-sections that rose linearly with energy and also that the neutrino cross-sections were about three times the anti-neutrino cross-sections.

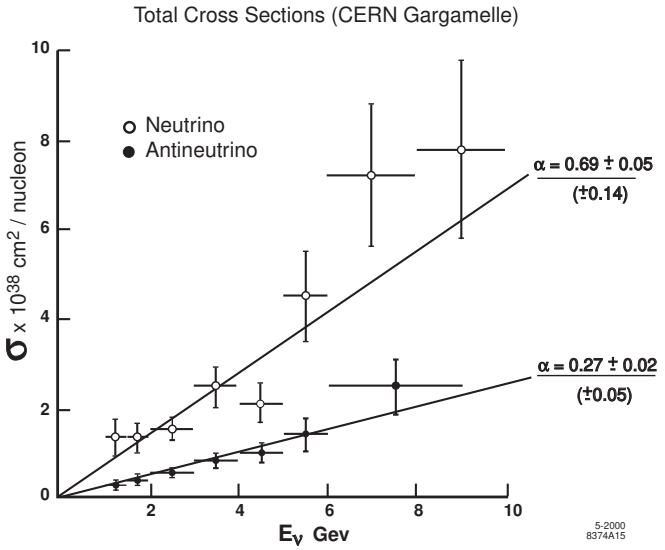

Figure 14 Neutrino and anti-neutrino total cross-sections as a function of energy, showing a linear rise as energy increases, and a ratio of neutrino to anti-neutrino cross-sections near 3, as expected in the quark models of inelastic scattering.

All this in marvelous agreement with the quark picture. The neutrino experiments could extract a structure function F_2, and in the quark model that structure function should be 5/18 of the corresponding F_2 derived from the electron experiments.

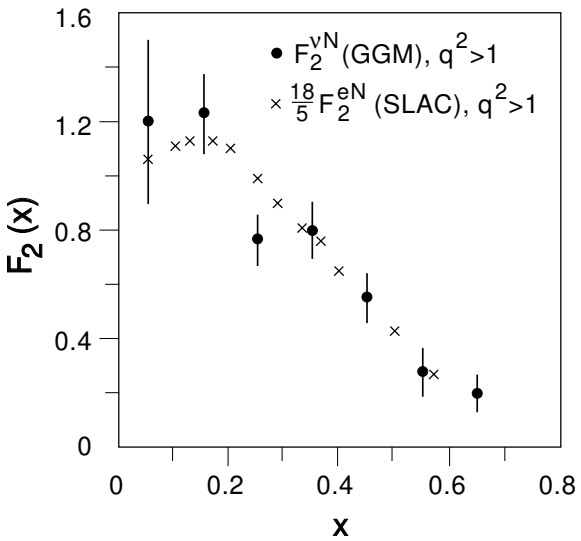

Figure 15 A comparison of structure functions, F_2, for neutrino and electron scattering demonstrating the power of the quark model that predicts a ratio of the structure functions equal to 18/5.

After this, I could sleep peacefully, secure in the belief that the quarks were surely there. This is a good place to stop this narrative. By 1972, the year that QCD first saw the light of day, a majority of the community was convinced that the proton was made of quarks and gluons. A few stubborn theorists would hold out until the discovery of the charmed quark in 1974.

The two groups at Stanford would continue with inelastic scattering experiments for two or three more years but our monopoly was ending, and the field would move on to higher energies and higher Q^2s. Soon, muon and neutrino experiments would vastly extend the range of the inelastic scattering data in ν and Q^2. The muon scattering experiments at Fermilab demonstrated scale breaking conclusively. Eventually HERA would reach Q^2s over three orders of magnitude greater than we could see at SLAC.

Looking back, it was during that first run at 6º that the point of discovery was reached. The large inelastic cross-sections were quite evident in that data, and the preliminary data on scaling already indicated to a thoughtful few that Bj was on the right track. It is interesting to speculate about what might have happened if Bj had not been at SLAC and had not taken an interest in deep inelastic experiments. Our large cross-sections would still have been a surprise, but I suspect the data would have attracted little serious attention until separations of W_1 and W_2 had been made. Our cross-sections were smaller than the estimates Gottfried had made with the simple three-quark model (in 1967) so many theorists would have claimed that the quark model was not working. (There are neutral gluons so the quarks carry only part of the proton momentum.)

Feynman might not have taken such an interest in the experiments if he had not been told about Bj's scaling, even with the meager support it had in the earliest data. In the end, I suppose we would still have made almost all the same measurements. With the large cross-sections, the small longitudinal cross-sections and the difference between proton and neutron there would have been enough clues to put things on track. But not for quite a while. I believe the thought that Bj put into predicting (or as he put it, "guessing about") what we would find at SLAC had a big impact on the direction that theory took in response to our data. Here a slide that Bjorken prepared once upon a time:

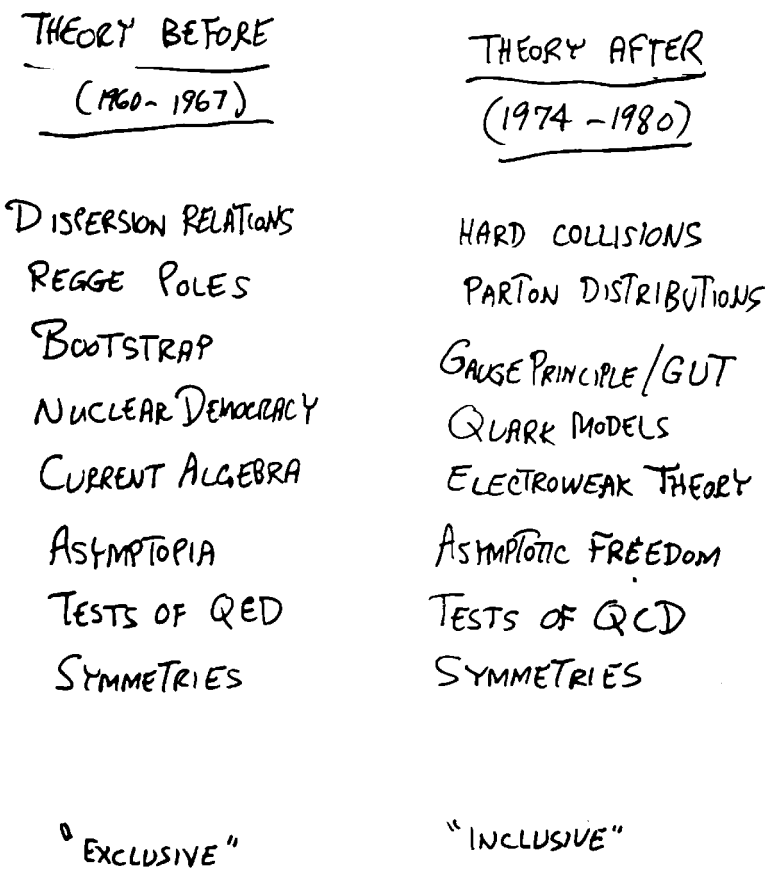

Figure 16 A slide prepared by J.D. Bjorken outlining the shift in emphasis in particle theory as the quarks displaced the hadrons as the basic building blocks of matter, the electromagnetic and weak interactions were combined, and new quarks and leptons were discovered

As an experimentalist, I take great pride in the fact that our 30 year old data is still sometimes included in graphs of the structure functions – albeit way down in the lower left hand corner.

The proton was elementary for about 60 years, but has not been elementary for the last thirty years. How long will all the quarks and leptons stay elementary?

---

* Because of the nature of this talk, I have not included references to the many contributions mentioned in the text. Many of them are very well known, and references to others can be found in the comprehensive Nobel lectures of Friedman, Kendall, and me in Rev. Mod. Phys., Vol. 63, No.3, 573-595. For those interested in talks similar to this one, Friedman's and Bjorken's recollections were published in *The Rise of the Standard Model: Particle Physics in the 1960s and 1970s* (Cambridge University Press, 1997), 566-599.