The Lesson of Quantum Theory, edited by J. de Boer, E. Dal and O. Ulfbeck

Observability of Quarks

James D. Björkén

Fermi National Accelerator Laboratory
Batavia, Illinois, USA

Contents

Abstract

Even if stable hadrons with fractional charge do not exist, most of the criteria of observability used for ordinary elementary particles apply in principle to quarks as well. This is especially true in a simplified world containing only hadrons made of top quarks and gluons. In the real world containing light quarks, essential complications do occur, but most of the conclusions survive.

1. Introduction

It is an honor and privilege to be here to participate in this centenary of Niels Bohr's birth. I am not at all of his generation. I glimpsed him only once in the lunchroom, where a friend pointed him out to me during my first visit to Copenhagen as a fresh postdoc. So his personal influence on me is indirect—mainly through the style and atmosphere of the institute which he created, which to this day so splendidly and directly perpetuates his influence on science and his way of doing science.

The topic I have chosen to discuss—quarks—is not of Niels Bohr's generation. Nevertheless, the issue of how we observe them and how the interior machinery of that observation process works is very much of his generation. And the topic, as it turns out, is even very Scandinavian. The deeper side is studied here in Copenhagen, especially by Holger Bech Nielsen and his colleagues. The more pragmatic side which has its heritage in Bohr's work on propagation of particles through matter, can be found across the Øresund in Lund. So in choosing this topic for a talk here, I risk uttering mere banalities. So be it.

The problem addressed in this talk originated almost as soon as the quark hypothesis was enunciated: if all hadrons are made of fractionally charged constituents, why do we not eventually reach an energy scale of collisions where the constituents are liberated, thereby yielding at least one stable, isolated hadron of net fractional charge? (This goes under the name of the confinement problem.) And given the empirical absence of fractionally charged objects in bulk matter that this is *not* the case, what meaning is there in ascribing reality to these constituents within hadrons? In particular, *how* do the quarks confine themselves even in the most violent of collisions?

Nowadays, the problem is believed to be resolved in the context of the theory of the strong force, quantum chromodynamics, or QCD. This theory did not emerge until a decade after the emergence of the quark, and it was at least another half-decade before it was generally accepted. While to this day QCD is not universally accepted, it is not my purpose here to entertain any doubts about it, but rather to assume that QCD is true. Likewise, I will not try to look at the question from very much of a historical perspective, but go directly to the modern viewpoint, expressed in as simple terms as I can muster.

2. *Quantum chromodynamics without light quarks*

An immediate nonrelativistic answer to the confinement problem is found in the simple harmonic oscillator. If the quarks in hadrons were bound together by harmonic oscillator-like forces, then they never would be "ionized". In order to separate them by a macroscopic amount, one could, with enough energy, accomplish this: they could be placed into a macroscopic orbit. The problem with quarks lies in reconciling this old-fashioned viewpoint with relativistic quantum mechanics. Surprisingly, QCD seems to allow this to happen, at least in a simplified, albeit artificial limit.

Let us start with a review of the essential features of QCD as a theory of the strong force. The most remarkable is the renunciation of the Yukawa picture of meson exchange as the essence of the strong force. Indeed the essence of the QCD strong force is best seen if *all* known mesons—and their quark constituents—are disregarded. This leaves only the unknown—or at least not very well known—tricolored top quarks and the gluon carriers of the QCD force as the remaining degrees of freedom. In this limit, the natural range of the strong force emerges in full clarity as being determined by the QCD confinement scale-parameter Λ. This parameter, with dimensions of mass, is by chance believed to have about the same value as the pion mass, even when pions are removed from the theory.

Thus the confinement distance $\hbar/\Lambda c$ is of order 10^{-13} cm. For distances small compared to this, the QCD force is approximately inverse-square and not too strong; its "fine structure constant" is small compared to unity, and there are many analogies to quantum electrodynamics (QED). But at large distances it is believed—and there are good reasons to believe so—that the force becomes constant and the flux lines becomes concentrated in a tube of roughly the size $\hbar/\Lambda c$.

This top-quark limit has a splendid simplicity, largely devoid of all the complications of relativistic quantum fields—in particular multiparticle production and pair creation. Why is this? The flux tube has dimensions large compared to the Compton wavelength of the top quark, known to be less than 10^{-15} cm. Hence the color field contained in the flux tube is too feeble to pair-create the top quarks; this mechanism is indeed exponentially suppressed. And emission of gluons or quark–antiquark pairs by short-distance mechanisms, while occasionally present, is suppressed because of the smallness of the QCD fine-structure constant governing these processes.

The net result of this is that the effective Hamiltonian controlling the dynamics of a $\mathrm{t\bar{t}}$ meson or ttt baryon is no worse than a relativistic potential-model. And the harmonic-oscillator analogy therefore still holds, the only differences being relativistic kinetic energies for the quarks and between them a potential energy which depends linearly, not quadratically, on their separation.

How now do we observe the quark? An easy way to try is to illuminate, say, a $\mathrm{t\bar{t}}$ meson with a weakly interacting probe, such as a photon or a lepton. The amplitude A_n for finding the system in quantum state $|n\rangle$ is essentially

$$A_n \sim \langle \mathrm{n} | e^{i\boldsymbol{q}\cdot\boldsymbol{x}} | 0 \rangle . \tag{1}$$

If the excitation energy is large (implying large momentum $\boldsymbol{q}$ transferred to the quark), the states $|n\rangle$ can be approximated by their semiclassical WKB formulae. In the region of overlap with the initial state $|0\rangle$, these are essentially plane waves. The important excitation energies will then be established as

$$E_q \sim \left(\sqrt{q^2 + m^2} - m\right) + \text{binding corrections}, \tag{2}$$

in accordance with classical kinematics and the Bohr correspondence principle. Hence, if a coarse-grained energy average is admitted, a wave packet

$$\psi_q(x) = e^{i\boldsymbol{q}\cdot\boldsymbol{x}}\, \psi_0(x) \tag{3}$$

is created by the collision, which propagates classically toward the turning point.

The picture could hardly be simpler. The probability for the collision to take place is given by a perturbative calculation. Coherence between, say, the contributions of interaction of the probe with t and $\mathrm{\bar{t}}$ will for large $\boldsymbol{q}$ clearly be negligible, and the impulse approximation and semi-classical picture of the subsequent motion may be justified.

Need we observe the struck quark? What happens to it? It is easiest to first view the evolution in the center-of-mass reference frame. In that frame the t and $\mathrm{\bar{t}}$ quarks simply oscillate back and forth between their classical turning points. The string tension, or energy per unit length of the QCD string, is about 1 GeV/fermi, i.e. about one proton mass per proton diameter. Were the t and t̄ quarks to be given relativistic momenta by the collision, for example 20 TeV, they would have not inconsiderable oscillation amplitudes. For 20 TeV the maximum separation is of order 0.4 Å, almost atomic dimensions. Eventually the oscillations will be damped. Two mechanisms come to mind but there may be more.

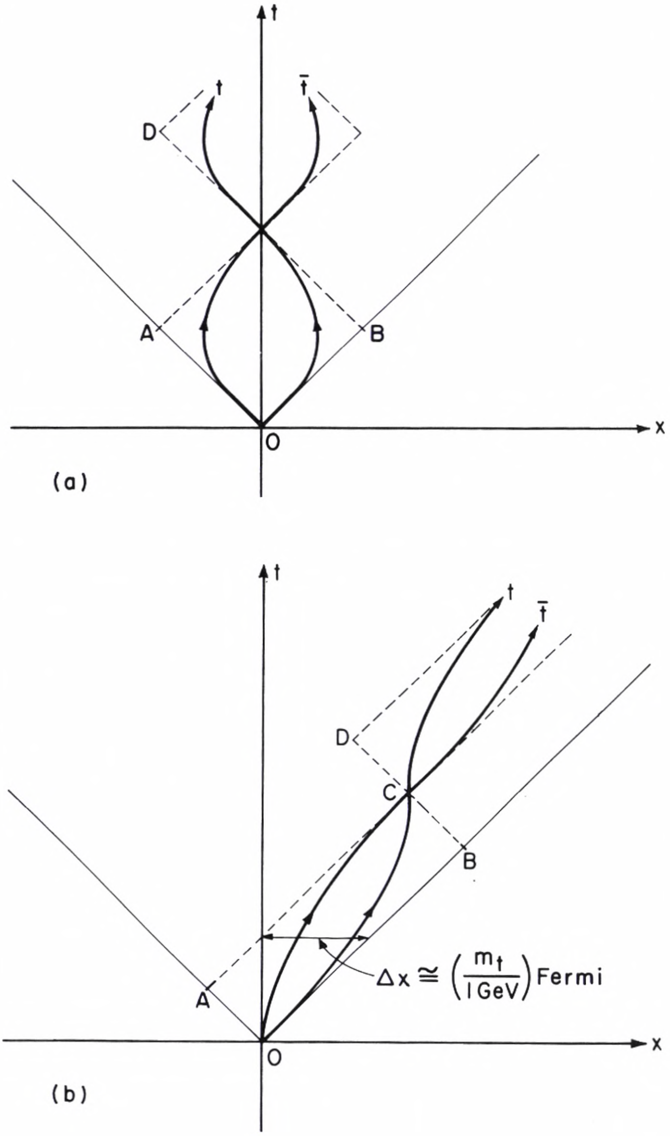

Fig. 1. (a) Space-time picture of t–t̄ motion in the center-of-mass frame; (b) the same in the laboratory frame. (tt̄ initially at rest; momentum q imparted to the t quark only).

The first is the emission of gluonia or glueballs. These are globs of pure gluons—perhaps better characterized as bits of closed flux-tube—which have a mass and size of the order of the confinement scale. (The typical mass estimates are a little larger than the proton mass.) Explicit computation is most easily done in the frame where the glueball in question is emitted at the turning point. But there is poor overlap between the wavefunctions of the initial and final t quarks (the level density is too high). This appears to imply a low probability per oscillation cycle for glueball emission.

A second mechanism is the emission of hard, "perturbative" gluons at birth (analogous to internal bremsstrahlung in QED) and at every half-period when the t and $\bar{t}$ pass by each other. This mechanism appears to be the most important. There may be other damping mechanisms that I have not found. But it is a near certainty that the oscillations will be highly underdamped.

There is an additional subtlety which occurs when our process of Compton scattering from a $t\bar{t}$ meson is viewed not in the center-of-mass frame but in the laboratory frame. The t quark struck by a photon recoils with a momentum (and energy) q, large compared with the rest mass m. This quark indeed moves a long distance, with a constant momentum loss of 1 GeV/fermi, while being decelerated by the string. However, at the other end of the string the antiquark is being accelerated. Soon it is traveling at essentially the speed of light behind the decelerating t quark. And some straightforward relativistic kinematics shows that, *no matter how large the initial momentum* q, the $\bar{t}$ antiquark and t quark never separate by a distance greater than a fixed amount proportional to the t quark mass, essentially,

$$\Delta x \leq \left(\frac{m_{\mathrm{t}}}{1\ \mathrm{GeV}}\right) \times (1\ \mathrm{fermi}). \tag{4}$$

When the leading t quark finally is decelerated to rest, the antiquark passes it up and the roles reverse; this commences the second half of the oscillation period. All this is shown in the space-time diagrams in fig. 1.

Thus it is *not* possible in the laboratory frame to observe the t quark in isolation, no matter how high the energy. But, fortunately, we may take recourse to the more practical colliding-beam processes $e^+e^- \to t\bar{t}$, or gluon + gluon $\to t\bar{t}$ in hadron colliders. These do provide concrete, in-principle, ways of producing macroscopically isolated quarks.

So if one had enough energy and the will, this oscillating top quark could, in a world devoid of light quarks, be observed in just as real terms as any other elementary particle.

3. *Effects of the ordinary light quarks*

The real world contains much more than top quarks, and these create fundamental complications. An essential change is that the stable string of the previous section cannot exist. It becomes unstable, due to pair creation of the light up and down quarks and antiquarks by the strong color fields in the flux tubes. There is now a

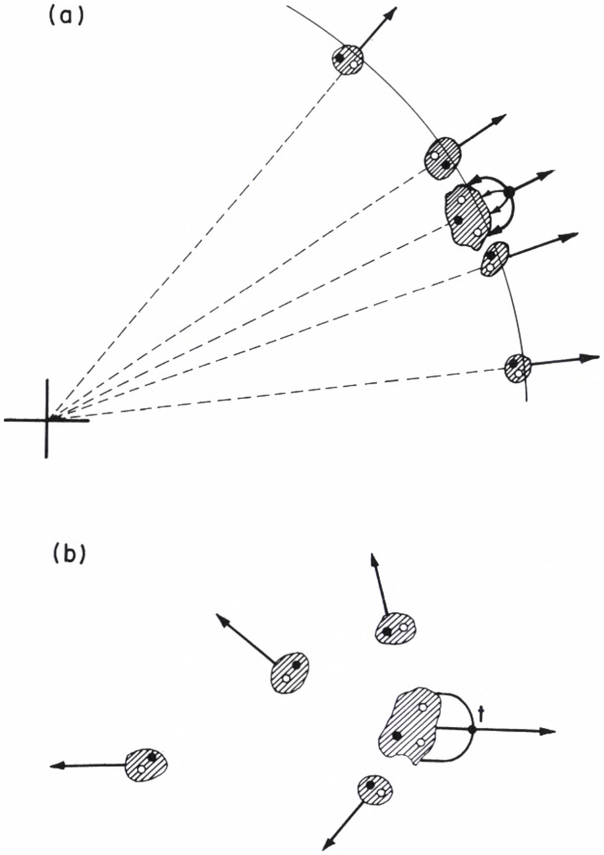

Fig. 2. Sketch of a t quark and its associated "fireball" cloud of quarks, antiquarks, and gluons during its evolution into a hadron jet; (a) laboratory frame, and (b) rest frame of "fireball".

good match between the size of the QCD flux tube and the light-quark Compton wavelength. The energy stored in the string can be used to create these $q\bar{q}$ pairs, and the string breaks into many pieces on a natural time scale $\lesssim 10^{-23}$ s. Since the string pieces contain quarks as well as glue, they are simply ordinary hadrons.

The lifetime of a piece of string can be estimated from experimental observed lifetimes of highly excited $c\bar{c}$ or $b\bar{b}$ meson states. These observations give a width per unit length of order tens of MeV/fermi of string length. This time scale is barely long enough to maintain viability of the concept of flux tube in the presence of the light-quark instability mechanism.

Let us now again look at our $t\bar{t}$ system—again literally with a highly inelastic γ-ray. Now the struck t quark, which we assume to be relativistic, does not grow a long string. Instead the incipient string invariably fragments into mesons. As the t recedes from the spectator, what emerges is a system as shown in fig. 2. The t quark again loses energy at a rate of order 1 GeV/fermi of transit. But this energy is no longer stored in the lengthening string, but is liberated into its decay products. The decay products emitted at late times will be of higher momenta and found adjacent to the excited system containing the top quark. It is tempting to think of that system, with dimensions of the order of one fermi, as a "fireball", emitting mesons as it cools off, and finally becoming a top meson or baryon. But this is a little over-simplified. A large fraction of the momentum imparted to the top quark in the original collision remains with the final hadron containing that quark. But the rest mass of the "fireball" in the early stages of the collision process is much larger than the top quark mass, while in the final stages it is very close to the top-quark mass. Thus the "fireball" mass *decreases* during the evolution of the event much more than its momentum does. Hence the Lorentz $\gamma \simeq p^*/m^* \simeq (1 - v^2/c^2)^{-1/2}$ (and therefore velocity) of the "fireball" *increases*—the "fireball" is accelerated. It is something like a rocket. In the "fireball" rest frame one sees the top quark at the front edge behind which there is a gluon "wake" which creates the quark–antiquark pairs. These quarks and antiquarks materialize into mesons which are emitted out the *back* of the fireball (fig. 2b).

The time-scale for this process is again set by the top-quark energy loss of ~ 1 GeV/fermi, just as for the elastic string. But because the process is dissipative, it terminates at a time of the same order as, but somewhat less than, what was needed to reach the turning point in the simplified "elastic" situation. This is still a large time at high energies. It again scales linearly with energy, as must be the case from basic special relativity: rapidly moving clocks slow down, so the laboratory time to get the job done grows accordingly. Thus the 200 GeV jets found at the CERN $\mathrm{Sp\bar{p}s}$ collider already evolve over a distance scale of up to 10^{-11} cm.

The picture we sketched implies that the quantum numbers of the source of the jet (t quark in our example) are linked to the portion of the jet carrying most of the momentum. Hence the most energetic hadrons of the jet will carry these quantum numbers. This is found to hold experimentally. On average the bottom meson carries more than 70% of the total b-jet momentum, and a charm meson about 50% of the momentum of a c-jet. For light quarks, the net fractional charge of the quark is, on average, found in the leading particles. Evidently this can hold only statistically. But this has been checked both in e^+e^- annihilation and hadron–hadron collisions.

4. *Summary of ways to observe quarks*

This is not the place to recite a long compilation of evidence for the quarks, but is meant only to underline the fact that we "see" them in ways not very different from the way we "see" electrons or protons. The methods include spectroscopy, inelastic scattering, and secondary interactions. We discuss these in turn very briefly.

4.1. Spectroscopy

The pattern of energy levels of a bound system tells us about its structure. This goes back, of course, to Niels Bohr himself. And the long history of spectroscopy which is relevant for the quark structure of hadrons goes back at least twenty-five years. The spectroscopy of baryons provided especially beautiful evidence for quarks even though, to this day, it is not obvious why a nonrelativistic quark model should work so well. More recently, the spectroscopy of $c\bar{c}$ (ψ) and $b\bar{b}$ (Υ) systems, which looks so similar to positronium spectroscopy, is equally decisive in convincing us that these states are built from fractionally charged quarks.

4.2. Inelastic scattering

The presence of electrons in matter can be inferred from the kinematics of the Compton effect. Many similar examples exist, not all of which use photons. Inelastic scattering of electrons from nuclei directly exhibit the presence of individual nucleons and determine their internal motion within the nucleus. For quarks in hadrons, lepton rather than photon scattering has also played the leading role. While lepton scattering from a $t\bar{t}$ system can be viewed as we have in the previous sections, there were grave obstacles in doing so for ordinary hadrons. They are not as reliably describable in terms of potential models. The light-quark pointlike constituents can be expected to move at relativistic velocities, spoiling the impulse approximation. In addition, virtual and real pair creation can be expected to be important as well.

The resolution of these difficulties came in exploiting the relativistic nature of the problem. When the hadron to be probed moves ultra-relativistically, its internal clocks slow down, and the external lepton probe, moving in the opposite direction, sees essentially a static distribution of constituents during the period of collision. The picture therefore reverts to something very similar to the familiar examples from atoms, nuclei, and even excitations in condensed matter. The initial internal motion of the constituent is slow compared to its motion after being struck by the probe. Free-particle kinematics can be used to estimate the collision probability, and the distribution of the scattered probe-particles again measures the initial velocity distribution of the constituent.

4.3. Secondary interactions

The previous methods observe the quark as it is bound within the hadron, just as the analogous methods observe the electron as it is within the atom, or the nucleon within a nucleus. To many, the essence of a real observation of a particle "in isolation" would be to follow and observe its subsequent motion. This implies—especially here in Copenhagen—additional interactions with a medium through which the particle propagates. For example, the ionization loss of a charged particle provides a mechanism by which its path can be followed and indeed defined, in the sense of quantum measurement theory. In the simplified case of section 2, QCD with only top quarks, this could be done straightforwardly given high enough excitation energy. In the general case which includes light quarks, the top quark is invariably immersed in its "fireball" of glue and $q\bar{q}$ pairs. It again would propagate a distance comparable to the "elastic" case and clearly leaves behind a track as well, but one somewhat harder to interpret.

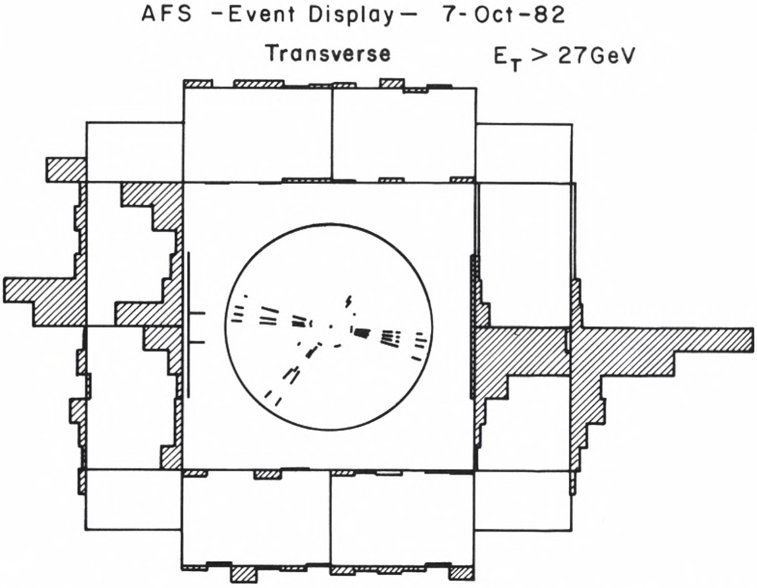

Fig. 3. View along the beam direction of a two-jet final state, most probably a quark–antiquark elastic scattering via gluon exchange, as seen at the CERN Intersecting Storage Rings by the Axial-Field Spectrometer experiment. (The date of the event is singularly appropriate).

Nevertheless, there is a practical way of probing the structure of such a newly formed quark system. It consists of highly inelastic lepton scattering from *nuclei*. If the energy scale in the collision exceeds hundreds of GeV—something attainable especially well in upcoming muon-scattering experiments at Fermilab—the quark system will traverse a considerable amount of nuclear matter before becoming independent hadrons, and its interior structure can thereby be probed. The main effects on the quark motion are anticipated to be multiple scattering and bremsstrahlung through the *strong* force. This gives, for large atomic numbers, a characteristic broadening of the angular distribution of the most energetic hadrons, as well as an attenuation (because of the gluon bremsstrahlung in nuclear matter) in the number of energetic hadrons.

In addition to these means, the quark and gluon jets seen in e^+e^- and $p\bar{p}$ collisions are in some sense the residue of the track of the quark fireball as it propagated through the vacuum. A nice example of this, kindly provided to me by Knud Hansen, is shown in fig. 3.

Carlo Rubbia pointed out to me that the ultimate high-energy physics experiment would be to somehow find the magnetic monopoles and antimonopoles anticipated in grand-unified theories and annihilate them. These monopoles might have a mass of at least 10^{15} proton masses. Were they to annihilate they would liberate quarks of comparable energies. The characteristic distance the quark would travel before full "hadronization" occurred would approach one meter. In this case macroscopic means could be used to follow and (again in the sense of measurement theory) define the course of the quark and its wake of gluons, strings, and $q\bar{q}$ pairs. It may even be that the mean ionization density of the "fireball", which fluctuates in charge as it emits the hadrons comprising its jet, corresponds (when averaged over many events) to a fractionally charged object.

5. Reflections and conclusions

When I look at the preceding arguments, they seem so self-evident that it is hard to recognize a problem at all. Was there ever a problem? The answer, I think, is yes. It existed in acute form before the development of QCD, and was divided into two parts: *why* quarks should be confined within hadrons, and then *how* they did not get out in high-energy collisions. While vague ideas about strings were available, there was little in the way of a relativistic theoretical structure within which such ideas of confinement could be developed. *How* quarks did not get out, and the importance of large distance scales in this process, could be—and was—attacked in the interim, even without appreciation of the color degree of freedom and QCD. The simple example of the top quark bound to elastic strings came with the full comprehension of QCD as the theory of the strong force. In addition the ψ and Υ spectroscopy provided much-needed stimulation from experiment.

What most distinguishes the observability of quarks from the observability of other particles is the technical complication of the light quarks. This makes the traversal of an energetic quark through vacuum a dissipative process, something like (but not identical to) ionization loss of a charged particle in matter. Instead of a mean energy loss of 2 MeV/gm cm^{-2}, we have a value of order 1 GeV/fermi. And the presence of light quarks also implies that a quark will not be found in isolation, but will inevitably be accompanied by a polarization cloud of quarks, antiquarks, and glue, which screens - in *any* frame of reference - its color field and fractional charge at distances beyond the confinement scale of 10^{-13} cm.

To me this additional complication is more technical than truly fundamental, but others may well disagree. The quark has been observed, even in the absence of quark tracks, and there need be little if any mystery associated with that. The real mystery lies in the nature of the medium through which the quark propagates—that is, the nature of the vacuum itself. It has by now taken on much dynamical character of its own, very much like the ground states of the solid-state analogues. The question of the observability of the vacuum itself has become the big problem. I wonder what Niels Bohr would say about that.

Discussion, session chairman H. Bethe

Schrieffer: First a remark. There is a very beautiful example of the near-observability of fractional charge in condensed-matter physics. We have the advantage of being able to pull apart the fractional charge to very large distances in quasi one-dimensional conductors. But this is only in very few conductors. What I did not mention in my talk yesterday was that in most cases we have confinement forces between what is the analogue of solid-state quarks, if you like. These confinement

potentials are in fact linear as soon as you get outside the form factor of the excitations themselves. When you pull them far enough apart they break into what you might call color singlets, which have quite weak forces between them. This can be seen beautifully in photo-excitation experiments. There is a remarkable coincidence between the two fields, even if the origin is probably quite different.
Then a question: Do you have any idea as to how one can account for fractional charge starting out with integer charged fields in a relativistic context?

Björkén: There is an old and beautiful idea by Han and Nambu, which is now obsolete. The model has three triplets of quarks with integer charges. However, it simply does not fit well with present days' standard model.

Rubbia: Concerning the experimental verification of the top quark I would like to bring an update on this: There is a handful of events with two jets and an electron and a neutrino (or a muon and a neutrino) in the collider results. A possible interpretation for these events is a decay of a W particle into a top-quark and a beauty-quark. If this is correct then the top-quark mass is in the vicinity of 40 GeV.