The two major theories of twentieth century physics—relativity and quantum theory—were obtained quite differently. The essentials of relativity theory were formulated in the first decade of this century. But the discovery of quantum theory spanned three decades and involved many lucky 'accidents' like Planck's hypothesis, Bohr's atomic theory, de Broglie's hypothesis, Heisenberg's noncommuting dynamical variables and Schrödinger's wave equation. The attempts to unify relativity and quantum theory again involved many accidents and lucky breaks. The first formulation of a relativistic wave equation with a standard probability interpretation was made by Dirac in 1928. But when the physical interpretation of the Dirac equation was finally worked out, it turned out that the wavefunction was not a Schrödinger wave function but a Heisenberg operator! It is only in the interaction-free limit that it could be interpreted as a wavefunction. Dirac assiduously searched for a relativistic Schrödinger equation describing a composite particle; and it is only in the nineteen seventies that he succeeded in this search. However, he found this description was inconsistent in the presence of electromagnetic interaction, and Dirac seems to have abandoned this programme some time ago. N. Mukunda and his collaborators, including E.C.G. Sudarshan found a way out of this difficulty. Dirac himself hailed this as a major advance as this appears as the culmination of the theoretical framework begun by Schrödinger and so successfully pursued by Dirac. In this article Professor Mukunda traces these exciting developments and the involvement he and his collaborators have had in them. He points out that this is a framework which should fit into any relativistic system rather than being a specific theory.

THE DIRAC EQUATION—OLD AND NEW

N. MUKUNDA

Centre for Theoretical Studies and Department of Physics, Indian Institute of Science, Bangalore 560 012, India

THE DIRAC EQUATION OF 1928

THE relativistic wave equation for the electron was discovered by Dirac in 1928[1]. Without fear of exaggeration one can say that it is one of a small number of fundamental equations on which most of present day physics rests — the others being Maxwell's equations for electromagnetism, Einstein's equations for gravitation, the Schrödinger equation of quantum mechanics, and the Yang-Mills equations for non Abelian gauge theories of more recent vintage.

The Schrödinger equation can rightly be regarded as the replacement in quantum mechanics for Newton's equations of motion in classical mechanics. It thus sets the general pattern for description of dynamical systems in the quantum domain, and so can be used for any quantum system. On the other hand the remaining equations mentioned above are more specific in content. Thus Maxwell's equations describe the behaviour of the electromagnetic field and the manner in which this field interacts with other physical systems; Einstein's equations accomplish the same things for gravitation; the Dirac equation gives an excellent account of the properties and behaviour of the electron, the μ meson and such other leptons; while the equations of Yang and Mills are a nontrivial and beautiful generalization of the structure of

Maxwell's equations from Abelian to non Abelian internal symmetry.

Dirac's equation of 1928 was the first spectacularly successful combination of the principles of quantum mechanics and of special relativity. There had been a previous attempt to combine these principles, and it had led to the equation of Klein and Gordon. Even Niels Böhr had felt that that equation was quite satisfactory. Dirac however was convinced that it was not, for two reasons: it was in conflict with the probabilistic interpretation of quantum mechanics, and it did not conform to the principles of the transformation theory of quantum mechanics which Dirac himself had set up. Apart from this, it is interesting to read his own description of the motivations and the beginnings of the work leading to his equation: he started 'playing with the equations rather than trying to introduce the right physical idea. It is my habit that I like to play about with equations, just looking for mathematical relations which may be do not have any physical meaning at all'.... 'I was not trying to solve directly some physical problem but to look for some pretty mathematics'. In Dirac's hands, such motivations led to profound physical consequences!

Dirac originally wrote his equation in the form

$$i\hbar \frac{\partial \psi}{\partial t} = (c\, \alpha . \mathbf{p} + \beta\, mc^2)\, \psi . \qquad (1)$$

The quantities α and β are four-dimensional hermitian matrices obeying characteristic algebraic relations. They act on the wave-function ψ which is a four-component column vector, each entry a function of space and time coordinates. An equivalent but 'more relativistic' notation expresses the above equation in the form

$$(\gamma^\mu \partial_\mu + m)\, \psi = 0 \qquad (2)$$

using matrices γ^μ related to the α and β. Because of the presence of these matrices, relativistic invariance of the equation leads to a specific transformation law for ψ under Lorentz transformations: its components go into linear combinations of themselves according to the spinor representations of the Lorentz group. As a natural consequence of this transformation law, this equation showed the existence of the intrinsic spin angular momentum for the electron. It also turned out to be possible to extend the above equation in a simple way to include the effect of an external electromagnetic field: the operator ∂_μ of differentiation with respect to the space-time coordinates x^μ is to be replaced by the combination $\partial_\mu - ie\, A_\mu(x)$ where $A_\mu\,(x)$ is a vector potential for the given external field, so the equation then reads

$$(\gamma^\mu (\partial_\mu - ie\, A_\mu) + m)\, \psi = 0. \qquad (3)$$

When this was done, on the one hand, the equation gave automatically the correct value for the spin magnetic moment of the electron; and on the other, when A_μ represents the Coulomb field of a proton, the fine structure of the energy levels of the hydrogen atom came out correctly.

There was one other feature which initially caused considerable confusion in the physical interpretation of the equation: this was the occurrence of negative energy solutions symmetrically with the positive energy ones describing the electron. Dirac's initial suggestion that these might correspond to protons was soon shown, especially through the analysis of Weyl, to be untenable. Dirac soon found a new interpretation and predicted that these solutions described 'anti electrons' – particles with the same mass and spin as the electron, but with opposite charge[2]. When these particles were experimentally detected in 1932, that became yet another major triumph of Dirac's equation.

THE MAJORANA EQUATION—1932

At just about this time, in an effort to get rid of this 'problem', the Italian physicist

Majorana constructed a really ingenious wave equation which had no negative energy solutions at all[3]. He did however retain two essential features of Dirac's equation: the operators ∂_μ appeared only linearly, and the equation was consistent with the probability interpretation of quantum mechanics.

Majorana's equation is usually written as

$$(\Gamma^\mu \partial_\mu + m_\circ)\,\psi = 0 \qquad (4)$$

In the Dirac case, the essential algebraic reason for the occurrence of both signs of the energy is the fact that the matrix γ^0 [the same as β in (1)] has both positive and negative eigenvalues. Majorana therefore wanted the matrix Γ^0 in his equation to be positive definite. This could however be accomplished only by allowing the wave-function ψ to have infinitely many independent components! Thus Majorana's equation is the first of the so-called infinite-component relativistic wave equations. Moreover, it involves four matrices Γ^μ which are all infinite dimensional and hermitian, with Γ^0 being nonsingular as well. Under a Lorentz transformation, the infinite number of components of ψ go into linear combinations of themselves in an irreducible way. The Majorana wave function thus belongs to an infinite dimensional irreducible representation of the Lorentz group, unlike the four component Dirac wave function. This 'Majorana' representation of the Lorentz group happens also to be unitary and it is remarkable that Majorana was able to construct such a representation so soon after the inauguration of relativistic quantum mechanics. The systematic study of such representations of the Lorentz group came much later, in the work of Dirac, Harishchandra and Gel' fand and Naimark.

There are in actual fact two equations of the form (4) built on two Majorana representations of the Lorentz group. As a consequence of the fact that ψ has infinitely many components, (4) turns out to describe not one particle with one mass and spin value, but a whole series or tower of particles. The 'integer spin' Majorana equation has solutions corresponding to particles with spins 0, 1, 2,..... but with the unphysical feature that the mass decreases monotonically as the spin increases: mass as a function of spin turns out to be

$$m\,(s) = m_0/(s + \tfrac{1}{2}) \qquad (5)$$

Similarly, the 'half-integer spin' equation yields spin values $s = 1/2, 3/2, 5/2, \ldots$ and the same unphysical mass-spin relation as above. For this reason, Majorana's equations have remained more or less a mathematical curiosity.

As mentioned earlier, the negative energy solutions were avoided by ensuring that Γ^0 was positive-definite. This however really applies only to the set of time-like solutions of (4). It was pointed out in 1948 by Bargmann that, essentially because all four matrices Γ^μ are hermitian, Majorana's equations possess light-like and unphysical space-like solutions as well. Thus in effect Majorana's construction amounts to trading negative time-like solutions for space-like ones. While this circumstance, if anything, makes the Majorana equations a little bit more unphysical, the equations nevertheless do possess a rather intricate and potentially useful mathematical structure. A thorough analysis of these equations and their solutions, from the viewpoint of the representation theory of the Poincare group, was made by some of us sometime ago, and the completeness of the set of all solutions was also proved[4,5]. This study showed that, at least in a mathematical sense, there is no inconsistency in including the effect of an external electromagnetic field in the Majorana equation much like the passage in the Dirac case from equation (2) to (3).

THE NEW DIRAC EQUATION—1971

We have devoted some space to a description of the Majorana equations because their underlying mathematical structure turns out to be relevant to the new Dirac equations, to which we now turn.

It has been known for some time, to a considerable extent through yet another piece of work by Dirac[6] that the two unitary 'Majorana' representations of the Lorentz group, which figure in the wave equations (4) can be built up in a simple and tractable way using quantum mechanical 'position' and 'momentum' operators. Let us consider two such canonical pairs, that is two position operators q_1, q_2 and two momenta p_1, p_2 obeying the usual quantum mechanical commutation relations

$$[q_j, p_k] = i\,\delta_{jk} \qquad (6)$$

We assume of course that these are hermitian. There is an essentially unique way in which to set up a Hilbert space H_0 on which these relations can be realised irreducibly. One description of this space, for instance, is to imagine it to consist of square-integrable wave functions $\psi\ (q_1, q_2)$ on which the momenta p_j act as differential operators with respect to q_j; other descriptions are of course possible. It now turns out that on this space a certain unitary representation of the de Sitter group SO (3,2) can be naturally and easily constructed. If one writes down all possible bilinear expressions in the four operators q_j, p_j, one has ten independent symmetric combinations and six independent antisymmetric ones. But because of the postulated commutation relations (6), the latter are quite trivial: each antisymmetric expression is either zero or a pure number. The ten symmetric combinations which survive in fact act as generators of the unitary SO(3,2) representation acting on H_0. By restricting oneself to the subgroup of SO(3,2) corresponding to Lorentz transformations on space-time, the subgroup SO(3,1), one finds that one is dealing in fact with the two Majorana representations! The Lorentz transformations require only six independent symmetric bilinears in q_j and p_j to generate them: the remaining four operators are none other than the 'matrices' Γ^{μ} in Majorana's wave equation.

To summarize: the Majorana representations of the Lorentz group are most naturally realized in terms of two quantum mechanical position-momentum pairs. A simple way to picture the wave functions of this space is to take them to be functions of q_1 and q_2: one could have taken the wave function ψ in Majorana's equation (4) to be of this kind, in addition to depending on the space-time coordinates x^{μ}.

We now come to Dirac's new wave equation, presented in 1971[7]: it really stands 'in between' his equation of 1928 and Majorana's of 1932, since elements of structure from both sources are combined in an ingenious fashion, and it can surely be regarded as yet another beautiful product of Dirac's habit of playing with equations!

The new Dirac equation can be written in the form:

$$(\gamma^{\mu}\partial_{\mu}+m)\,Q\psi = 0. \qquad (7)$$

Let us briefly describe the elements that go into it and the way they are combined. The γ^{μ} are the 4×4 matrices that are familiar from the relativistic electron equation (2). The symbol Q denotes a 4-component column vector whose entries are the four Hermitian operators q_1, q_2, p_1, p_2 in that sequence. So the matrices γ^{μ} act on the operator column vector Q. The wave function ψ depends on the space-time variables x^{μ} and moreover, for each x, is a vector lying in the space H_0! Thus the operator ∂_{μ} acts on the x-dependence of ψ, while the q's and p's comprising Q act on ψ by virtue of its being a vector in H_0 (q's and p's are operators on H_0). This new Dirac equation is, in actual fact, a set of four equations to be obeyed by one quantity ψ

which is of the same nature as the wave function in Majorana's wave equations.

Equations (7) are relativistically invariant. The new Dirac wave-function ψ is defined to transform under Lorentz transformations exactly like the ψ in (4), *i.e.*,*via* the Majorana representation. With respect to this representation of the Lorentz group it transpires that the four-component operator column vector Q behaves as a *spinor* just like the ψ in the old Dirac equation. Thus, all in all, Q acting on the new ψ behaves just like the old ψ under Lorentz transformations, and this makes equations (7) relativistically invariant.

It is when one searches for the solutions to (7) that one finds some perhaps unexpected features. First of all it turns out that if ψ obeys (7), it then necessarily obeys the Klein Gordon equation for mass m, as well as Majorana's wave equation (4) (with m in place of m_0). The former means that the new Dirac equation has solutions corresponding to physical particles of mass m alone, and the latter then means that only solutions with positive energy appear. These results can also be confirmed by explicit calculation. It is interesting to see what the wave function ψ looks like when the energy and momentum have definite values P_μ : apart from factors, we have

$$\psi = \exp(-iP\cdot x)\cdot\exp\Big\{-\tfrac{1}{2}(q_1^2 + q_2^2 + i\,P_1\,(q_1^2 - q_2^2) - 2i\,P_2\,q_1\,q_2)/(P_0 + P_3)\Big\} \quad (8)$$

Had we searched for a solution with negative energy, we would have found that the dependence on q_1, q_2 is non-normalizable and so unacceptable.

This solution, and therefore the new Dirac equation, describes a particle with mass m, spin zero and positive energy. The system (7) is thus a clever way of isolating just one out of the infinity of solutions of the 'integer spin' *Majorana equation* (4), *and in* particular of avoiding the unphysical space-like solutions of the latter. But this, as we shall soon see, is achieved at a rather heavy price.

It is tempting to interpret the above solution of the new Dirac equation as describing, in a relativistically consistent way, a composite particle with some internal structure corresponding to two constituents. However, this must be done with caution. We saw *earlier that with respect to* Lorentz transformations Q behaves like a spinor, similar to the ψ of the old Dirac equation. Thus the 'positions' q_1 and q_2 do not transform vectorially like space-time positions, and one cannot think of the system described by Dirac's new equation as a composite one consisting of two conventional particles. In spite of this limitation, Dirac has developed a semiclassical picture of the object described by his equation[8]: it is an extended object in the shape of a shell which pulsates as it moves along. This pulsation is the analogue of the 'trembling motion' or Zitterbewegung in the case of the relativistic electron equation.

One gets the feeling that very little use has been made, in the solution (8), of the rich internal structure that the new Dirac wavefunction ψ was originally endowed with. This is because one is ultimately describing a spin zero particle with a unique mass m, and nothing else. Attempts have been made by Biedenharn and his collaborators to generalise Dirac's equation to get particles with non zero spin value[9]. However, in these generalisations, the equations involve progressively higher space-time derivatives of ψ.

THE PROBLEM OF ELECTROMAGNETIC INTERACTION

We come now to the major problem faced by the new Dirac equation — a problem that forced Dirac some years ago to give up this line of inquiry. If we imagine the extended object carrier charge e and is placed in an external electromagnetic field with vector potential $A_\mu(x)$, and if we suppose that the effect of the field on the system is given by the replacement $\partial_\mu \rightarrow \partial_\mu - ie\, A_\mu(x)$ as in the cases of the old Dirac equation as well as Majorana's equation, we obtain an inconsistency. That is to say, for a nonvanishing external field, the system

$$(\gamma^\mu(\partial_\mu - ie\, A_\mu(x)) + m)\, Q\psi = 0, \quad (9)$$

forces ψ itself to vanish! Recall that Majorana's equation does not face a similar problem.

This rather disappointing result was stated by Dirac and later explicitly proved by Biedenharn and collaborators. The essential reason for this result is this: if one tries to construct any antisymmetric bilinear expression in the internal variables q_j, p_j the result is either zero or a pure number. This comes about because these are *boson* variables obeying the simple commutation relations (6).

The higher spin generalisations of Biedenharn *et al* also suffer from this same problem.

Recently two possible ways out of this situation, intended to save the overall mathematical structure of Dirac's construction but permitting electromagnetic interactions, have been suggested. One of them involves going back to a classical Lagrangian framework and setting up a model that makes full use of the internal boson structure, so that after quantisation one describes a whole sequence of particles with mass increasing monotonically with respect to spin; and such that each particle in this spectrum can couple to the electromagnetic field[10]. Another alternative, which is fully quantum mechanical from the start, has been suggested by Sudarshan, Chiang and the present author[11,12], and Dirac has expressed the opinion that this might very well solve the problem. The idea here is to use paraboson internal variables q_j and p_j rather than Dirac's bosons[13]. This replacement does not harm the relativistic invariance of Dirac's new equation at all. Moreover, in the simplest version, the free equation again describes a spin zero, mass m particle with positive energies. The major difference is that antisymmetric bilinear expressions in the paraboson q_j and p_j do not vanish, so it appears now that one can introduce the electromagnetic interaction in a consistent way.

One difference between bosons and parabosons is this: we cannot any longer exhibit the internal structure as directly as in (8), since the wavefunction ψ is not simply a function of q_1, q_2 and x^μ. The operators q_1 and q_2 themselves are somewhat nontrivial and in particular do not commute. In spite of this, a straightforward generalisation to nonzero spins seems possible while retaining the basic form (7) of Dirac's new equation; and even with this generalisation, electromagnetic coupling remains consistent. Work is currently in progress to study these generalised equations in some detail, and to see if the kind of description of internal structure found by Dirac in the course of his 'playing about with equations, just looking for mathematical relations' may not after all be of relevance for the real world. This has happened so often with Dirac that it may well happen once more!

1. Dirac, P. A. M., *Proc. R. Soc.*, 1928, **A117**, 610.
2. Dirac, P. A. M., *Proc. R. Soc.*, 1931, **A133**, 60.
3. Majorana, E., *Nuovo Cimento*, 1932, **9**, 335.
4. Sudarshan, E. C. G. and Mukunda, N., *Phys. Rev.*, 1970, **D1**, 571.
5. Mukunda, N., *Phys. Rev.*, 1969, **183**, 1486.
6. Dirac, P. A. M., *J. Math. Phys.*, 1963, **4**, 901.
7. Dirac, P. A. M., *Proc. R. Soc.*, 1971, **A322**, 435.
8. Dirac, P. A. M., *Proc. R. Soc.*, 1972, **A328**, 1.
9. Biedenharn, L. C., Han, M. Y. and van Dam, H., *Phys. Rev.*, 1973, **D8**, 1735.
10. Mukunda, N., van Dam, H., and Biedenharn, L. C., Phys. Rev., 1980, **D22**, 1938.
11. Mukunda, N., Chiang, C. C. and Sudarshan, E. C. G., *Proc. R. Soc.*, 1981 (in press).
12. Sudarshan, E. C. G., Mukunda, N. and Chiang, C. C., *Phys. Rev.*, 1981 (in press).
13. Green, H. S., *Phys. Rev.*, 1953, **90**, 270.