Where We Are Now

Steven Weinberg

I knew Robert Oppenheimer for only a few years when he was director of the Institute for Advanced Studies, but I knew of him as one of the great men of our times for many years before. It is a high honor to receive an award bearing his name, especially an award that has previously been received by Dirac, Dyson, Salam, and Serber. I am very grateful.

In considering what to talk about here, I picked up Dirac's little book, *The Development of Quantum Theory*, which contains the text of the talk he gave on receiving the first Oppenheimer prize. It was fascinating for me to read Dirac's reminiscences of the heroic period at the birth of quantum mechanics, and I wished I had heard his talk. I don't think reminiscences of my own times would be anywhere near so interesting—as Dirac says, there has not been any revolution in physics since the early 1930's even remotely as important as the development of quantum mechanics, a fact of which physicists of my generation are painfully aware. Therefore, rather than reminisce about the mini-revolutions, palace coups, and constitutional crises of my own times, I would like instead to look at the physics of particles and fields from the perspective of the present moment, and try to describe where I think we are now.

To talk about where we are now, it is necessary to have some rough idea of where we wanted to go. Different physicists have different motivations, and I can only speak with certainty about my own. To me, the reason for spending so much effort and money on elementary particle research is not that particles are so interesting in themselves—if I wanted a perfect image of tedium, one million bubble chamber photographs would do very well—but rather that as far as we can tell, it is in the area of elementary particles and fields (and perhaps also of cosmology) that we will find the ultimate laws of nature, the few simple general principles which determine why all of nature is the way it is. I am not under any illusion that discoveries in elementary particle physics are going to make life any easier for the biologist or solid-state physicist—it may well be that we already know enough about atoms, radiation, and so on for all such purposes. I am also not under any illusion that research on elementary particles is the only kind of "basic" science—in fact, it seems to me that the truly great scientific revolutions like Darwin's theory of evolution were great because they provided crucial missing links in our understanding of the deductive order of nature, whether these links were at the roots or way out on the branches of the deductive tree. The search for the ultimate laws of nature is only one of the aims of basic science, but it is this aim that is the particular concern of the physics of particles and fields, and so, when I ask where we in particle physics are now, I mean how far are we from understanding the laws of nature in their simplest and most general form. The answer I would like to propose here is that, although we are very far from this goal, we may be much closer than is generally realized.

The most fundamental principles of physics that we know are those provided by the two great discoveries of the 20th century—relativity and quantum mechanics. Relativity (strictly speaking, special relativity) sets the space-time stage on which physical processes are played out, and quantum mechanics provides the language, a language of probabilities, in which the script is written. However, relativity and quantum mechanics do not immediately provide either the script or the cast of characters; they don't obviously lead to a unique theory of elementary particles. One way to estimate how far we are from understanding the ultimate laws of nature is to ask how many additional assumptions we need to add to relativity and quantum mechanics to be able to deduce the observed properties of the elementary particles. If many additional assumptions are needed, then we are obviously far away from the ultimate laws, while if relativity and quantum mechanics are nearly sufficient by themselves, then we may not be so far from our goal.

The reason I take such an optimistic view of where we are now is that relativity and quantum mechanics, taken together but without any additional assumptions, are extraordinarily restrictive principles. Quantum mechanics without relativity would allow us to conceive of a great many possible physical systems. Open any textbook on nonrelativistic quantum mechanics and you will find a rich variety of made-up examples—particles in rigid boxes, particles on springs, and so on—which don't exist in the real world but are perfectly consistent with the principles of quantum mechanics. The same is true of relativity without quantum mechanics. However, when you put quantum mechanics together with relativity, you find that it is nearly impossible to conceive of any possible physical systems at all. Nature somehow manages to be both relativistic and quantum mechanical, but these two requirements restrict it so much that it has only a limited choice of how to be—hopefully a very limited choice.

I would like to try to explain this crucial near-incompatibility of relativity and quantum mechanics in nonmathematical terms. The arguments I will use are not precisely the same as those I would present to an audience consisting solely of physicists, but close enough.

In my view, the essence of the conflict between relativity and quantum mechanics arises from the peculiar view of time introduced into physics by special relativity. Before Einstein, one simply spoke of the time of an event, with the tacit assumption that all observers would see an event occurring at the same time. However, in relativity theory not only the time of an event, but even the order of events, depends on the state of motion of the observer. That is, if one observer sees event A occur

The author is professor of physics at the Massachusetts Institute of Technology, Cambridge 02139. This article is the text of a lecture delivered when he received the 1973 J. Robert Oppenheimer memorial prize at the University of Miami Center for Theoretical Studies, Coral Gables, Florida.

before B, another observer who is moving sufficiently fast relative to the first may see B occur before A. It is important to qualify this immediately: The order of events depends upon the state of motion of the observer only if the two events are close enough in time and far enough in space so that no light signal can propagate from one to the other—otherwise, all observers will agree on which comes first. For this reason, no paradox can arise in classical relativistic mechanics—if one event *causes* another event, then it must be possible for some sort of signal to travel from one event to the other, and since no such signal can exceed the speed of light, the two events must be close enough in space and far enough in time so that all observers will agree that the cause precedes the event. This is especially the case if one event is the creation of a particle and the other is its destruction; here the particle itself is the signal which connects the two events, and as long as the particle's speed cannot exceed that of light, all observers will agree that the particle is created before it is destroyed.

However, in quantum mechanics the uncertainty principle prevents us from getting off so easily. The uncertainty principle tells us that we cannot simultaneously measure the position and the speed of a particle at a given time with unlimited accuracy, because the act of measuring the position and time changes the speed, and vice versa. Thus, if we specify the exact positions and times of two events, then we cannot be sure that a particle traveling from one to the other will have a speed less than that of light. If one observer sees a particle created at A and destroyed at B, then the possibility that the particle travels from one event to the other at speeds greater than that of light means that there is a possibility that a sufficiently rapid observer will see the particle destroyed at B before it is created at A!

We only know one way out of this difficulty: The rapidly moving observer who sees event B before A must interpret what he sees as the creation of a particle at B which is subsequently destroyed at A. But how is this possible? If a particle with one unit of positive charge is created at A and destroyed at B, then, because charge is conserved, the rest of the system must lose one unit of charge at A and gain one unit at B. But if a second observer sees the particle created at B and destroyed at A, he must also see the system gain one unit of charge at B and lose one at A, so he must interpret what he sees as the creation and subsequent destruction of a particle carrying one unit of *negative* charge. This line of reasoning obviously applies to any kind of particle, and to any conserved quantity like charge, and leads us to the conclusion that relativistic quantum theory must have an antiparticle, with opposite values of charge and all other conserved attributes, for each type of particle, except for those purely neutral particles (like photons) that are their own antiparticles.

Of course, to carry out actual physical calculations, we need more than this hand-waving about the uncertainty principle—we need a mathematical formalism which incorporates the deep relation between particle creation or destruction and antiparticle destruction or creation. This formalism is known as quantum field theory. It was developed during the 1920's and 1930's by Dirac, Jordan, Wigner, Pauli, Weisskopf, and others, but it received its modern form in the late 1940's through the work of Feynman, Dyson, Schwinger, and Tomonaga. In fact, the particular contribution made in the 1940's was the development of a method of calculation which preserves the symmetry between particles and antiparticles at every stage in the calculations. When a particle theorist uses Feynman diagrams to map out a series of steps in a physical process, every line in the diagrams represents either a particle running from A to B or an antiparticle running from B to A. It is for this reason, and only for this reason, that such calculations are manifestly consistent with special relativity at every stage. I don't mean to say that we have proved that any dynamical theory consistent with relativity and quantum mechanics must take the form of a quantum field theory; but I believe this is true, and at any rate there are no known counterexamples.

Once we accept the general ideas of quantum field theory, a great many other conclusions follow, beyond the mere existence of antiparticles. Here is a partial list:

1) There is a symmetry principle known as CPT invariance, which says that two observers, who distinguish particles and antiparticles oppositely *and* distinguish right and left oppositely, *and* distinguish the past from the future oppositely, will measure the same probabilities for corresponding events.

2) There is a connection between spin and statistics. That is, particles either avoid being in the same state (like electrons in an atom) or prefer being in the same state (like photons in a laser) depending on their spin.

3) The only kinds of force between particles are those which arise from exchange of other particles. This conclusion, expressed mathematically as an hypothesis of analyticity, can itself be used as the basis for a separate logical development, known as S-matrix theory, in which one tries to avoid the explicit use of quantum field theory.

4) The only kinds of force which can have long range (in the sense that an inverse-square law like Newton's theory of gravitation is long range, but an inverse cube is not) are essentially just those we know do exist, that is, gravitation, electricity, and magnetism. (There are important qualifications to this, having especially to do with possible scalar fields, but we need not go into this here.)

5) There cannot exist any charged particles of zero mass.

These conclusions are reasonably noncontroversial, by which I mean not that they have been proved rigorously, but that most theorists would expect that they could be proved in any relativistic quantum theory. Also, they are, as far as our experiments show, actually true of nature. However, impressive as they are, these deductions do not take us very far toward an understanding of all of particle physics on the basis of relativity and quantum mechanics alone. In order to make further progress toward this goal, we have to screw up our courage, and draw some conclusions that are very controversial indeed.

This new line of argument starts with the problem of infinities. As an example, consider the inverse square law of force, familiar from Newton's law of gravitation or Coulomb's law of electrostatics. If the force between two pieces of matter decreases like the inverse square of their separation, then the force must increase without limit as the pieces come closer together. In particular, the energy of a point particle produced by the forces between its parts must be infinite. To be sure, this sort of infinity can arise in nonquantum relativistic mechanics, but the problem of the infinities is very much worse in relativistic quantum mechanics, because we have so little flexibility in fooling around with the laws of force, and also because the symmetry between particle annihilation and antiparticle creation prevents us from putting any limit on the number of particles that can be produced at high energy. In fact, calculations in quantum field theory are infested with a rich variety of nonsensical infinities, which paralyzed theoretical research until the developments of the late 1940's revealed a way out.

In certain narrowly restricted kinds of theory, it is possible to absorb all the infinities of the theory into a redefinition, a "renormalization," of the physical parameters of the theory. For instance, suppose that the quantity m which we insert in our equations to represent the mass of the electron is not the true mass at all, but a negative infinite quantity, which when added to the positive infinite energy of electrostatic repulsion yields a finite quantity—the observed electron mass. Theories in which the infinities can be absorbed in this way are called "renormalizable."

Now, it is a characteristic of any renormalizable theory that once we specify the types of elementary particles it describes, and give the values of a *finite* number of fundamental constants, the whole theory is uniquely determined; there is only one possible way for nature to behave. For instance, in the theory of electrons and photons, these constants are just the mass and charge of the electron; once we specify the values of these constants, there is only one possible theory of photons and electrons. Furthermore, this theory seems to work. The renormalizable theory of photons and electrons known as quantum electrodynamics has been used in a variety of calculations, and the results agree with experiment to a fantastic degree of accuracy. Without renormalization theory, we would not only be unable to do these calculations, we would not even understand why the simplest properties of the electron, such as its strength as a magnet, have the values they have. For many years it was believed to be impossible to construct any renormalizable theory of the weak interactions (interactions responsible for certain kinds of radioactivity, and also for certain steps in the nuclear processes that heat the stars) but this has now been shown to be possible if the theory of weak interactions is unified with the theory of electromagnetism. (My own work in recent years has been mostly in this area.)

But is it really true that nature can find no way of eliminating infinities other than renormalization? That is, is renormalizability really a logical consequence of quantum mechanics and special relativity? The question is open. Some of my colleagues point out that a nonrenormalizable theory always has built into it a characteristic unit of length, and that physical processes which involve separations and wavelengths much larger than the unit of length will automatically look as if the underlying theory were renormalizable. I understand that something like this happens near a critical point in the theory of phase transitions. Personally, I have always leaned toward an orthodox renormalizationist position. This is partly because when I was a graduate student renormalization theory was the branch of elementary particle physics which was toughest mathematically, so I made a large investment of time in learning the theory, and, since I didn't understand what I read, in writing papers about it. One of these papers actually used the Heine-Borel theorem, a fact of which I've always been proud. (With advancing age, I now tend to choose problems because they are mathematically easy, rather than hard. This is known as wisdom.) However, a deeper reason for taking renormalization theory seriously is just that it is such a restrictive theory, and it is our business in particle theory to search out the restrictions which limit nature's freedom.

Where then are we now? If we accept quantum field theory and renormalizability as inescapable consequences of quantum mechanics and relativity, then we find such powerful constraints that most of the freedom of choice of nature is gone. Therefore, by the test I mentioned earlier, we are justified in supposing that quantum mechanics and relativity take us pretty far toward the ultimate laws of nature. But not far enough. We still have very little idea why the particles that exist are the ones that must exist, or why the constants of nature have the values they have. Our best guess is that the answer has something to do with the symmetries of nature, but that is another story.

Finally, I want to admit that the implicit background of what I have said here, a picture of the sciences branching out in logical order from particle physics, which itself has a few basic principles more or less like the principles of relativity and quantum mechanics, may be entirely wrong. Perhaps the logical tree isn't a tree at all, but something else, perhaps something with loops. For instance, according to a joke that went around when I was an undergraduate, the laws of nature are not fixed at all, but are revised from time to time by a committee of dead physicists in heaven. If so, then there is a logical circularity in nature, with particle physics following from defunct psychology, and vice versa. Who knows? More seriously, the laws of nature are discovered by human beings, and it may not be possible permanently to divorce the content of these laws from the psychology of their discoverers. Or perhaps there is no logical order to nature at all. Ernst Mach rejected the whole notion of a hierarchy of the sciences, and he resisted the atomic explanation of chemistry, because he didn't believe that chemistry *needed* an explanation in terms of more fundamental truths.

In the last analysis, it seems to me that the best reason for believing in a deductive order of nature with its roots in particle physics is that it allows us to make sense in asking, not only *how* nature behaves, but *why* it behaves the way it does. We feel we ought to know how the facts of nature follow from the ultimate laws of nature, and, not knowing, we feel a sense of mystery which helps to direct us to the work that still needs to be done.

Notes

1. I thank G. Holton, L. Weinberg, and V. F. Weisskopf for their help in the preparation of this lecture.