PAUL A. M. DIRAC

Theory of electrons and positrons

Nobel Lecture, December 12, 1933

Matter has been found by experimental physicists to be made up of small particles of various kinds, the particles of each kind being all exactly alike. Some of these kinds have definitely been shown to be composite, that is, to be composed of other particles of a simpler nature. But there are other kinds which have not been shown to be composite and which one expects will never be shown to be composite, so that one considers them as elementary and fundamental.

From general philosophical grounds one would at first sight like to have as few kinds of elementary particles as possible, say only one kind, or at most two, and to have all matter built up of these elementary kinds. It appears from the experimental results, though, that there must be more than this. In fact the number of kinds of elementary particle has shown a rather alarming tendency to increase during recent years.

The situation is perhaps not so bad, though, because on closer investigation it appears that the distinction between elementary and composite particles cannot be made rigorous. To get an interpretation of some modern experimental results one must suppose that particles can be created and annihilated. Thus if a particle is observed to come out from another particle, one can no longer be sure that the latter is composite. The former may have been created. The distinction between elementary particles and composite particles now becomes a matter of convenience. This reason alone is sufficient to compel one to give up the attractive philosophical idea that all matter is made up of one kind, or perhaps two kinds of bricks.

I should like here to discuss the simpler kinds of particles and to consider *what can be inferred about them from purely theoretical arguments.* The simpler kinds of particle are:

 (i) the photons or light-quanta, of which light is composed;

 (ii) the electrons, and the recently discovered positrons (which appear to be a sort of mirror image of the electrons, differing from them only in the sign of their electric charge) ;

(iii) the heavier particles - protons and neutrons.

Of these, I shall deal almost entirely with the electrons and the positrons - not because they are the most interesting ones, but because in their case the theory has been developed further. There is, in fact, hardly anything that can be inferred theoretically about the properties of the others. The photons, on the one hand, are so simple that they can easily be fitted into any theoretical scheme, and the theory therefore does not put any restrictions on their properties. The protons and neutrons, on the other hand, seem to be too complicated and no reliable basis for a theory of them has yet been discovered.

The question that we must first consider is how theory can give any information at all about the properties of elementary particles. There exists at the present time a general quantum mechanics which can be used to describe the motion of any kind of particle, no matter what its properties are. The general quantum mechanics, however, is valid only when the particles have small velocities and fails for velocities comparable with the velocity of light, when effects of relativity come in. There exists no relativistic quantum mechanics (that is, one valid for large velocities) which can be applied to particles with arbitrary properties. Thus when one subjects quantum mechanics to relativistic requirements, one imposes restrictions on the properties of the particle. In this way one can deduce information about the particles from purely theoretical considerations, based on general physical principles.

This procedure is successful in the case of electrons and positrons. It is to be hoped that in the future some such procedure will be found for the case of the other particles. I should like here to outline the method for electrons and positrons, showing how one can deduce the spin properties of the electron, and then how one can infer the existence of positrons with similar spin properties and with the possibility of being annihilated in collisions with electrons.

We begin with the equation connecting the kinetic energy W and momentum $p_r, (r = 1, 2, 3)$, of a particle in relativistic classical mechanics

$$\frac{W^2}{c^2} - p_r^2 - m^2c^2 = 0 \qquad (1)$$

From this we can get a wave equation of quantum mechanics, by letting the left-hand side operate on the wave function ψ and understanding W and p_r to be the operators $ih\partial/\partial t$ and $-ih\partial/\partial x_r$. With this understanding, the wave equation reads

$$\left[\frac{W^2}{c^2} - p_r{}^2 - m^2c^2\right]\psi = 0$$

Now it is a general requirement of quantum mechanics that its wave equations shall be linear in the operator W or $\partial/\partial t$, so this equation will not do. We must replace it by some equation linear in W, and in order that this equation may have relativistic invariance it must also be linear in the p's.

We are thus led to consider an equation of the type

$$\left[\frac{W}{c} - \alpha_r p_r - \alpha_0 mc\right]\psi = 0$$

This involves four new variables α_r and α_0, which are operators that can operate on ψ. We assume they satisfy the following conditions,

for

$$\alpha_\mu{}^2 = 1 \qquad \alpha_\mu \alpha_\nu + \alpha_\nu \alpha_\mu = 0$$

$$\mu \neq \nu \text{ and } \mu, \nu = 0, 1, 2, 3$$

and also the α's commute with the p's and W. These special properties for the α's make Eq. (3) to a certain extent equivalent to Eq. (2), since if we then multiply (3) on the left-hand side by $W/c + \alpha_r p_r + \alpha_0 mc$ *we* get exactly *(2)*.

The new variables α which we have to introduce to get a relativistic wave equation linear in W, give rise to the spin of the electron. From the general principles of quantum mechanics one can easily deduce that these variables *a* give the electron a spin angular momentum of half a quantum and a magnetic moment of one Bohr magneton in the reverse direction to the angular momentum. These results are in agreement with experiment. They were, in fact, first obtained from the experimental evidence provided-by spectroscopy and afterwards confirmed by the theory,

The variables α also give rise to some rather unexpected phenomena concerning the motion of the electron. These have been fully worked out by Schrödinger. It is found that an electron which seems to us to be moving slowly, must actually have a very high frequency oscillatory motion of small amplitude superposed on the regular motion which appears to us. As a result of this oscillatory motion, the velocity of the electron at any time equals the velocity of light. This is a prediction which cannot be directly verified by experiment, since the frequency of the oscillatory motion is so high and its

amplitude is so small. But one must believe in this consequence of the theory, since other consequences of the theory which are inseparably bound up with this one, such as the law of scattering of light by an electron, are confirmed by experiment.

There is one other feature of these equations which I should now like to discuss, a feature which led to the prediction of the positron. If one looks at Eq. (1), one sees that it allows the kinetic energy W to be either a positive quantity greater than mc^2 or a negative quantity less than $-mc^2$. This result is preserved when one passes over to the quantum equation (2) or (3). These quantum equations are such that, when interpreted according to the general scheme of quantum dynamics, they allow as the possible results of a measurement of W either something greater than mc^2 or something less than $-mc^2$.

Now in practice the kinetic energy of a particle is always positive. We thus see that our equations allow of two kinds of motion for an electron, only one of which corresponds to what we are familiar with. The other corresponds to electrons with a very peculiar motion such that the faster they move, the less energy they have, and one must put energy into them to bring them to rest.

One would thus be inclined to introduce, as a new assumption of the theory, that only one of the two kinds of motion occurs in practice. But this gives rise to a difficulty, since we find from the theory that if we disturb the electron, we may cause a transition from a positive-energy state of motion to a negative-energy one, so that, even if we suppose all the electrons in the world to be started off in positive-energy states, after a time some of them would be in negative-energy states.

Thus in allowing negative-energy states, the theory gives something which appears not to correspond to anything known experimentally, but which we cannot simply reject by a new assumption. We must find some meaning for these states.

An examination of the behaviour of these states in an electromagnetic field shows that they correspond to the motion of an electron with a positive charge instead of the usual negative one - what the experimenters now call a positron. One might, therefore, be inclined to assume that electrons in negative-energy states are just positrons, but this will not do, because the observed positrons certainly do not have negative energies. We can, however, establish 'a connection between electrons in negative-energy states and positrons, in a rather more indirect way.

We make use of the exclusion principle of Pauli, according to which

there can be only one electron in any state of motion. We now make the assumptions that in the world as we know it, nearly all the states of negative energy for the electrons are occupied, with just one electron in each state, and that a uniform filling of all the negative-energy states is completely un-observable to us. Further, *any unoccupied negative-energy state, being a departure from uniformity, is observable and is just a positron.*

An unoccupied negative-energy state, or *hole,* as we may call it for brevity, will have a positive energy, since it is a place where there is a shortage of negative energy. A hole is, in fact, just like an ordinary particle, and its identification with the positron seems the most reasonable way of getting over the difficulty of the appearance of negative energies in our equations. On this view the positron is just a mirror-image of the electron, having exactly the same mass and opposite charge. This has already been roughly confirmed by experiment. The positron should also have similar spin prop-erties to the electron, but this has not yet been confirmed by experiment.

From our theoretical picture, we should expect an ordinary electron, with positive energy, to be able to drop into a hole and fill up this hole, the energy being liberated in the form of electromagnetic radiation. This would mean a process in which an electron and a positron annihilate one another. The converse process, namely the creation of an electron and a positron from electromagnetic radiation, should also be able to take place. Such processes appear to have been found experimentally, and are at present being more closely investigated by experimenters.

The theory of electrons and positrons which I have just outlined is a self-consistent theory which fits the experimental facts so far as is yet known. One would like to have an equally satisfactory theory for protons. One might perhaps think that the same theory could be applied to protons. This would require the possibility of existence of negatively charged protons forming a mirror-image of the usual positively charged ones. There is, how-ever, some recent experimental evidence obtained by Stern about the spin magnetic moment of the proton, which conflicts with this theory for the proton. As the proton is so much heavier than the electron, it is quite likely that it requires some more complicated theory, though one cannot at the present time say what this theory is.

In any case I think it is probable that negative protons can exist, since as far as the theory is yet definite, there is a complete and perfect symmetry between positive and negative electric charge, and if this symmetry is really fundamental in nature, it must be possible to reverse the charge on any kind

of particle. The negative protons would of course be much harder to produce experimentally, since a much larger energy would be required, corresponding to the larger mass.

If we accept the view of complete symmetry between positive and negative electric charge so far as concerns the fundamental laws of Nature, we must regard it rather as an accident that the Earth (and presumably the whole solar system), contains a preponderance of negative electrons and positive protons. It is quite possible that for some of the stars it is the other way about, these stars being built up mainly of positrons and negative protons. In fact, there may be half the stars of each kind. The two kinds of stars would both show exactly the same spectra, and there would be no way of distinguishing them by present astronomical methods.