Representations of the Dirac Equation

S. K. BOSE, A. GAMBA, AND E. C. G. SUDARSHAN*†
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received October 20, 1958)

A representation of the Dirac equation which displays its extreme relativistic properties is discussed. The "spin" appears naturally related to the "position."

I.

It is well known that the usual form of the Dirac equation for a particle of spin $\frac{1}{2}$ does not lend itself easily to a simple interpretation in terms of physical quantities. For example, the velocity operator has only $\pm c$ as eigenvalues and the usual relation between velocity and momentum is lost.

It was shown by Foldy and Wouthuysen[1] that, by means of a canonical transformation, the Dirac equation can be written in a form which can be more easily interpreted in terms of classical quantities. The simplest operators in this C (for classical) representation are immediately related to significant classical quantities. To distinguish these new operators from those having the same names in the usual D (for Dirac) representation, Foldy and Wouthuysen call their observables "mean," leaving the usual name to the D operators. We shall refer to the various quantities simply as D or C and use the same letter as an index, whenever ambiguity could arise.

The C representation is very convenient in discussing the nonrelativistic limit, when the momentum p is small compared to the mass m of the particle (limit towards Pauli equation).

The purpose of this paper is to investigate an E (for extreme relativistic) representation, which is convenient in discussing the extreme relativistic approximation, i.e., when the mass m is small compared to the momentum p of the particle (limit towards Weyl equation). The most simple E operators will be shown to possess a direct physical meaning.

II.

The Dirac Hamiltonian,

$$H^D=\boldsymbol{\alpha}\cdot\mathbf{p}+\beta m, \qquad (1D)$$

is transformed into the following:

C-representation Hamiltonian	E-representation Hamiltonian[2]
$H^C=\beta E, \qquad (1C)$	$H^E=\dfrac{\boldsymbol{\alpha}\cdot\mathbf{p}}{p}E, \qquad (1E)$

where $E^2=p^2+m^2$, by means of the following canonical transformation:

$$H^C=e^{iS}H^De^{-iS}, \qquad (2C) \qquad H^E=e^{iT}H^De^{-iT}, \qquad (2E)$$

$$S=-\frac{i}{2}\beta\frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{p}\tan^{-1}\left(\frac{p}{m}\right), \quad (3C) \qquad T=\frac{i}{2}\beta\frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{p}\tan^{-1}\left(\frac{m}{p}\right), \quad (3E)$$

$$e^{\pm iS}=\frac{E+m\pm\beta\boldsymbol{\alpha}\cdot\mathbf{p}}{\{2E(E+m)\}^{\frac{1}{2}}} \quad (4C) \qquad e^{\pm iT}=\frac{E+p\mp\beta(\boldsymbol{\alpha}\cdot\mathbf{p}/p)m}{\{2E(E+p)\}}. \quad (4E)$$

In both C and E representations the positive- and negative-energy states are kept separate. The projection operator

$$\Lambda_\pm{}^D=\tfrac{1}{2}(1\pm H^D/E), \quad (\Lambda_\pm{}^D)^2=\Lambda_\pm{}^D, \qquad (5D)$$

for positive- (and negative-) energy states in D representation, is correspondingly transformed by the same canonical transformation into the projection operator

$$\Lambda_\pm{}^C=\tfrac{1}{2}(1\pm\beta), \quad (5C) \qquad \Lambda_\pm{}^E=\tfrac{1}{2}(1\pm\boldsymbol{\alpha}\cdot\mathbf{p}/p). \quad (5E)$$

The similarity of the two representations is apparent: the quantity which multiplies E in the transformed Hamiltonians is constructed in a similar way from the surviving term of the original D representation when going to the appropriate C or E limit.

Other similarities of the two transformations will appear later. At this point we would like, instead, to stress the differences. First of all the two transformations, starting from D representation, are going in opposite directions, to different limits. D representation is, so to speak, somewhat in between C and E representations, as suggested by the alphabetic order of the letters used to denote them.

Secondly, by simply looking at Eqs. (1C), (1D) one easily realizes that C representation is isotropic in the ordinary 3-dimensional space, where E representation, containing explicitly the momentum $\mathbf{p}$, establishes a kind of polarity in space. This is to be expected, since at the extreme relativistic limit the momentum still determines a privileged direction in space, whereas the extreme classical limit goes over to the particle at rest.

---

* On leave of absence from the Tata Institute of Fundamental Research, Bombay, India.

† Present address: Jefferson Physical Laboratory, Harvard University, Cambridge, Massachusetts.

[1] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1949), hereafter referred to as F.-W.

[2] *Note added in proof.*—This transformation has been previously given by M. Cini and B. Touschek [Nuovo cimento **7**, 422 (1958)].

TABLE I. Relevant E operators in both E and D representations.

E dynamical variables	E representation expressions[a]	D representation expressions[a]
E position	$\mathbf{x}_\pm = \Lambda_\pm{}^E \mathbf{x} \Lambda_\pm{}^E = \left\{ \mathbf{x} \pm \frac{1}{2ip}\left[\boldsymbol{\alpha} - \frac{(\boldsymbol{\alpha}\cdot\mathbf{p})\mathbf{p}}{p^2} \right] \right\} \Lambda_\pm{}^E$	$\left\{ \mathbf{x} - \frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{2ip^2}\left[\boldsymbol{\alpha} - \frac{m\beta}{E^2}\mathbf{p} \right] \right\} \Lambda_\pm{}^D$
E momentum	$\mathbf{p}_\pm = \Lambda_\pm{}^E \mathbf{p} \Lambda_\pm{}^E = \mathbf{p}\Lambda_\pm{}^E$	$\mathbf{p}\Lambda_\pm{}^D$
E velocity	$d\mathbf{x}_\pm/dt = i[H^E, x_\pm] = \pm \mathbf{p}_\pm/E = \mathbf{p}_\pm/E_\pm$ where $E_\pm = \pm E$	$\pm(\mathbf{p}/E)\Lambda_\pm{}^D$
E orbital angular momentum	$\mathbf{L}_\pm = \Lambda_\pm{}^E \mathbf{x}\times\mathbf{p}\Lambda_\pm{}^E = \left\{ \mathbf{x} \pm \frac{1}{2ip}\boldsymbol{\alpha} \right\} \times \mathbf{p}_\pm$	$\left[\mathbf{x}\times\mathbf{p} + \frac{1}{2}\boldsymbol{\sigma} - \frac{1}{2}\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})\mathbf{p}}{p^2} \right] \Lambda_\pm{}^D$
E spin angular momentum	$\boldsymbol{\sigma}_\pm = \Lambda_\pm{}^E \frac{\boldsymbol{\alpha}\times\boldsymbol{\alpha}}{2i} \Lambda_\pm{}^E = \mp \gamma_{5\pm}\frac{\mathbf{p}_\pm}{p}$	$\frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{p}\frac{\mathbf{p}}{p}\Lambda_\pm{}^D$

[a] The upper (lower) sign has to be taken for positive- (negative-) energy states.

The third difference arises in the following connection If one wishes to keep positive- and negative-energy states separate, the only expectation values of an operator A, which may occur, are

$$\langle\psi_+|A|\psi_+\rangle, \quad \langle\psi_-|A|\psi_-\rangle, \tag{6}$$

where the $+$ $(-)$ sign denotes positive- (negative-) energy states. Since

$$\langle\psi_\pm|A|\psi_\pm\rangle = \langle\psi|\Lambda_\pm A\Lambda_\pm|\psi\rangle = \langle\psi|A_\pm|\psi\rangle, \tag{7}$$

instead of discussing a general operator A and then considering the consequences of the fact that only terms of type (6) do occur, one may alternatively and more simply drop restriction (6) and discuss instead the operator

$$A_\pm = \Lambda_\pm A \Lambda_\pm. \tag{8}$$

However, all the relevant C operators ($\mathbf{x}^C$, $d\mathbf{x}^C/dt$, $\mathbf{p}^C$, $\mathbf{x}^C\times\mathbf{p}^C$, $\boldsymbol{\sigma}^C$) commute with the operator $\Lambda_\pm{}^C$, so that in every relation involving such C operators (commutation relations, etc.), one can simply shift the $\Lambda_\pm{}^C$ on one side of each expression, and the discussion of the $A_\pm$ operators becomes trivially related to the similar relation for the original A operators. This is why the difference between $A_\pm$ and A operators is not even mentioned in the F.-W. paper. However, the significant operators in E representation do not all commute with $\Lambda_\pm{}^E$ and we have to take into account such a distinction and use the $\pm$ operators consistently.

A table of important C and D operators is already available in the F.-W. paper. Therefore, we shall simply give a table of the relevant E operators, with their expressions in D representation. Their C-representation expressions could be easily derived by means of the following canonical transformation, which holds for an arbitrary operator A:

$$A_{C\text{ repres.}} = UA_{E\text{ repres.}}U^{-1}, \tag{9}$$

where

$$U^{\pm 1} = \frac{1}{\sqrt{2}}\left(1 \pm \frac{\beta\boldsymbol{\alpha}\cdot\mathbf{p}}{p}\right). \tag{10}$$

Since nothing new is to be obtained by such a transformation, which is not yet apparent in D representation, the explicit C expressions are not given.

The following comments to Table I are in order:

(1) The E position operator $\mathbf{x}_\pm$ satisfies the usual commutation relation with the E momentum operator $\mathbf{p}_\pm$. However the different components of $\mathbf{x}_\pm$ do not commute. The commutator for the x and y components is the following:

$$[x_\pm, y_\pm] = \left(\frac{ip_z}{2p^3}\gamma_5\Lambda_\pm{}^E\right)_{E\text{ repres.}} = \left(\pm\frac{\sigma_z}{2ip^2}\Lambda_\pm{}^D\right)_{D\text{ repres.}}. \tag{11}$$

(2) The time derivative of the E position operator (the E velocity) is related to the E momentum by the relation to be expected from relatively, as in the case of the homonominous C operators.

(3) Both the E orbital angular momentum and E spin angular momentum are constants of motion. They are orthogonal and parallel to the E momenta, respectively. Their moduli have a simple physical meaning, being simply the *total* angular momentum along the two directions. This is most easily shown by looking to their D-representation expressions. Even the factor $\frac{1}{2}$ ($\mathbf{J} = \mathbf{L} + \frac{1}{2}\boldsymbol{\sigma}$), which creeps in both in C and D representations, is automatically obtained in the E representation. There is no need for adding the quantity $\frac{1}{2}\boldsymbol{\sigma}$ to the orbital angular momentum. Simply by properly defining the orbital angular momentum as $\mathbf{x}_\pm\times\mathbf{p}_\pm$, one finds the *total* angular momentum (still a vector orthogonal to $\mathbf{p}_\pm$), i.e., a quantity which automatically incorporates the relativistic (spin) effects.

(4) The E spin operator, $\boldsymbol{\sigma}_{\pm}$, is also simply related to the E chirality operator $\gamma_{5\pm}$ which in the conventional D representation is in turn simply related to the longitudinal polarization (helicity) of the particle. The relation is not an identity because of the sign $\pm$ (see Table I). This shows most clearly why any chirality invariance requirement, such as has been used in the theory of weak interaction,[3] results in opposite helicities for particles and antiparticles.

The discussion of the E representation for Dirac particles in interaction with external fields will be dealt with in a subsequent communication.

ACKNOWLEDGMENTS

We wish to acknowledge useful discussions with C. J. Goebel, R. E. Marshak, K. T. Mahantappa, and S. Okubo.

[3] F. C. G. Sudarshan and R. E. Marshak, Proceedings of the Padua-Venice Conference on Mesons and Newly Discovered Particles, 1958; Phys. Rev. **109**, 1860 (1958). See also R. P. Feynman and M. Gell-Mann, Phys. Rev. **109**, 193 (1958).

PHYSICAL REVIEW VOLUME 113, NUMBER 6 MARCH 15, 1959

A Soluble Problem in Dispersion Theory*

M. L. GOLDBERGER AND S. B. TREIMAN
Palmer Physical Laboratory, Princeton, University Princeton, New Jersey
(Received November 12, 1958)

The Lee model is modified by addition of a new field θ' and a weak coupling $N+\theta \rightarrow N+\theta'$, which leads to instability of the V particle: $V \rightarrow N+\theta \rightarrow N+\theta'$. The decay amplitude is calculated to lowest order in the weak coupling by dispersion relation methods. In effect we are required to study a set of simultaneous dispersion relations. The problem is completely soluble and serves to clarify the essential structure of dispersion methods. The results agree with what one obtains, more easily in the present case, by direct methods.

I. INTRODUCTION

THE Lee model[1,2] of a soluble field theory has come to play a role similar to that of, say, the harmonic oscillator in classical mechanics. Once a model is known to be soluble by simple and straightforward methods, it is not difficult to find indirect and not-so-simple methods of solution which may nevertheless be relevant and useful in other contexts. In this essentially pedagogical spirit we discuss here the dispersion relation approach to the Lee model. The original model is slightly altered however, by addition of a weak coupling which leads to instability of one of the particles of the theory. This modification provides a physical motivation for studying matrix elements which are uninteresting in the original model and thus, as is desirable, forces us to study a set of simultaneous dispersion relations.

A second reason for enlarging the Lee dynamics in this way has to do with a dispersion relation treatment of $\pi \rightarrow \mu+\nu$ decay which we undertook previously.[3] In the present case we deal again with a decay process, and it is possible to test for errors of principle in the dispersion relation approach. This is worth while, for when applied to particle decay the dispersion methods treat renormalization questions in a way which has disturbed some of our colleagues.[4] What we find in the present model is that the dispersion approach leads to the correct solution. A practical attack on more realistic particle decay problems of course requires many approximations and assumptions beyond a commitment to dispersion relations. But granted the basic analyticity assumptions, it appears that no errors of principle enter into the application of the dispersion relation methods.

The Lee model deals with N, θ, and V fields which are coupled according to the interaction $V \rightleftarrows N+\theta$. The corresponding particles N and θ are stable; and with a suitable choice of parameters a stable V particle also exists.[5] Let the respective masses be m_N, μ, and m_V, where $m_V < m_N+\mu$. We now introduce an additional field θ', corresponding to a particle of mass μ', where $m_N+\mu' < m_V < m_N+\mu$. We also introduce a direct *weak* interaction $N+\theta \rightleftarrows N+\theta'$, which we always treat to lowest order. As a consequence of this interaction the V particle becomes unstable, decaying into $N+\theta'$ through the sequence $V \rightarrow N+\theta \rightarrow N+\theta'$. Our problem is to calculate the decay amplitude—to first order in

* Work supported in part by the Office of Scientific Research, Air Research and Development Command.

[1] T. D. Lee, Phys. Rev. **95**, 1329 (1954).

[2] G. Källen and W. Pauli, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **30**, No. 7 (1955).

[3] M. L. Goldberger and S. B. Treiman, Phys. Rev. **110**, 1178 (1958).

[4] We want to thank especially S. Barshay, N. Kroll, A. Pais, M. Ruderman, and J. C. Taylor for discussions and communications. We also thank R. Haag for informative discussions on the Lee model.

[5] V. Glaser and G. Källen, Nuclear Phys. **2**, 706 (1957).