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

Particles, Fields, and now Strings

Steven Weinberg * **

Physics Department, The University of Texas
Austin, Texas, USA

We are celebrating the 100th birthday of a great man, so it is natural here to think back over the history of his and our field. In thinking about any sort of history it is often a useful approximation to see it in terms of the struggle between great opposites, whether Revolution and Reaction, or Cross and Crescent, or even Guelf and Ghibelline. (Somehow historians always manage to find opposites that start with the same letter.) In our field a fair amount of the history, if not of the last hundred at least of the last fifty years, can be seen as the struggle between two grand views of the world: as a world made of particles, or as a world made of fields. This opposition crystallized in the 1950s into a more specific distinction, between theoretical physics as quantum field theory, and as S-matrix theory. Just in the last year or so this old oscillation between quantum field theory and S-matrix theory has entered a new cycle, in the recent revival of what are called string theories. I will come back to string theories later in this talk, but it will be useful first to try to set the historical stage.

As a theory of everything, I suppose quantum field theory starts with the pair of papers written by Heisenberg and Pauli in 1929. Before that there was only a quantum field theory of the electromagnetic field, limited to photons. The Heisenberg–Pauli formalism was further developed in 1934 in papers by Pauli and Weisskopf and Furry and Oppenheimer. Quantum field theory presented a grand view: The fundamental ingredients of our universe are fields, and the particles are mere quanta of energy and momentum of the fields. But despite the grandness of this vision, quantum field theory did not really catch on the 1930s. I wasn't doing physics then, but my impression is that there was great doubt about the future of quantum field theory for two particular reasons. One is that a more dualistic approach, especially associated with Dirac, survived and for many purposes was as good as quantum field theory. In Dirac's view (which I think he kept to the end) there are fields, but only the electromagnetic fields, and there are particles, electrons of both positive and negative energy, which are eternal but which can be raised from negative energy to positive energy states, giving the appearance of pair creation.

* Author's note: This talk is largely based on one I gave at the close of the American Physical Society Division of Particles and Fields meeting in Eugene, Oregon, on August 15, 1985. I was asked at Eugene to assess future prospects for elementary particle physics in the light of its past history, and I thought that nothing else would be so appropriate for me to talk about at the Bohr Centenary.

** Discussion on p. 238.

This Dirac hole theory was for most purposes as good as quantum field theory. A notable exception of great importance in the history of physics was Fermi's theory of beta decay, in which for the first time a physical process was described in a way that could not have been understood within the old hole theory. Pais emphasized the importance of Fermi's theory of beta decay in the development of quantum field theory in a talk I heard him recently give in Princeton. It is surprising that the success of Fermi's theory did not convert more of the theorists of the 1930s into enthusiasts for quantum field theory.

Another reason was that although quantum field theory worked well in electrodynamics and was beginning to give a reasonable description of the weak interactions, it had some obvious failures. Most important of all were the ultraviolet divergences, discovered when quantum field theory was applied to the electromagnetic self-energy of the electron by Oppenheimer and independently by Waller in 1930. These infinities have been on our minds ever since. Also there were problems in the application of quantum electrodynamics to electromagnetic cosmic ray showers, which were actually somewhat of a red herring, because these problems are now believed to have been due to the production of mesons. There were also some wrong calculations done, which never helps. Many of these problems were of course equally problems for hole theory and quantum field theory, and it didn't seem very important to distinguish the two, as both were failing.

The first revival of quantum field theory took place in the arena of quantum electrodynamics in the late 1940s through the work of Feynman, Schwinger, Tomonaga, Dyson, Salam and others. The success was brilliant. In a few short years it was learned how to develop a Lorentz invariant perturbation scheme for quantum electrodynamics, how to carry out calculations in such a way that the infinities were not only swept under the rug, but dealt with in a physically satisfying way by renormalization of physical constants, and how these calculations not only could be done, and gave physically sensible answers, but gave answers that agreed with experiment to large numbers of decimal places.

It was, I suppose, as exciting a time as theoretical physicists have had since the 1920s. I heard about all this when I was an undergraduate a few years later, and it seemed to me obvious that what I wanted to work on was this brilliant, successful new field, quantum field theory, which had been so beautifully vindicated a few years earlier. I came to Copenhagen as a first-year graduate student in 1954, knowing nothing whatever about quantum field theory, but all keen to learn the technology of Feynman diagrams and renormalization. I was very disappointed and surprised to learn that already a disillusionment had set in, and that the theorists at Copenhagen, some of the leading spirits of the age, were already beginning to have renewed reservations about the future significance of quantum field theory in physics. These reservations we all remember. The first was that the renormalization theory which had worked so wonderfully when applied to quantum electrodynamics didn't work at all when applied to weak interactions in the only theory that was then extant, the Fermi theory of beta decay. Also, although there weren't necessarily any problems with renormalizability, the strong interactions were too strong to allow any use of perturbation theory.

There was another problem which is perhaps less direct, but which I think was also important. Quantum field theory has built into it the sense of an elementary particle as being different from something like an atom or a piece of chalk. An elementary particle in quantum field theory is a particle whose field appears in the fundamental Lagrangean or field equations. This seemed reasonable as long as the elementary particles were just a few in number, the electron, the photon, the proton, and then a little bit later the neutron and the neutrino. But beginning in the 1950s and then with a rush in the 1960s, the old distinction between elementary and composite particles began to look less and less physical. Particles were discovered like the ρ meson; in what sense could you say that it was not elementary and the proton was? It really didn't seem to make any sense to choose a few particles which just happen to be the ones that were historically discovered first, the proton, the neutron, and so on, and make a fundamental field theory out of those particular particles.

In the mid-1950s in Copenhagen and then in Princeton I began to hear about an alternative to quantum field theory. Already in 1937, John Wheeler had introduced the S-matrix, and then independently Heisenberg in 1943, and Christian Møller here in Copenhagen in 1945 had developed the beginnings of what might be called S-matrix theory, a theory in which the fundamental object of interest is not the evolution of a state vector in time, but rather the probability amplitude that given an initial state at $t = -\infty$, a particular final state will be observed at $t = +\infty$.

In this early work, up to the end of the 1940s, there was no real idea what sort of dynamical scheme might be used to calculate the S-matrix without referring back to a quantum field theory. But then, picking up the earlier work on electromagnetism of Kramers, Kronig, Toll and Wheeler, in the mid-1950s Chew, Goldberger, Gell-Mann, Low, Nambu, Thirring and others began to apply these ideas to the strong interactions, and there developed the beginnings of a self-contained S-matrix theory of elementary particle dynamics. The starting idea was that the unitarity principle would relate the imaginary part of any scattering amplitude to the total cross-section, which involves squares of scattering amplitudes, and the causality principle would relate the real and the imaginary part of the scattering amplitude to each other, and so in this way one would have a closed self-consistent set of non-linear equations, which one might either solve perturbatively, or perhaps in some nonperturbative way. This approach became quite popular, in part I think because it satisfied a deep philosophical preconception that many physicists have shared since earliest times.

Physics of course deals with observables, and in the end we always have to relate our theories to doable experiments. But there is the more stringent demand, which from time to time emerges in the history of physics, that not only must the end product of our calculations relate to observables, but that every ingredient in our theories must be directly observable. This requirement was very important in Bohr's thinking. We remember the work that Bohr and Rosenfeld did on the measurability of the electric field, a homework problem that had been given them by Landau and Peierls. Bohr regarded it as a matter of great importance to establish that the electromagnetic field was in fact as measurable as the commutation relations allowed it to be. In S-matrix theory one could feel satisfied that one was again dealing only with observables, whereas that was not the case in quantum field theory.

This last remark may seem a little strange, after what I've just said about Bohr and Rosenfeld showing that the electromagnetic field was observable. Remember, however, they were not working in what we now call quantum field theory, but rather in the old hole theory, in which the only fields were the electromagnetic fields. In a quantum field theory, the electron field is hardly observable, because it is massive and charged and a Fermi field. Quantum field theory deals with fields that we are not going to measure; they only appear in the theory as parts of the machinery for calculating the S-matrix. And so, by turning away from quantum field theory toward S-matrix theory, the physicists of the 1950s could enjoy an act of purification, an expelling of the unobservables.

It soon became clear that in order to develop S-matrix theory in this way, one had to go far beyond the forward-scattering dispersion relations that were first studied. After all, unitarity relates every physical process to every other physical process. A full fledged S-matrix theory dealing with processes of arbitrary complexity began to be developed at the end of the 1950s, especially at Berkeley in a group including Mandelstam and Stapp, centered around Geoffrey Chew.

The group in Berkeley, and many others, felt that S-matrix theory was the most interesting thing going on in physics, but I found it forbidding for various reasons, and did not get into it myself. One thing I did do that was motivated by S-matrix theory I will mention here because it comes up later on in this lecture. I tried to see how the well-known properties of electromagnetism and gravitation could be understood in a purely S-matrix context without any ideas about quantum field theory. I showed that any mass zero, spin one particle would, according to the axioms of S-matrix theory, have to behave pretty much the way we know photons do, and any mass zero, spin two particles would have to behave the way we believe gravitons do, as was also remarked by Feynman. I will come back to this point later.

The S-matrix theory which began in the mid-1950s had by the mid-1960s already collapsed, at least as a grand program for solving the problems of physics. This was for a number of reasons. (There are always a number of reasons for everything.) One reason is that in order to understand how to deal with physical processes involving anything more complicated than elastic scattering, it was necessary to understand the analytic behavior of functions of many complex variables, to understand functions that have simultaneous cuts in more than one variable at a time. In fact, this is necessary not just to do calculations, but even to state the axioms of S-matrix theory. This kind of mathematics of many complex variables is notoriously difficult; it may or may not be beyond the capacity of the human brain to understand the analytic structure of the full multi-particle S-matrix, but I was able to show quite rigorously that it was beyond me.

A second problem with S-matrix theory was that there never was any rational approximation scheme; there was no small parameter in the theory. And a third problem was that despite the wonderful, logical-positivistic flavour of S-matrix theory, it really wasn't very successful in predicting anything much about the real world. The forward-scattering dispersion relations and Regge pole systematics, of course, were and continue to be very important, but beyond this very little about the strong interactions was successfully predicted by the S-matrix theories. In particular, although they concentrated very much on pion scattering, they missed the important point (revealed through the application of current algebra in the mid-1960s) that pion interactions at low energy are quite weak.

However, before the end of S-matrix theory, there was a last curious spurt of activity. Since it was so difficult even to formulate exactly what the axioms of S-matrix theory were, let alone apply them, a number of theorists began just to guess at candidate scattering-amplitudes which would satisfy a number of our general ideas about the properties of physical scattering amplitudes, ideas having to do with crossing symmetry, Regge asymptotic behavior, and so on. Veneziano, in particular, discovered a remarkable formula for a candidate meson–meson scattering amplitude. This amplitude can be expressed *either* as a Breit–Wigner sum of an infinite number of resonances, *or* as the sum of an infinite number of exchanged resonances (which give it the desired Regge asymptotic behavior at high-energy and fixed-momentum transfer), a property known as "duality".

Great effort went into the exploration of these dual models, starting with the paper of Veneziano. They were very rapidly generalized to multiparticle processes, by Bardacki and Ruegg, Goebel and Sakita, Chan and Tsun, and Koba and Nielsen in 1969, and it was then pointed out by Nambu, Nielsen, and Susskind that the problem to which the dual scattering amplitudes were the solution was the problem of the motion of a relativistic string. That is, the infinite number of particles whose exchange produces the Regge trajectory are in one to one correspondence with the modes of vibration of a relativistic string. For the Veneziano model in its original form the string is open, with free ends which travel at the speed of light. The strings have a certain tension, and since this was all in the context of strong interaction physics, this string tension was imagined to be about a pion mass per fermi, or roughtly (100 MeV)^2. There were other models developed; another model called the Virasoro–Shapiro model was found to be based on a closed string, with a string tension that again was taken to be about (100 MeV)^2. Spinors were incorporated into string theories by Neveu and Schwarz, Ramond, and others, and it was this formalism that led Wess and Zumino in 1974 to their development of supersymmetry.

It is really remarkable, now looking back at the period of the late 1960s and early 1970s, how much work went into these string theories without the slightest encouragement from experiment. In fact, the string theories incorporated features that were not only not confirmed by experiment, but were in gross contradiction with experiment. One of these features was that these theories contained massless particles, and given the background of these theories, these were taken to be massless strongly interacting particles, which clearly would have already been observed. In the open string theory, there was a massless spin one particle, and in the closed string theory massless particles of spin two and also spin zero. Another problem was that when one tried to take into account the unitarity corrections to the S-matrix elements, it was found that these theories really only made sense in 26 dimensions, or if you included fermions in the theory, then in 10 dimensions. That was quite embarrasing. Schwarz and Scherk in 1974 suggested that these theories should not be thought of as describing the strong interactions; rather, the mass-zero spin-two particle that appeared in the closed-string theory should be identified as the graviton. In other words, the string tension would not be of order (100 MeV)^2, but would be something like the square of the Planck mass, $(10^{38}\ \mathrm{GeV})^2$. One does not assume general covariance or the equivalence principle here, but these theories are Lorentz covariant, and thus as I mentioned earlier, any mass-zero spin-two particle in them would have to interact like the graviton of general relativity. However, this proposal received little attention, I think above all because it was embarrasing to be working on theories that made sense only in 10 or 26 dimensions.

However, perhaps even more influential than any failure of the dual models or string theory was a second revival of quantum field theory, which began at just about the time of the development of string theory. I need not dwell on this, as I suppose it is familiar to everyone here. First, it was shown by 't Hooft and others that the earlier suggestion of a spontaneously broken renormalizable gauge theory of electroweak interactions would indeed work mathematically. Very soon thereafter, the discovery of neutral currents and the measurement of their properties showed that not only could an electroweak theory be constructed along these lines, but the already constructed $SU(2) \times U(1)$ theory was in fact the correct theory of electroweak interactions. At the same time, there was also the development at last of a theory of strong interactions, based on much the same Yang–Mills mathematical structures, in which instead of the gauge symmetry being spontaneously broken, it provides a trapping mechanism which hides the underlying constituents, the gluons and quarks, from our view. We had in the early 1970s a fully developed standard model of weak, electromagnetic, and strong interactions, which apparently was capable of describing everything to do with accessible particle physics. Exciting experiments then later in the 1970s gave a final confirmation to these theories, showing that in fact the neutral currents did behave in the way they were supposed to, the quarks and leptons really are what we think they are (except that there are more of them), and the W and the Z are really there.

Once again it became possible to think of a quantum field theory as being a fundamental theory of nature. In particular it made sense once again to talk about particles as being elementary, because we no longer had to think of the proton or the neutron or the ρ meson as elementary particles. The elementary particles had been reduced to a fairly manageable set; leptons, quarks, gluons, W, Z, the photon, and maybe a few Higgs bosons. One thing that especially excited me, was the fact that at last in these theories we could understand in a natural way why symmetries like parity and strangeness and isotopic spin were symmetries of some interactions and not of others. All the mysterious facts that my generation of physicists had had to learn as empirical rules when we were graduate students suddenly began to make sense rationally.

I tried for this talk to find a couple of quotes in the literature to exemplify the mood of enthusiasm in the mid-1970s, and then the sense of disillusionment that set in a few years after. After a good deal of searching I found two perfect quotes, both by me. In 1977 in an article in Daedalus I had the effrontery to say,

> "If something like a set of ultimate laws of nature were to be discovered in the next few years [and thank goodness I said in parentheses 'an eventuality by no means expected'], these laws would probably have to be expressed in the language of quantum field theory."

Just four years later in 1981 I was at another anniversary, the 50th Anniversary of the Lawrence Berkeley Laboratory, and I gave a talk that also dealt with the oscillation between S-matrix and field theory. At the end of the talk I concluded with the words,

> "These have been exciting times. Quantum field theory is riding very high, and one might be forgiven for a certain amount of complacency with it. But perhaps we will now see another swing away from quantum field theory. Perhaps that swing will be back in the direction of something like S-matrix theory."

Well, I was at Berkeley, and I wanted to say nice things about S-matrix theory, but I had two serious reasons for this sense of caution about the future of quantum field theory.

One reason is that theorists had failed to make further progress in explaining or predicting the properties of elementary particles, beyond the progress that was already well in hand by the early 1970s. The standard model has many loose ends; mass ratios, coupling constants, a whole menu of quarks and leptons, and we have simply not succeeded in explaining it. Several attractive ideas were tried: grand unification, technicolor, preons, supersymmetry, Kaluza–Klein theories, and all that, but despite so much clever mathematical work, almost nothing has come out in the way of concrete numbers that could be compared with experiments. Perhaps the only success in that hard quantitative sense was the grand-unification prediction of $\sin^2 \theta$, a prediction that does seem to be quite robust, and also to agree with experiment. It soon became clear that in trying to make the next step beyond the standard model we would probably have to understand physics at or near the Planck scale, partly because that is where whatever grand unified group there might be would break down, and also because after all gravity exists, and gravity becomes a strong interaction at the Planck scale. Unfortunately, throughout the 1970s most of us saw no hope for a quantum field theory of gravity.

The second reason that I gave in 1981 for being skeptical about the future of quantum field theory is that we could understand its successes in the low-energy range, up to a TeV or so, without having to believe in quantum field theory as a fundamental theory. There is a folk theorem (a term of Wightman, meaning something which is generally known to be true although it has not been proved) that says that any theory which satisfies the axioms of S-matrix theory, and contains only a finite number of particles with mass below some M, will at energies below M look like a quantum field theory involving just these particles. That is, quantum field theory by itself has no content; it is just a way of calculating the most general scattering amplitudes that obey the axioms of S-matrix theory. Of course, one might argue that the quantum field theories that we have developed, like quantum chromodynamics, are not just any old quantum field theories, but simple, even beautiful, field theories. However, another folk theorem tells us that in the effective field theory, to which essentially any theory reduces at sufficiently low energy, the non-renormalizable interactions are all suppressed by powers of the underlying fundamental mass scale, which one might imagine is the Planck mass scale. Thus, the physics we see at accessible energies should be described by a renormalizable effective field theory, and we know that the interactions in such theories are always limited in number and complexity. (This is a view of the rationale for renormalizable field theory that is somewhat different from the one described at this meeting by Arthur Wightman.) To see anything else we would have to do experiments at the Planck scale, except that a few interactions though very weak may be detectable for special reasons, like for example the special circumstance that gravity adds up coherently, so that although the gravitational interaction is fantastically weak, we can still measure its effect for macroscopic bodies.

To summarize: quantum field theory has not done much for us lately, and what it had done for us earlier we can understand without having to believe that quantum field theory is in any sense fundamental.

When I made these remarks in 1981, I had no idea what direction S-matrix theory would take as a possible replacement for quantum field theory. In fact, Schwarz had been forcefully advocating string theory as the only hope for a quantum theory of gravity ever since his 1974 work with Scherk, but he was ignored by almost everyone (myself included). In 1980 Green and Schwarz proved the space–time supersymmetry of superstring theory. In the following two years they developed a new supersymmetric formalism for superstrings and used it to invent the Type II superstring theories and to prove their finiteness at one loop. At the 1982 Solvay Conference in Austin, Zumino mentioned this work of Green and Schwarz as the natural candidate for a finite theory of gravitation. Then in 1983, at the Fourth Workshop on Grand Unification, Witten gave an influential talk about $d = 10$ superstring theory, and other theorists began to take such theories seriously as a promising approach to quantum gravity. The time was ripe for this suggestion, as it had not been in 1974, partly because we were all so frustrated within everything else, and also because several years of work on Kaluza–Klein theories had made us comfortable with the idea that space–time might really have more than 4 dimensions, with all but 4 wrapped up in a compact manifold of very small circumference.

Witten in his talk also pointed out the theoretical obstacle, having to do with a hexagon anomaly, that impeded this development. Later, with Alvarez-Gaumé, he discovered the cancellation of this anomaly in one example of a superstring theory, that unfortunately seemed phenomenologically unpromising. Then the great breakthrough came last year when Green and Schwarz demonstrated that there were a few potentially realistic superstring theories, with very specific gauge groups, in which the anomalies that Witten had worried about cancelled. This started an explosion of interest in string theory, which has not yet even peaked.

I suppose that this should be scored as a victory for the S-matrix approach. String theory grew out of S-matrix theory, but in a sense it has some of the features of both S-matrix theory and quantum field theory — the experts have not yet settled down in their view of what string theory really is. Indeed, this is one of the things that makes the theory hard to learn; not everyone will tell you the same thing about what it is you are supposed to be learning.

On one hand there is a view of string theory which takes seriously that these are theories of strings. Instead of quantum fields that are time-dependent functions of the position in space of a particular particle, you have quantum fields that are time-dependent functionals of the configuration in space of a moving string. This second quantized quantum field theory of strings unfortunately does not yet exist, but many of the leading experts in this area are working very hard to develop it. The particles are the normal modes of these strings, so when you calculate an S-matrix element (which in the end is what you always have to do) you imagine a string in a particular normal mode colliding with another string in another normal mode, and perhaps two strings joining together to make a single string, and then that single string breaking apart to be two other strings, which finally wind up in two other normal modes. Calculations are not actually done that way. The description I just gave is often presented in public talks about string theories by the experts, but as far as I can tell they do not actually do calculations that way. It is too hard, and the formalism has not been fully developed yet.

There is another approach to string theory, which is the one that almost everyone actually uses. In this other approach, one starts with the observation that a string moving through space sweeps out a two-dimensional surface in space–time. You can describe the string by giving the space–time coordinates x^μ as functions of two parameters. One parameter, σ, tells you where along the string you are, and the other parameter, τ, tells you how long the string had to move. So the string theory can be regarded as a quantum field theory in *two* dimensions, the "fields" being taken as the d quantities $x^\mu(\sigma, \tau)$, (where $d = 4$ or 10 or 26 or whatever) with perhaps some spinors $\psi(\sigma, \tau)$ as well. The interpretation of this two-dimensional field theory in terms of physical processes in d space–time dimensions has to be accomplished by asking what sort of quantum averages in the two-dimensional world have the unitarity and Lorentz transformation properties that we require for the S-matrix in d-dimensional space–time, so we have a curious blend here of quantum field theory and S-matrix theory.

It is natural to write the Lagrangean for the two-dimensional theory so that it is independent of the choice of the parameters σ and τ, which of course requires the introduction of a two-by-two metric tensor. As emphasized by Polyakov, it turns out then that the old string theory is not only a generally covariant two-dimensional field theory, but is invariant as well under conformal transformations, in which the metric is multiplied with an arbitrary function of the two-dimensional coordinates. This may sound like I am getting into technicalities here, but the addition of conformal invariance to two-dimensional general covariance and d-dimensional Lorentz invariance has an overwhelmingly important consequence: the string Lagrangean must be that of a *free* field theory in two dimensions, with the non-triviality of the S-matrix in flat d-dimensional space–time arising not from interaction terms in the Lagrangean but from the non-trivial topology of the Riemann surface described by the 2×2 metrics. So here we have the realization of an old dream of what a fundamental theory ought to be: interactions are not something we insert in a more-or-less arbitrary way into a Lagrangean (and might if we wished be left out altogether), but are inevitable consequences of the nature of the theory's degrees of freedom.

The differences between these two views of string theory can be illustrated by considering how each would deal with a "one-loop" calculation of the S-matrix for two-particle scattering in a closed string theory. In the two-dimensional field theory approach, one imagines the two-dimensional space to form a torus, and carries out a free-field quantum average of a product of four "vertex functions" of position on the torus, one for each incoming or outgoing particle. On the other hand, in the second-quantized string theory approach, one imagines two closed strings in different normal modes approaching, joining to form one string, then breaking up again into two closed strings, then joining again to form one closed string, and finally breaking apart again to form two closed strings in definite normal modes. This gives one more of a sense of the physical reality of strings, but it is not a very elegant way to describe a torus.

I have not been entirely impartial here in drawing the contrast between the two leading approaches to string theory. My preference for the two-dimensional field theory approach may be due in part to the fact that this is the only approach that I have so far been able to learn. Certainly one should try to understand all possible approaches. Also, it may be, as often argued, that the second-quantized field theory of strings offers the best hope for an understanding of non-perturbative effects. Nevertheless, since string theory is supposed to be better than quantum field theory, it does not seem clear to me that the best strategy is to make string theory look as much as possible like field theory, only with strings instead of particles, which I take is the spirit of the second-quantized approach.

In the last few minutes, I want to take up the question that has doubtless been on the minds of all those of you who have not yet become string mavens (meyvn, mevinim, Jiddish for expert; eds.). The question is: Is it safe to ignore string theory, and hope that it will go away? For a theorist, the question takes the form whether it is necessary to learn all about automorphic functions, Riemann surfaces, Virasoro algebras, and all that, or just bypass all this effort and wait for the next fashion in theoretical physics. For the experimentalist, the question is whether it is worthwhile beginning to think of possibly testing these theories?

In trying to answer these questions, I must say right away that there is not the slightest shred of experimental evidence for string theory. The same was also true of the other theories that we developed in our desperate attempt to go beyond the standard model, in particular for supersymmetry and Kaluza–Klein theories, which have now been incorporated into superstring theory. Never has so much brilliant mathematics been done by physicists with so little encouragement from experiment. Furthermore, just as for supersymmetry and Kaluza–Klein theories, the string theories have still not settled down so that they could make very definite predictions which could be tested experimentally. However, in this respect I think string theories are really different from the Kaluza–Klein and supersymmetry theories. Supersymmetry is a symmetry like Lorentz invariance; it allows a tremendous variety of possible dynamical theories. Likewise, Kaluza–Klein theory is just a general idea, that there might be some higher number of dimensions which are compactified; again, this idea allows a great many specific theories. String theories, on the other hand, are very rigid. There are almost no string theories at all, and of the few possibilities only one at present (the "heterotic" superstring of Gross, Harvey, Martinec and Rohm) seems at all promising phenomenologically. As I discussed earlier, what you are really doing in string theory is studying two-dimensional conformal gravity (actually supergravity), which is a free-field theory, so that the only interactions are those that arise from the topology of the two-dimensional manifolds. Also, the topology of a two-dimensional manifold is completely specified by the number of handles that you put on it (assuming it an orientable closed surface). And so there is nothing you can tinker with in these theories; they are either right or wrong as they stand. We do not have any experimental evidence for string theories, and I cannot really tell the experimentalists what they should look for, but the string theorists in the next few years should be able to come up with definite predictions, which can then be tested. Already, there is an indication that any realistic string theory when compactified down to four dimensions will probably contain at least an extra U(1), so that it is worth looking for one more gauge symmetry in addition to the $SU(3) \times SU(2) \times U(1)$ of the standard model. But the precise features of this extra U(1) are certainly not yet predicted. The biggest gap that will have to be crossed before such predictions can be made is in understanding the dynamics of the compactification from 10 to 4 space–time dimensions. Candelas, Horowitz, Strominger, Witten and others have made some progress in understanding the general features of this compactification, but much remains to be done.

I have remarked that there is no evidence for string theories, but there are other criticisms of a more fundamental nature. Georgi has argued to me that it really would be very unlikely that string theory should provide a fundamental theory of gravity and everything else, in part because after all string theories developed out of the original guess by Veneziano of a scattering amplitude for strong interaction meson–meson scattering. Why in the world should we believe that mathematical structures that grew out of the attempt to understand the strong interactions should be applicable not to the strong interactions but to everything, including gravity? I do not agree with this argument, because I do not agree with its historical basis. The string theories and the dual models which preceded them did not grow out of an attempt to understand the detailed empirical facts of the strong interactions; in fact, they never were much good at that. They grew out of an attempt to find some solution to the problem of constructing scattering amplitudes that satisfy the basic axioms of analyticity, unitarity, and so on. And, these are of course the same problems that we are all solving, whether at 100 MeV or at the Planck scale. There is a good chance that the way of satisfying these fundamental S-matrix principles that is provided by string theory is unique, at least if we want to include gravitation. The string theorists of the late 1960s and early 1970s guessed that the kind of dual model which they were proposing would turn out (when suitably unitarized) to be the more or less unique solution of the axioms of S-matrix theory. And maybe it is, not as applied only to the strong attractions, but as applied to everything, including gravitation. Maybe we really do now know the laws of nature, and the only thing that is left is to work hard for the next few years trying to figure out how the ten dimensions get compactified to four, and then finding out what low-mass quarks and leptons and gauge bosons are left over, calculate their mass ratios, check to see whether they agree with experiment, and then go home.

Relying on a general sense that nothing ever in this life works out the way we want it to ("we" here means theorists), my guess is that there are many surprises that will be provided to us both by imaginative theorists and by enterprising experimentalists before we finally get to the solution to our problem. Nevertheless, I would argue that it is not safe to ignore string theories and wait for the next change in fashion. These theories are much too promising and too beautiful for us not to take them very seriously and explore their consequences for as long as it takes.

Discussion, session chairman A. Salam

Nielsen: There is at least one aspect of the program in which uniqueness seems doubtful. That is the compactification of the six extra dimensions. Of course it may be determined dynamically.

Weinberg: It is true that you have to be cautious concerning uniqueness. I should have mentioned that when you compactify the free two-dimensional theory (whose fields $x^{\mu}(\sigma, \tau)$ lie on a ten-dimensional manifold) to something which is four-dimensional space–time and something else, you find that the two-dimensional string theory is not free and perhaps strongly interacting.

One can say that the fundamental actions are either stated at the ten-dimensional level or at the four-dimensional level. I lean towards the first point of view, and hope that compactification is a dynamical afterthought. Perhaps this happens in a unique way, but perhaps there are still a few parameters which are undetermined (which would be disappointing since then quantities like α could not be calculated without knowing the initial conditions of the universe).

Johnson: Early in your talk you pointed out that renormalizable theories are theories in which the details of the cut-off are irrelevant. We have a similar situation in condensed matter physics, where the type of cut-off does not affect critical exponents. Why then do you need a fundamental theory if it is irrelevant? How can you tell what the theory is on the basis of physics at our own scale in which we only know what the order parameters are?

Weinberg: The fundamental theory is not really irrelevant in that it tells us where the parameters in the effective theory come from. By way of analogy with critical phenomena, there is certainly universality for critical exponents, but the Curie point of a metal must be determined from the microscopic theory. Here we want to see not only the 17 or so free parameters of the standard model but also the 3, 2 and 1 in $SU(3) \times SU(2) \times U(1)$ and the dimensionality of space–time coming out of the final theory.

I think that in the next few years the theory will prove to be either stunningly successful in explaining things we already know or else turn out to be quite wrong.

Bleuler: Will a general unification explain the enormous gap between gravity and other interactions and also why there are four types of interaction?

Weinberg: That is the right question to ask. We have learned, however, to ask it another way: why are all the mass scales, such as particle masses, so small compared to the Planck mass? There is a natural way of getting such hierarchies with supersymmetry which can protect the masses of certain particles from becoming of the order of the Planck mass.

Of course, the superstring formalism is so restrictive it may fail — but that's what is good about it.

Bleuler: Would there be a change in the definition of four-dimensional space?

Weinberg: Yes, a radical change, as the x^{μ}'s are now dynamical variables in a two-dimensional theory.

Fowler: When John Schwarz talks about superstrings, he describes photinos, etc. Why is there not some hope that these can be discovered experimentally?

Weinberg: Actually, there is hope. The difficulty in finding superpartners is that they are not protected by gauge symmetry from getting large masses (whereas the particles we see are protected). Some theories say that photinos, quarks, etc. are right around the corner, whereas others say that they're at the Planck scale. In the former case there is a good chance that photinos make up the dark matter in the universe.

Lee: You made a remark about this theory being either proved right or dying, in the next few years. I don't think it will do that. Very rarely in our experience have either people or nature not been ingenious enough to keep theories workable for long periods of time. Perhaps we will neither prove it right nor wrong and it will reappear in a few decades.

Weinberg: Yes, and we may have a proliferation of string theories, but then I say to hell with it.

Salam: I do not think we should take this attitude. I am recalling a remark by Niels Bohr in 1948, when I first heard him. He was speaking of the fundamental length. He wanted to have a fundamental length at 1 GeV. Now we do have a fundamental length with a local field theory, which is an amazing statement to be made. We have a theory of which my student Chris Isham said that he joined the quantum-gravity bandwagon in order to get quantum physics understood in terms of general covariance. That is to say, Einstein would be supreme and Planck would come out of it. What has happened is exactly the opposite: you get the theory of Einstein as a very special case of the string theories and Planck's constant has to be put in by hand. These are very important remarks which have come out of these string theories. I do not agree with T.D. Lee. But even if the theories do not satisfy the uniqueness principle advocated by its adherents, it will have a very strong place.

I would like to end by a remark by Witten when he said that 57 years of point quantum field theory have come to an end. We shall now in the future have string field theories of which point quantum field theories will be special cases.