Zitterbewegung in Relativistic Spin-0 and -$\frac{1}{2}$ Hamiltonian Theories*

Ralph F. Guertin and Eugene Guth†
Physics Department, Rice University, Houston, Texas 77001
(Received 10 April 1972; revised manuscript received 11 October 1972)

By working in the Heisenberg picture, one of the present authors has previously given a general treatment of the charge-space *Zitterbewegung* of the coordinate in theories involving a Hamiltonian that factorizes the Klein-Gordon equation. In the present paper, this treatment is expanded, with particular emphasis being placed on the peculiar doubling of the dimension of the Hamiltonian due to *Zitterbewegung*. The two-component Weyl theory of the neutrino and the four-component Dirac and two-component Sakata-Taketani theories for massive particles of spin $\frac{1}{2}$ and 0, respectively, are discussed from the above viewpoint. The recent two-component theory of Biedenharn, Han, and van Dam (BHV) for massive spin-$\frac{1}{2}$ particles is analyzed. From our viewpoint, a Poincaré-invariant theory of this type necessarily has four components. We offer a two-component interpretation of the BHV theory at the expense of Poincaré invariance. This version cannot be generalized for an arbitrary electromagnetic field, in contrast to the four-component version. We also discuss the relation of the (first order in the time derivative) four-component BHV theory to the (second order) two-component Kramers theory.

I. INTRODUCTION

The question of the dimensionality of equations describing relativistic particles with spin is an interesting one, and the spin-$\frac{1}{2}$ case provides some pertinent observations. To describe a massless spin-$\frac{1}{2}$ particle one commonly employs the two-component first-order Weyl equation, whereas massive spin-$\frac{1}{2}$ particles can be described by the four-component first-order Dirac equation or by the two-component second-order Kramers equation.[1,2] Many authors have been concerned with this apparent distinction between the massive and massless cases,[3] since there appears to be a "doubling" of the spin space or of the time derivative in the former but not in the latter. Recently this question has been taken up in several interesting papers by Biedenharn, Han, and van Dam[4,5] (BHV), who have introduced an equation which is of first order in the time derivative, which describes massive spin-$\frac{1}{2}$ particles, and which they formally consider as a two-component equation. One of the purposes of the present paper is to emphasize that a Poincaré-invariant theory based on the BHV equations is a four-component one. The description of a particle in an arbitrary electromagnetic field also requires four-component BHV equations, and we find that an alternative interpretation of the physical situation is required if a two-component formalism is employed.

The difference in dimension of the Dirac and Weyl equations can be understood by examining the different role that spin plays in massive and massless theories. Let us recall that the four solutions of the free Dirac equation for each momentum value include both signs of the energy and both signs of the helicity, whereas the two solutions of the Weyl equation have opposite signs of both the energy and the helicity. It is well known that an irreducible representation of the proper orthochronous Poincaré group describing either massive or massless particles has a fixed sign of the energy.[6,7] If this representation describes a massive spin-J particle then there are $2J+1$ helicity states for each value of the momentum, but if it describes a massless spin-J particle there is only one helicity,[8] $+J$ or $-J$. Thus, in both the Dirac and Weyl equations there is a doubling of the space by the inclusion of both positive- and negative-energy representations of the proper orthochronous Poincaré group, and the Weyl equation provides no motivation for seeking a first-order two-component massive spin-$\frac{1}{2}$ equation.

By considering the spin-0 case one can illustrate the role that the sign of the energy plays in the doubling, while avoiding any complications due to the presence of spin. A massive spin-0 particle can be described by either the one-component second-order Klein-Gordon equation or by the two-component Sakata-Taketani equation,[9,10] which is of first order in the time derivative. In the latter, the dimensionality accounts for the 2 degrees of freedom in the "charge" space of positive and negative energies, whereas in the former this freedom is accounted for by the second-order time derivative. Feshbach and Villars[10] have shown the equivalence of the two theories in the presence of an arbitrary electromagnetic field. Brown[2] has stressed that there is a similar relationship between the Dirac equation and the

Kramers equation, pointing out that the latter "separates the spin and charge degrees of freedom, describing the former by the familiar Pauli spin matrices and the latter by the space-time dependence of the wave function."

In this paper we shall continue the earlier development by one of the present authors of a unified Hamiltonian theory of relativistic free-particle equations.[11] By working in the Heisenberg picture, we are able to show the relationship of the doubling of the spin space that is present in a large class of such Hamiltonians to rapid oscillations or "*Zitterbewegung*" of the canonical coordinate between states of positive and negative energies. The free Weyl, Dirac, Sakata-Taketani spin-0, and BHV Hamiltonians provide particular examples of such theories. Our approach serves to illustrate that the origin of the doubling in the BHV theory is the same as in the other theories.

We begin in Sec. II by confining our attention to theories that provide a representation of the proper orthochronous Poincaré group and define what we mean by a Hamiltonian that factorizes the Klein-Gordon operator. In addition to the usual canonical quantization rules in the Heisenberg picture, we assume that the angular momentum operator is expressible as the sum of a spin-J and an orbital part. In such theories the Hamiltonian, the canonical momentum, and the canonical helicity are mutually commuting – we find that the matrix elements of the canonical coordinate between simultaneous eigenstates of all three exhibit rapid oscillations between states having the same values of the canonical momentum but opposite signs of the energy (there is a certain exception that does not include the equations with which we are concerned). The presence of both positive- and negative-energy states is then seen to require that the Hamiltonian for a massive particle be a $2(2J+1)$ by $2(2J+1)$ matrix operator. Further observations concerning the nature of the *Zitterbewegung* are made for theories that are invariant under spatial inversion. To specifically illustrate our results we then consider the Dirac and the Sakata-Taketani spin-0 Hamiltonians. By examining their Heisenberg equations of motion, we find that in the Sakata-Taketani theory a set of operators called the τ matrices plays a role similar to that of the familiar ρ matrices in the Dirac theory. Finally, we discuss the extension of our approach to the Weyl equation.

In Sec. III we discuss the Poincaré-invariant theory based on the BHV equations and show why it is a four-component one. Although two-component operators H_+ and H_- that separately factorize the Klein-Gordon operator exist, a four-component Hamiltonian formed from them is required to obtain a Poincaré-invariant theory. This doubling of the spin space is explained by showing that the theory satisfies the general requirements of Sec. II; in particular, the *Zitterbewegung* of the canonical coordinate connects the subspaces associated with H_+ and H_-. After finding the operators that play the same role in the BHV theory that the ρ matrices do in the Dirac theory, we discuss the equivalence of these two free-particle theories, employing a nonlocal unitary operator given by BHV.[5]

A possible interpretation of the two-component BHV equations in a vanishing electromagnetic field, a theory which is not Poincaré-invariant, is provided in Sec. IV.[12] It appears that one obtains a genuine splitting into two two-component subspaces, each possessing its own Hamiltonian, by adding to the free Dirac Hamiltonian a particular potential term.

In Sec. V the electromagnetic four-potential is introduced into the (four-component) free BHV equations in the usual gauge-invariant manner, and we once more obtain a theory equivalent to that of Dirac. By treating the even and odd parts (under spatial inversion) of the potentials as the components of a matrix in the "stigma" space of BHV, we rewrite our equations in a form that BHV formally regard as two-component. To emphasize the four-component nature of the BHV theory we consider the limit of a vanishing field and also the special case of electromagnetic potentials that yield an "even environment." After going to the Schrödinger picture, we review the relationship of the BHV equations to the Kramers equation for a spin-$\frac{1}{2}$ particle in an electromagnetic field. Finally, we discuss the impossibility of introducing the electromagnetic potentials into the two-component interpretation of Sec. IV.

Our treatment is confined entirely to the one-particle interpretation of the various equations and no attempt is made to consider second quantization. The possibility that the BHV theory might lead to different results than the Dirac theory in certain cases is reviewed in Sec. VI.

II. FACTORIZATIONS OF THE KLEIN-GORDON EQUATION FOR FREE RELATIVISTIC PARTICLES AND ZITTERBEWEGUNG

Let us consider a relativistic free-particle theory in which the generator of time displacements (Hamiltonian) is H_c and the generator of spatial displacements (canonical momentum) is $\vec{p}_c$. In addition, let $\vec{J}_c$ be the generator of rotations and $\vec{K}_c$ the generator of "boosts." These operators are required to satisfy the Lie algebra of the proper orthochronous Poincaré group[7,13]:

$$[p_{ci}, p_{cj}] = 0, \tag{2.1a}$$

$$[p_{ci}, H_c] = 0, \tag{2.1b}$$

$$[J_{ci}, p_{cj}] = i\sum_k \epsilon_{ijk} p_{ck}, \tag{2.1c}$$

$$[J_{ci}, H_c] = 0, \tag{2.1d}$$

$$[J_{ci}, J_{cj}] = i\sum_k \epsilon_{ijk} J_{ck}, \tag{2.1e}$$

$$[p_{ci}, K_{cj}] = -i\delta_{ij} H_c, \tag{2.1f}$$

$$[H_c, K_{ci}] = -ip_{ci}, \tag{2.1g}$$

$$[J_{ci}, K_{cj}] = i\sum_k \epsilon_{ijk} K_{ck}, \tag{2.1h}$$

$$[K_{ci}, K_{cj}] = -i\sum_k \epsilon_{ijk} J_{ck}. \tag{2.1i}$$

As a result of these equations, the transformation properties of the Hamiltonian under finite proper orthochronous Poincaré transformations are expressed by the relations

$$\exp(i\vec{d}\cdot\vec{p}_c)H_c\exp(-i\vec{d}\cdot\vec{p}_c) = H_c, \tag{2.2a}$$

$$\exp(-iDH_c)H_c\exp(iDH_c) = H_c, \tag{2.2b}$$

$$\exp(-i\vec{\theta}\cdot\vec{J}_c)H_c\exp(i\vec{\theta}\cdot\vec{J}_c) = H_c, \tag{2.2c}$$

$$\exp(-i\vec{\lambda}\cdot\vec{K}_c)H_c\exp(i\vec{\lambda}\cdot\vec{K}_c) = H_c\cosh|\vec{\lambda}| + \frac{\vec{\lambda}\cdot\vec{p}_c}{|\vec{\lambda}|}\sinh|\vec{\lambda}|. \tag{2.2d}$$

Here the components of the displacement parameters $\vec{d}$ and D are numbers that commute with all the operators of the theory. Likewise, the components of $\vec{\theta}$, which indicate both the magnitude of the angle and the direction of the axis of rotation, and the components of $\vec{\lambda}$, which indicate both the magnitude and direction of the boost, are numbers that commute with each of the operators. The parameters also commute with the discrete symmetry operators when the latter are included.

Since $H_c^2 - \vec{p}_c^{\,2}$ commutes with each of the generators in (2.1), we can take it to be a positive multiple of the unit matrix. Our actual concern will be only with cases for which $\vec{p}_c^{\,2} = \vec{p}^{\,2}$, where $\vec{p}$ is the familiar operator that takes on the form $-i\vec{\nabla}$ in a coordinate representation. Then

$$H_c^2 = \vec{p}^{\,2} + m^2, \tag{2.3}$$

where $m > 0$. A Hamiltonian satisfying (2.3) is said to "factorize the Klein-Gordon equation."

We assume the usual canonical commutation relations for $\vec{p}_c$ and the canonical coordinate $\vec{x}_c$:

$$[x_{ci}, x_{cj}] = 0, \tag{2.4a}$$

$$[p_{ci}, p_{cj}] = 0, \tag{2.4b}$$

$$[x_{ci}, p_{cj}] = i\delta_{ij}. \tag{2.4c}$$

To avoid any ambiguity, we shall take $\vec{x}_c$ to be $\vec{x}$ when $\vec{p}_c = \vec{p}$; in other cases we require that $\vec{x}_c^{\,2} = \vec{x}^{\,2}$.

We confine our attention to representations in which the generator of rotations has the form

$$\begin{aligned}\vec{J}_c &= \vec{x}_c\times\vec{p}_c + \vec{S}^{(J)}\\ &= \vec{x}\times\vec{p} + \vec{S}^{(J)}.\end{aligned} \tag{2.5}$$

Here the $\vec{S}^{(J)}$ are the usual $2J+1$ by $2J+1$ matrices satisfying the algebra

$$[S_i^{(J)}, S_j^{(J)}] = i\sum_k \epsilon_{ijk} S_k^{(J)}. \tag{2.6}$$

This assures that the theory describes massive particles of spin J. From (2.1b), (2.1d), and (2.5) we see that in general none of the three components of $\vec{S}^{(J)}$ commutes with H_c but that

$$S^{(J)}(\vec{p}_c) = \vec{S}^{(J)}\cdot\frac{\vec{p}_c}{|\vec{p}|}, \tag{2.7}$$

which we call the canonical helicity operator, does. Therefore, we can find simultaneous eigenstates of H_c, $\vec{p}_c$, and $S^{(J)}(\vec{p}_c)$, the latter of which has $2J+1$ eigenvalues.

As usual, the Hamiltonian is to be expressed as a function of the canonical momentum. Aside from the infinite dimensionality associated with the operator $\vec{p}_c$, what is the "dimension" of H_c? It is clear that (2.5) and the commutation relations (2.1) require H_c to be at least a $2J+1$ by $2J+1$ matrix operator having the spin degrees of freedom, but are there any compelling reasons for the space of this operator to be even larger?

In seeking the answer to our question, let us follow the advice of Dirac,[14] "to get a physical understanding of any quantum theory, it is best to use the Heisenberg equations of motion. . . ." Therefore, we assume the Heisenberg equations of motion for any dynamical variable $\mathcal{O}_c$:

$$\dot{\mathcal{O}}_c = i[H_c, \mathcal{O}_c] + \frac{\partial\mathcal{O}_c}{\partial t}. \tag{2.8}$$

The consistency of (2.1b) with (2.8) requires that $\dot{\vec{p}}_c = 0$, as should be the case for a free particle.

One of the authors[11] has shown elsewhere that, as a consequence of (2.1b), (2.3), (2.4), and (2.8), we have

$$\begin{aligned}i\ddot{\vec{x}}_c &= -2H_c\dot{\vec{x}}_c + 2\vec{p}_c\\ &= 2\dot{\vec{x}}_c H_c - 2\vec{p}_c.\end{aligned} \tag{2.9}$$

With the aid of this result, the solution of the equation of motion for $\vec{x}_c$ is found to be[11]

$$\begin{aligned}\vec{x}_c(t) = {}&\tfrac{1}{4}\ddot{\vec{x}}_c(0)[1-\exp(-i2H_ct)]H_c^{-2}\\ &+ \vec{p}_c H_c^{-1}t + \vec{x}_c(0).\end{aligned} \tag{2.10}$$

We also find that

$$[\dot{x}_{ci}, p_{cj}] = 0\,. \tag{2.11}$$

Then from (2.9) and (2.11) we have

$$\ddot{x}_{ci}H_c = -H_c\ddot{x}_{ci}\,, \tag{2.12a}$$

$$\ddot{x}_{ci}p_{cj} = p_{cj}\ddot{x}_{ci}\,. \tag{2.12b}$$

Only in the special case $\dot{\vec{x}}_c = H_c^{-1}\vec{p}_c$ [e.g., if $H_c = (\vec{p}^2 + m^2)^{1/2}$] does $\ddot{\vec{x}}_c$ vanish identically and therefore make no contribution to (2.10). In other cases one finds from (2.12) that the only nonvanishing matrix elements of $\ddot{\vec{x}}_c$ between simultaneous eigenstates of $\vec{p}_c$ and H_c are those that have the same canonical three-momentum and opposite signs of the energy. Thus, for each value of the canonical three-momentum and each of the $2J+1$ values of the canonical helicity, both positive- and negative-energy solutions occur and there is a "doubling" of the space of the Hamiltonian to $2(2J+1)$ dimensions. Since the eigenvalues of H_c have absolute values greater than or equal to m, we find from (2.10) that $\vec{x}_c$ exhibits rapid oscillations of frequency $\omega \geq 2m$ between states of positive and negative energy; following Schrödinger[15] we refer to this phenomenon as "*Zitterbewegung*" in the "charge space" of positive- and negative-energy states.

It is interesting to note that the spin operator $\vec{S}^{(J)}$ also exhibits rapid oscillations in charge space. From (2.1d), (2.5), and (2.8),

$$\ddot{\vec{S}}^{(J)} = -\ddot{\vec{x}}_c \times \vec{p}_c\,, \tag{2.13}$$

with the solution

$$\vec{S}^{(J)}(t) = \tfrac{1}{4}\ddot{\vec{S}}^{(J)}(0)[1 - \exp(-i2H_c t)]H_c^{-2} + \vec{S}^{(J)}(0)\,. \tag{2.14}$$

It follows from (2.12) and (2.13) that

$$\ddot{\vec{S}}^{(J)}H_c = -H_c\ddot{\vec{S}}^{(J)}\,, \tag{2.15}$$

so the states connected by the "*Spinbewegung*" indeed have the same four-momentum and opposite signs of the energy.

Some further interesting observations concerning the *Zitterbewegung* can be made if the $2(2J+1)$-component operators of our theory are invariant under spatial inversion. In this case, there exists an operator $\mathcal{S}_c$ satisfying[13,16]

$$\mathcal{S}_c\vec{x}_c\mathcal{S}_c = -\vec{x}_c\,, \tag{2.16a}$$

$$\mathcal{S}_c\vec{p}_c\mathcal{S}_c = -\vec{p}_c\,, \tag{2.16b}$$

$$\mathcal{S}_c H_c\mathcal{S}_c = H_c\,, \tag{2.16c}$$

$$\mathcal{S}_c\vec{J}_c\mathcal{S}_c = \vec{J}_c\,, \tag{2.16d}$$

$$\mathcal{S}_c\vec{K}_c\mathcal{S}_c = -\vec{K}_c\,, \tag{2.16e}$$

$$\mathcal{S}_c^2 = 1. \tag{2.16f}$$

An eigenstate of $\mathcal{S}_c$ will be said to have "relativistic parity" +1 or −1 corresponding to its eigenvalue. Since $\mathcal{S}_c$ commutes with H_c and $\vec{J}_c$, relativistic parity is invariant under time translations and rotations in three-space. It is not, however, invariant under spatial translations or boosts.

One can find simultaneous eigenstates of H_c, $\vec{J}_c^2$, J_{c3}, and $\mathcal{S}_c$. From (2.5), (2.8), (2.9), and (2.16a)–(2.16c) it follows that

$$\mathcal{S}_c\ddot{\vec{x}}_c = -\ddot{\vec{x}}_c\mathcal{S}_c\,, \tag{2.17a}$$

$$\ddot{\vec{x}}_c\vec{J}_c^2 = \vec{J}_c^2\ddot{\vec{x}}_c\,. \tag{2.17b}$$

Thus, when the theory is invariant under spatial inversion, the only nonvanishing matrix elements of $\ddot{\vec{x}}_c$ between simultaneous eigenstates of $\vec{J}_c^2$, H_c, and $\mathcal{S}_c$ are those between states having the same total angular momentum and the same absolute value of the energy but opposite signs of both the energy and the relativistic parity. This provides another way of viewing the rapid oscillations of the coordinate in (2.10).

The simplest form that a Hamiltonian meeting our requirements can have is[11]

$$H_c = b\dot{\vec{x}}_c \cdot \vec{p}_c + g\,, \tag{2.18}$$

where $\dot{\vec{x}}_c$ is either independent of $\vec{p}_c$ or is linear in $\vec{p}_c$. Here b is a numerical factor having the value 1 in the former case and $\frac{1}{2}$ in the latter case. Although it commutes with $\vec{x}_c$ and $\vec{p}_c$, g need not commute with $\dot{\vec{x}}_c$.

Upon substituting $\vec{p}_c = \vec{p}$, $b\dot{\vec{x}}_c = \rho_1\vec{\sigma}$, and $g = \rho_3 m$ into (2.18), one obtains the Dirac Hamiltonian

$$H_D = \rho_1\vec{\sigma}\cdot\vec{p} + \rho_3 m\,. \tag{2.19}$$

As is well known, H_D, $\vec{p}_D = \vec{p}$, and the operators

$$\vec{J}_D = \vec{x}\times\vec{p} + \tfrac{1}{2}\vec{\sigma}\,, \tag{2.20a}$$

$$\begin{aligned}\vec{K}_D &= \vec{x}H_D - t\vec{p} - i\rho_1\tfrac{1}{2}\vec{\sigma}\\ &= \tfrac{1}{2}(\vec{x}H_D + H_D\vec{x}) - t\vec{p}\end{aligned} \tag{2.20b}$$

satisfy the Poincaré algebra (2.1) and describe spin-$\frac{1}{2}$ particles.[13]

On the other hand, in (2.18) we can make the identification $\vec{p}_c = \vec{p}$, $b\dot{\vec{x}}_c = (2m)^{-1}\tau_+\vec{p}$, and $g = \tau_3 m$, where $\tau_\pm = \tau_3 \pm i\tau_2$. Here τ_1, τ_2, and τ_3 have formally the same multiplicative properties as the Pauli matrices and may be represented by them at $t=0$. The result is the Sakata-Taketani Hamiltonian[9,10]

$$H_S = \tau_+\frac{\vec{p}^2}{2m} + \tau_3 m \tag{2.21}$$

describing spin-0 particles. In this case the Poincaré algebra (2.1) is satisfied by H_S, $\vec{p}_S = \vec{p}$, and the operators

$$\vec{J}_S = \vec{x}\times\vec{p}\,, \tag{2.22a}$$

$$\vec{K}_S = \tfrac{1}{2}(\vec{x}H_S + H_S\vec{x}) - t\vec{p}\,. \tag{2.22b}$$

All observables $\mathcal{O}_s$ in the present theory are pseudo-Hermitian, satisfying

$$\mathcal{O}_s = \tau_3 \mathcal{O}_s^\dagger \tau_3\,. \tag{2.23}$$

We see that H_D is a result of choosing the Hamiltonian to be linear in the canonical momentum, whereas H_S is a consequence of choosing it to be of second order in this operator – in both Hamiltonians the remaining operators are required to be independent of space-time. One easily verifies that $\ddot{\vec{x}} \neq 0$ in each case; i.e., for the Dirac theory

$$\ddot{\vec{x}}_D = -2\vec{\sigma}\times\vec{p} - 2m\rho_2\vec{\sigma}\,, \tag{2.24a}$$

and for the Sakata-Taketani theory

$$\ddot{\vec{x}}_S = i2\tau_1\vec{p}\,. \tag{2.24b}$$

Therefore, the charge-space *Zitterbewegung* of the coordinate is present and explains why the dimension of the Hamiltonian in each case is twice that required by the spin degrees of freedom. Since, as we shall see below, both theories are invariant under spatial inversion, the remarks following (2.17) concerning relativistic parity also apply to the *Zitterbewegung*.

To verify spatial inversion invariance we introduce the familiar operator $\mathcal{P}$ that inverts $\vec{x}$ and $\vec{p}$:

$$\mathcal{P}\vec{x}\mathcal{P} = -\vec{x}\,, \tag{2.25a}$$

$$\mathcal{P}\vec{p}\mathcal{P} = -\vec{p}\,, \tag{2.25b}$$

but which commutes with σ_i, ρ_i, and τ_i and with the parameters $\vec{d}$, D, $\vec{\theta}$, and $\vec{\lambda}$ in (2.2). The spatial inversion operators consistent with Eqs. (2.16) are

$$\mathcal{S}_D = \rho_3\mathcal{P} \tag{2.26a}$$

for the Dirac theory and

$$\mathcal{S}_S = \mathcal{P} \tag{2.26b}$$

for the Sakata-Taketani theory.

It is instructive to demonstrate the similar role that the ρ and τ matrices play in the Dirac and the Sakata-Taketani theories by examining their Heisenberg equations of motion.

We find that

$$\dot{\tau}_1 = -i2\tau_1 H_S\,, \tag{2.27a}$$

$$\dot{\tau}_+ = i2m\tau_1\,, \tag{2.27b}$$

$$\dot{\tau}_- = -i2m^{-1}H_S^2\tau_1\,. \tag{2.27c}$$

The solutions of Eqs. (2.27) are

$$\tau_1(t) = \tau_1(0)\exp(-i2H_S t)\,, \tag{2.28a}$$

$$\tau_+(t) = \tau_+(0) + \tau_1(0)mH_S^{-1}[1-\exp(-i2H_S t)]\,, \tag{2.28b}$$

$$\tau_-(t) = \tau_-(0) - \tau_1(0)m^{-1}H_S[1-\exp(-i2H_S t)]\,. \tag{2.28c}$$

Similarly we obtain

$$\dot{\rho}_1 = -2m\rho_2\,, \tag{2.29a}$$

$$\dot{\rho}_2 = -i2\rho_2 H_D\,, \tag{2.29b}$$

$$\dot{\rho}_3 = 2\vec{\sigma}\cdot\vec{p}\rho_2\,, \tag{2.29c}$$

where $\vec{\sigma}\cdot\vec{p}$ is a constant of the motion. Equations (2.29) have the solutions

$$\rho_1(t) = \rho_1(0) + i\rho_2(0)mH_D^{-1}[1-\exp(-i2H_D t)]\,, \tag{2.30a}$$

$$\rho_2(t) = \rho_2(0)\exp(-i2H_D t)\,, \tag{2.30b}$$

$$\rho_3(t) = \rho_1(0) - i\rho_2(0)\vec{\sigma}\cdot\vec{p}H_D^{-1}[1-\exp(-i2H_D t)]\,. \tag{2.30c}$$

One immediately notes the striking similarity of Eqs. (2.28) and (2.30) and the fact that both exhibit *Zitterbewegung*. To investigate the origin of this phenomenon in the present case, we resort to the equalities

$$\tau_1 H_S = -H_S\tau_1 \tag{2.31a}$$

and

$$\rho_2 H_D = -H_D\rho_2\,. \tag{2.31b}$$

Equations (2.31) and the fact that both τ_1 and ρ_2 commute with $\vec{p}$ show that the rapid oscillations in (2.28) and (2.30) occur between states having the same three-momentum but opposite signs of the energy.

In the general discussion at the beginning of this section, only Eqs. (2.1a), (2.1b), (2.3), (2.4), and (2.8) were necessary to obtain Eqs. (2.10) and (2.12) demonstrating the coordinate *Zitterbewegung*. Our result is therefore applicable to the case of zero mass, except that now the oscillations occur for all frequencies $\omega > 0$. The remaining assumptions (2.1c) through (2.1i) and (2.5) enabled us to be precise as to what was doubled by the presence of both signs of the energy in the massive case; i.e., the $2J+1$ spin components. However, if (2.5) is assumed for massless particles, then, except for $J=\frac{1}{2}$, subsidiary conditions are required to eliminate all helicities except $+J$ and $-J$.

For spin $\frac{1}{2}$ we have the Weyl Hamiltonian

$$H_W = \vec{\sigma}\cdot\vec{p}\,, \tag{2.32}$$

in which case the helicity operator (2.7) takes on the form

$$S^{(1/2)}(\vec{p}) = \tfrac{1}{2}\frac{H_W}{|H_W|}\,. \tag{2.33}$$

Thus the *Zitterbewegung* is between states having opposite signs of both the energy and of the helicity, in agreement with our remarks in the Introduction.

III. FOUR-COMPONENT POINCARÉ-INVARIANT BHV EQUATIONS

One form of the equations introduced by Biedenharn, Han, and van Dam[4,5] for a vanishing external electromagnetic field employs the Hamiltonian

$$H_B = \rho_1 \vec{\sigma} \cdot \vec{p} + \mathcal{P}_{\text{ext}} m . \tag{3.1}$$

This is the Hamiltonian form of Eqs. (56) and (56′) of Ref. 5 and is identical to Eq. (57) there.

In (3.1) we have

$$\mathcal{P}_{\text{ext}} = \mathcal{P} \varsigma_3 = \varsigma_3 \mathcal{P} . \tag{3.2}$$

Here $\mathcal{P}$, introduced in Eq. (2.25), is referred to by BHV as the "internal parity" operator. To obtain their "external parity" operator $\mathcal{P}_{\text{ext}}$ BHV introduce the operators ς_j, where ς is the ancient Greek symbol "stigma." As explained by BHV,[17] the ς matrices have formally the same multiplicative properties as the ρ matrices, but they commute with the ρ matrices. In addition ς_j commutes with $\vec{\sigma}$, $\vec{p}$, $\vec{x}$, and $\mathcal{P}$. It follows that

$$\mathcal{P}_{\text{ext}} \vec{p} \mathcal{P}_{\text{ext}} = -\vec{p} , \tag{3.3a}$$

$$\mathcal{P}_{\text{ext}} \vec{x} \mathcal{P}_{\text{ext}} = -\vec{x} , \tag{3.3b}$$

$$\mathcal{P}_{\text{ext}} \vec{\sigma} \mathcal{P}_{\text{ext}} = \vec{\sigma} , \tag{3.3c}$$

$$\mathcal{P}_{\text{ext}} \rho_j \mathcal{P}_{\text{ext}} = \rho_j . \tag{3.3d}$$

Furthermore, from (3.1) and (3.2), we see that

$$[H_B , \rho_1] = [H_B , \varsigma_3] = 0, \tag{3.4}$$

so ρ_1 and ς_3 can be taken to be constants of the motion. Therefore, ρ_1 plays a different role here than in the Dirac theory, where it is not a constant of the motion, and might be more appropriately labeled ρ_1^B in the present theory.

Following BHV an eigenstate of ρ_1 is said to have "chirality" +1 or −1 corresponding to its eigenvalue. Likewise, an eigenstate of ς_3 has "stigma" designated by its eigenvalue[18] +1 or −1. In the BHV terminology, one can choose chirality and stigma to be "sharp." It follows that the space of the Hamiltonian (3.1) can be divided into four subspaces characterized by their chirality and their stigma. In particular, we may rewrite (3.1) in the form

$$H_B = \tfrac{1}{2}(1 + \rho_1) H_B^{\text{I}} + \tfrac{1}{2}(1 - \rho_1) H_B^{\text{II}} , \tag{3.5}$$

where

$$H_B^{\text{I}} = \vec{\sigma} \cdot \vec{p} + \mathcal{P}_{\text{ext}} m , \tag{3.6a}$$

$$H_B^{\text{II}} = -\vec{\sigma} \cdot \vec{p} + \mathcal{P}_{\text{ext}} m . \tag{3.6b}$$

The advantage of using the form (3.5) is that the operator $\frac{1}{2}(1 + \rho_1) H_B^{\text{I}}$ has nonvanishing matrix elements only between states of positive chirality and the operator $\frac{1}{2}(1 - \rho_1) H_B^{\text{II}}$ has nonvanishing matrix elements only between states of negative chirality.

The equivalence relation

$$H_B^{\text{I}} = \mathcal{P}_{\text{ext}} H_B^{\text{II}} \mathcal{P}_{\text{ext}} , \tag{3.7}$$

which follows from (3.3) and (3.6), shows that the solutions of H_B^{II} can be obtained from those of H_B^{I} and vice versa. Thus, Eq. (3.1) contains redundant components and either of the operators (3.6) describes the same physical situation as H_B.

We can continue the decomposition into subspaces by using (3.2) to write

$$H_B^{\text{I}} = \tfrac{1}{2}(1 + \varsigma_3) H_+^{\text{I}} + \tfrac{1}{2}(1 - \varsigma_3) H_-^{\text{I}} , \tag{3.8a}$$

$$H_B^{\text{II}} = \tfrac{1}{2}(1 + \varsigma_3) H_+^{\text{II}} + \tfrac{1}{2}(1 - \varsigma_3) H_-^{\text{II}} , \tag{3.8b}$$

where

$$H_\pm^{\text{I}} = \vec{\sigma} \cdot \vec{p} \pm m \mathcal{P} . \tag{3.9a}$$

$$H_\pm^{\text{II}} = -\vec{\sigma} \cdot \vec{p} \pm m \mathcal{P} . \tag{3.9b}$$

The equivalent operators $\frac{1}{2}(1 + \varsigma_3) H_+^{\text{I}}$ and $\frac{1}{2}(1 + \varsigma_3) H_+^{\text{II}}$ have nonvanishing matrix elements only between states of positive stigma, whereas the equivalent operators $\frac{1}{2}(1 - \varsigma_3) H_-^{\text{I}}$ and $\frac{1}{2}(1 - \varsigma_3) H_-^{\text{II}}$ have nonvanishing matrix elements only between states of negative stigma.

From (3.9) and (2.25b),

$$(H_\pm^{\text{I}})^2 = (H_\pm^{\text{II}})^2 = \vec{p}^{\,2} + m^2 . \tag{3.10}$$

Thus each of the four operators in (3.9) factorizes the Klein-Gordon equation,[4,5] but is each of these operators the Hamiltonian for a Poincaré-invariant theory? In the remainder of this section we shall develop the Poincaré-invariant BHV theory according to the general treatment at the beginning of Sec. II and explain why only the equivalent four-component operators H_B^{I} and H_B^{II} are Hamiltonians. An alternative interpretation that allows one to regard each of the operators is (3.9) as Hamiltonians for a non-Poincaré-invariant theory will be discussed in the next section.

We shall confine our attention to H_B^{I} in (3.6a), since the operators for the representation using H_B^{II} can be obtained through the equivalence transformation (3.7). From (3.1)–(3.6) and (2.25b), we see that $[H_B , \vec{p}] \neq 0$ and also $[H_B^{\text{I}} , \vec{p}] \neq 0$ and $[H_B^{\text{II}} , \vec{p}] \neq 0$, so (2.1b) demonstrates that if any of these operators is indeed a free-particle Hamiltonian, $\vec{p}$ cannot be the canonical momentum operator; i.e., the generator of spatial displacements. However, let us introduce the operators

$$\vec{p}_B = \varsigma_1 \vec{p} , \tag{3.11a}$$

$$\vec{x}_B = \varsigma_1 \vec{x} \tag{3.11b}$$

[note that $\vec{p}_B$ and $\vec{x}_B$ satisfy the canonical commutation relations (2.4)], and rewrite (3.6a) in the form

$$H_B^{\mathrm{I}} = \varsigma_1 \vec{\sigma}\cdot\vec{p}_B + \mathcal{P}_{\mathrm{ext}} m . \tag{3.12}$$

Furthermore, we define

$$\vec{J}_B = \vec{x}_B \times \vec{p}_B + \tfrac{1}{2}\vec{\sigma} = \vec{J}_D , \tag{3.13a}$$

$$\vec{K}_B = \vec{x}_B H_B^{\mathrm{I}} - t\vec{p}_B - i\varsigma_1 \tfrac{1}{2}\vec{\sigma} = \tfrac{1}{2}(\vec{x}_B H_B^{\mathrm{I}} + H_B^{\mathrm{I}}\vec{x}_B) - t\vec{p}_B . \tag{3.13b}$$

It is easily verified that H_B^{I}, $\vec{p}_B$, $\vec{J}_B$, and $\vec{K}_B$ satisfy the Lie algebra (2.1). Thus H_B^{I} is indeed the Hamiltonian for a spin-$\frac{1}{2}$ theory that is invariant under proper orthochronous Poincaré transformations.

The theory is also invariant under spatial inversions, since if we make the identification

$$\mathcal{S}_B = \varsigma_3 , \tag{3.14}$$

the requirements (2.16) are satisfied. Consequently, the stigma of a particle is identical to the relativistic parity quantum number defined in our discussion following (2.16).[19] It is, as we saw there, invariant under time translations and spatial rotations.

Even though the four-component Hamiltonian can be split in every reference frame into two subspaces of sharp stigma, neither H_+^{I} nor H_-^{I} plays the role of a Hamiltonian in the present theory: (a) A Hamiltonian generates time displacements for *all* physical observables, but neither of these operators can generate the equations for $\vec{x}_B$ and $\vec{p}_B$, which have components connecting the two stigma subspaces; (b) furthermore, it is not possible to find a set of two-component operators satisfying the Poincaré algebra (2.1) with either H_+^{I} or H_-^{I} as the Hamiltonian.

Our conclusion that only H_B^{I} yields a Poincaré-invariant theory is in agreement with Ref. 5, but, since the two-component subspaces, H_+^{I} and H_-^{I} of H_B^{I} have opposite stigma (relativistic parity) BHV state,[20] "we have chosen to ignore the dimensionality associated with the space-time structure carried by parity," and they formally regard H_B^{I} as a two-component operator. To emphasize their viewpoint, they choose to write in (2.2)

$$\vec{d}\cdot\vec{p}_B = \varsigma_1 \vec{d}\cdot\vec{p} = \vec{d}_B\cdot\vec{p} , \tag{3.15a}$$

$$\vec{\lambda}\cdot\vec{K}_B = \varsigma_1\vec{\lambda}\cdot\vec{K}'_B = \vec{\lambda}_B\cdot\vec{K}'_B , \tag{3.15b}$$

where

$$\vec{K}'_B = \varsigma_1 \vec{K}_B , \tag{3.16}$$

and where

$$\vec{d}_B = \varsigma_1 \vec{d} , \tag{3.17a}$$

$$\vec{\lambda}_B = \varsigma_1 \vec{\lambda} \tag{3.17b}$$

are "parity-carrying operators[21]" (in the BHV terminology) that anticommute with $\mathcal{P}_{\mathrm{ext}}$. Then in place of (2.2a) and (2.2d) for finite spatial translations and boosts, respectively, they write

$$\exp(i\vec{d}_B\cdot\vec{p}) H_B^{\mathrm{I}} \exp(-i\vec{d}_B\cdot\vec{p}) = H_B^{\mathrm{I}} , \tag{3.18a}$$

$$\exp(-i\vec{\lambda}_B\cdot\vec{K}'_B) H_B^{\mathrm{I}} \exp(i\vec{\lambda}_B\cdot\vec{K}'_B) = H_B^{\mathrm{I}} \cosh|\vec{\lambda}| + \frac{\vec{\lambda}_B\cdot\vec{p}}{|\vec{\lambda}|}\sinh|\vec{\lambda}| . \tag{3.18b}$$

The operators $\vec{p}$ and $\vec{K}'_B$, the former of which does not have the significance of momentum in the present theory, have matrix elements only within each stigma subspace, and BHV formally ignore the doubling present in $\vec{d}_B$ and $\vec{\lambda}_B$.

We shall now demonstrate that, from the viewpoint advocated in this paper, the doubling present in the BHV free-particle theory is no more ignorable than is the doubling in the Dirac and Sakata-Taketani theories. To show that the doubling of the spin space in the BHV theory is associated with the presence of both positive and negative eigenstates of the Hamiltonian for each value of the canonical three-momentum and each value of the canonical helicity, it is sufficient, according to the discussion following (2.12), to verify that the operator $\vec{x}_B$ exhibits *Zitterbewegung* in the charge space; i.e., that $\ddot{\vec{x}}_B \neq 0$ in (2.10). From (3.3b), (3.11), (3.12), and (2.8) it follows that

$$\dot{\vec{x}}_B = \varsigma_1 \vec{\sigma} \tag{3.19}$$

and that

$$\ddot{\vec{x}}_B = -2\vec{\sigma}\times\vec{p}_B - 2m\varsigma_2\mathcal{P}\vec{\sigma} . \tag{3.20}$$

We note that $\ddot{\vec{x}}_B \neq 0$ and it anticommutes with ς_3, so the *Zitterbewegung* of the canonical coordinate does not occur within each two-component subspace having definite stigma [this agrees with our remarks following (2.17)].

Examination of (3.12) shows that H_B^{I} has the simple form (2.18) with $\dot{\vec{x}}_c$ independent of $\vec{p}_c$ and with g a nonlocal operator. Furthermore, we see that H_B^{I} can be obtained from the Dirac Hamiltonian (2.19) by the substitutions $\vec{p}\to\vec{p}_B$, $\rho_1\to\varsigma_1$, and $\rho_3\to\mathcal{P}_{\mathrm{ext}}$. This analogy can be carried further if we solve the equation of motion for the operators ς_1, $\overline{\mathcal{P}} = \varsigma_2\mathcal{P}$, and $\mathcal{P}_{\mathrm{ext}} = \varsigma_3\mathcal{P}$, using H_B^{I}. The final result is obtained from that for the ρ matrices in (2.30) by the substitutions $\rho_1\to\varsigma_1$, $\rho_2\to\overline{\mathcal{P}}$, $\rho_3\to\mathcal{P}_{\mathrm{ext}}$, $H_D\to H_B^{\mathrm{I}}$, and $\vec{\sigma}\cdot\vec{p}\to\vec{\sigma}\cdot\vec{p}_B$.

It has been shown by BHV[5] that the above relation between the free Dirac and BHV theories is in fact an equivalence. They introduced the unitary operator[22]

$$\Gamma = \Gamma^\dagger = \tfrac{1}{2}(1+\varsigma_1) + \tfrac{1}{2}(1-\varsigma_1)\mathcal{P} \qquad (3.21)$$

and considered the transformation

$$\mathcal{O}_B \to \mathcal{O}_B^\Gamma = \Gamma \mathcal{O}_B \Gamma \qquad (3.22)$$

for each operator $\mathcal{O}_B$. Then for the canonical coordinate

$$\vec{x}_B^\Gamma = \vec{x}, \qquad (3.23)$$

for the generators of proper orthochronous Poincaré transformations,

$$\vec{p}_B^\Gamma = \vec{p}, \qquad (3.24a)$$

$$(H_B^{\mathrm{I}})^\Gamma = \varsigma_1 \vec{\sigma}\cdot\vec{p} + \varsigma_3 m, \qquad (3.24b)$$

$$\vec{J}_B^\Gamma = \vec{x}\times\vec{p} + \tfrac{1}{2}\vec{\sigma}, \qquad (3.24c)$$

$$\vec{K}_B^\Gamma = \tfrac{1}{2}[\vec{x}(H_B^{\mathrm{I}})^\Gamma + (H_B^{\mathrm{I}})^\Gamma \vec{x}] - t\vec{p}, \qquad (3.24d)$$

and for the spatial inversion operator $S_B = \varsigma_3$,

$$S_B^\Gamma = \varsigma_3 \mathcal{P}. \qquad (3.25)$$

One finds that the transformed Hamiltonian (3.24b) is identical to the Dirac Hamiltonian (2.19), since only the algebra satisfied by the operators that appear in these equations is relevant. Similarly, the remaining operators in (3.24) and (3.25) are identical to the corresponding operators in the Dirac theory. It could also be demonstrated that the antiunitary (Wigner) time-reversal operator for the BHV theory maps into that for the Dirac theory under the transformation (3.21). Consequently, the free Dirac theory and the BHV theory with either of the Hamiltonians (3.6) describe equivalent representations of the Poincaré group with spatial inversion and antiunitary time reversal included. It is well known that, through a (nonlocal) Foldy-Wouthuysen transformation[13,23] on the Dirac theory, one can show that this representation is reducible to two inequivalent representations with masses $+m$ and $-m$.

While the Hamiltonians (3.24b) and (2.19) are identical physically and mathematically, it should be stressed that the properties of the ς matrices in a theory based on the Dirac Hamiltonian (3.24b) is different from their properties in a theory based on the BHV Hamiltonian (3.12). In the former theory these matrices are all invariant under translations, whereas in the latter ς_2 and ς_3 are not. Furthermore, in the Dirac theory the time dependence is given by (2.30), with the substitution $\rho_j \to \varsigma_j$, but in the BHV theory ς_1 and ς_2 "rotate" among themselves while ς_3 is time-independent. It is the operators ς_1, $\overline{\mathcal{P}} = \varsigma_2\mathcal{P}$, and $\mathcal{P}_{\mathrm{ext}} = \varsigma_3\mathcal{P}$ in the BHV theory that map into the operators ς_1, ς_2, and ς_3, respectively, under the transformation (3.21) between the two representations. For these reasons, it is best to follow BHV by using the ρ notation in the Dirac representation and the ς notation in their representation. But we also wish to emphasize the very different significance of ρ_1 in the eight-component operator (3.1) than in the Dirac Hamiltonian (2.19). In the former case, taking ρ_1 sharp yields a four-component theory describing the same physics as the Dirac equation.

Although the BHV and Dirac theories are equivalent for free particles, it is well known that such theories do not necessarily lead to identical results when interactions are introduced. We shall return to this matter in Secs. V and VI.

IV. A TWO-COMPONENT INTERPRETATION OF THE BHV THEORY

In the preceding section, we were unable to regard the individual two-component operators (3.9) that factorize the Klein-Gordon equation as Hamiltonians because of arguments (a) and (b) following Eq. (3.14). Is there some way of circumventing these two restrictions and finding a theory in which the two-component operators do play the role of Hamiltonians? Our previous arguments offer two clues: Argument (a), in particular, shows that, if we wish to interpret one or more of the operators $H_\pm^{\mathrm{I}}$ and $H_\pm^{\mathrm{II}}$ as Hamiltonians, it is necessary to seek a theory in which all observables split up into subspaces of "sharp" stigma, while (b) shows that we will have to abandon Poincaré invariance.

We could obtain the desired result if we could somehow justify the use of $\vec{p}$ and $\vec{x}$ as canonical operators, as they are in the Dirac theory. Let us therefore start with the free Dirac Hamiltonian (2.19) and add to it a potential

$$V_I = m(\mathcal{P} - \rho_3). \qquad (4.1)$$

The result is the new Hamiltonian

$$H_D' = \rho_1\vec{\sigma}\cdot\vec{p} + m\mathcal{P}, \qquad (4.2)$$

in which $\vec{p}$ is the canonical momentum operator. It is interesting to note that (4.2) can be obtained from (3.1) and (3.2) by taking the stigma sharp there; i.e., $\varsigma_3 \to +1$.

Since H_D' commutes with ρ_1 we may divide it into two subspaces as follows:

$$H_D' = \tfrac{1}{2}(1+\rho_1)H_+^{\mathrm{I}} + \tfrac{1}{2}(1-\rho_1)H_+^{\mathrm{II}}, \qquad (4.3)$$

where H_+^{I} and H_+^{II} are the two-component positive-stigma operators introduced in (3.9), and ρ_1 can be chosen to be sharp.

The Dirac canonical momentum operator $\vec{p}$ commutes with ρ_1 and therefore ρ_1 is translation invariant in this theory. However, neither $H_+^{\rm I}$ nor $H_+^{\rm II}$ is translation-invariant, since from (2.25b) and (3.9),

$$[H_+^{\rm I}, \vec{p}] = [H_+^{\rm II}, \vec{p}] = 2m\mathcal{P}\vec{p}\,. \qquad (4.4)$$

This fact has already been noted by BHV[4,5] and several other authors[12] who also pointed out that the form of the two-component operators $H_+^{\rm I}$ and $H_+^{\rm II}$ is not invariant under boosts. Just as the Dirac equation in a Coulomb potential centered at the origin takes on a more complicated form when one performs a translation or a boost, one can regard (4.2) as the simple form that the Hamiltonian has in a particular reference frame. For example, under a spatial translation

$$H'_D \to \exp(i\vec{d}\cdot\vec{p})H'_D\exp(-i\vec{d}\cdot\vec{p}) = \rho_1\vec{\sigma}\cdot\vec{p} + m\mathcal{P}(\vec{d})\,, \qquad (4.5)$$

where

$$\mathcal{P}(\vec{d}) = \exp(i\vec{d}\cdot\vec{p})\mathcal{P}\exp(-i\vec{d}\cdot\vec{p}) \qquad (4.6)$$

is the operator for spatial reflections about the point $-\vec{d}$.

Since $\vec{p}$, $\vec{x}$, and $\vec{\sigma}$ have nonvanishing matrix elements only between states with the same ρ_1 eigenvalue, it appears correct to regard $H_+^{\rm I}$ and $H_+^{\rm II}$ separately as Hamiltonians that generate time translations within each two-component subspace. Note that either $H_+^{\rm I}$ or $H_+^{\rm II}$ generates time displacement for all physical observables in the present theory, in agreement with requirement (a) following Eq. (3.14), and that we have abandoned the requirement (b) of Poincaré invariance.

From (4.4) and (2.8), the latter of which also applies when interactions are present, we have

$$\dot{\vec{p}} = i2m\mathcal{P}\vec{p}\,, \qquad (4.7)$$

which has the solution

$$\vec{p}(t) = \exp(i2m\mathcal{P}t)\vec{p}(0) = (\cos 2mt + i\mathcal{P}\sin 2mt)\vec{p}(0)\,. \qquad (4.8)$$

Thus, in the present case the canonical momentum operator exhibits *Zitterbewegung* of frequency $2m$ between states having the same energy.

Whether the two-component theory describes any realistic physical situation is not presently clear and we do not wish to pursue the matter further in the present paper. However, in Sec. V arguments will be presented which probably exclude the usual introduction of a general electromagnetic field into such a two-component theory.

V. THE BHV EQUATIONS IN AN EXTERNAL ELECTROMAGNETIC FIELD

A well-known prescription for obtaining a Hamiltonian for a particle in an external electromagnetic field is to make the substitution $\vec{p}_c \to \vec{p}_c - e\vec{A}(\vec{x}_c, t)$ in the corresponding free-particle Hamiltonian and to add to it a term $eA^0(\vec{x}_c, t)$, where $\vec{p}_c$ and $\vec{x}_c$ are the canonical variables introduced in Secs. II and III. This prescription is consistent with the requirement that a gauge transformation

$$A^0(\vec{x}_c, t) \to A^0(\vec{x}_c, t) + \frac{\partial}{\partial t} f(\vec{x}_c, t)\,, \qquad (5.1a)$$

$$\vec{A}(\vec{x}_c, t) \to \vec{A}(\vec{x}_c, t) - \vec{\nabla}_c f(\vec{x}_c, t) \qquad (5.1b)$$

should have no physical effect.

Thus, in place of the free Dirac Hamiltonian (2.19) one employs

$$H_D = \rho_1\vec{\sigma}\cdot[\vec{p} - e\vec{A}(\vec{x}, t)] + \rho_3 m + eA^0(\vec{x}, t)\,, \qquad (5.2)$$

and in place of the free BHV Hamiltonian in (3.6a) and (3.12),

$$H_B^{\rm I} = \varsigma_1\vec{\sigma}\cdot[\vec{p}_B - e\vec{A}(\vec{x}_B, t)] + \mathcal{P}_{\rm ext}m + eA^0(\vec{x}_B, t). \qquad (5.3a)$$

Note that from the Hamiltonian in (3.6b) one obtains

$$H_B^{\rm II} = -\varsigma_1\vec{\sigma}\cdot[\vec{p}_B - e\vec{A}(\vec{x}_B, t)] + \mathcal{P}_{\rm ext}m + eA^0(\vec{x}_B, t)\,, \qquad (5.3b)$$

which is equivalent to (5.3a),

$$H_B^{\rm II} = \mathcal{P}_{\rm ext}H_B^{\rm I}\mathcal{P}_{\rm ext}\,. \qquad (5.4)$$

Using the operator Γ defined in (3.21), we find that

$$(H_B^{\rm I})^\Gamma = \Gamma H_B^{\rm I}\Gamma = \varsigma_1\vec{\sigma}\cdot[\vec{p} - e\vec{A}(\vec{x}, t)] + \varsigma_3 m + eA^0(\vec{x}, t) \qquad (5.5)$$

is identical to the Dirac Hamiltonian (5.2), since once more only the algebra satisfied by the ρ or ς matrices is relevant. Thus, in the present case, the two theories are both mathematically and physically equivalent – every eigenstate of the BHV Hamiltonian (5.3a) maps one-to-one into an eigenstate of the Dirac Hamiltonian (5.2) and vice versa.

We can write

$$A^\mu(\vec{x}, t) = A_+^\mu(\vec{x}, t) + A_-^\mu(\vec{x}, t)\,, \qquad (5.6)$$

where

$$A_\pm^\mu(\vec{x}, t) = \tfrac{1}{2}[A^\mu(\vec{x}, t) \pm A^\mu(-\vec{x}, t)]\,. \qquad (5.7)$$

This gives a division of $A^\mu(\vec{x}, t)$ into parts that are even and odd under spatial inversions, since

$$\mathcal{P}A_\pm^\mu(\vec{x}, t)\mathcal{P} = A_\pm^\mu(-\vec{x}, t) = \pm A_\pm^\mu(\vec{x}, t)\,. \qquad (5.8)$$

Let us define matrices $A_B^\mu(\vec{x}, t)$ by

$$A_B^0(\vec{x}, t) = A^0(\vec{x}_B, t) = A_+^0(\vec{x}, t) + \varsigma_1 A_-^0(\vec{x}, t), \tag{5.9a}$$

$$\vec{A}_B(\vec{x}, t) = \varsigma_1 \vec{A}(\vec{x}_B, t) = \vec{A}_-(\vec{x}, t) + \varsigma_1 \vec{A}_+(\vec{x}, t). \tag{5.9b}$$

These have the interesting property[24]

$$\mathcal{P}_{\text{ext}} A_B^0(\vec{x}, t) \mathcal{P}_{\text{ext}} = A_B^0(\vec{x}, t), \tag{5.10a}$$

$$\mathcal{P}_{\text{ext}} \vec{A}_B(\vec{x}, t) \mathcal{P}_{\text{ext}} = -\vec{A}_B(\vec{x}, t). \tag{5.10b}$$

Equations (5.3) may now be written in the form[25]

$$H_B^{\text{I}} = \vec{\sigma} \cdot [\vec{p} - e\vec{A}_B(\vec{x}, t)] + \mathcal{P}_{\text{ext}} m + eA_B^0(\vec{x}, t), \tag{5.11a}$$

$$H_B^{\text{II}} = -\vec{\sigma} \cdot [\vec{p} - e\vec{A}_B(\vec{x}, t)] + \mathcal{P}_{\text{ext}} m + eA_B^0(\vec{x}, t), \tag{5.11b}$$

which look very much like two-component equations with only the spin degrees of freedom. BHV, in fact, choose to formally regard them as two-component operators. Note, however, that in the limit of a vanishing field, the Hamiltonians (5.11) reduce to the free BHV Hamiltonians (3.6), which we have seen can be viewed as four-component operators in the same sense that the Dirac and Sakata-Taketani Hamiltonians are, respectively, four and two-component ones.

In general, ς_3 does not commute with either of the Hamiltonians (5.11). However, for the particular case of an "even environment[26]" under spatial inversion, when $A_-^0(\vec{x}, t) = 0$ and $\vec{A}_+(\vec{x}, t) = 0$, one has from (5.6) and (5.11a)

$$H_B^{\text{I}} = \vec{\sigma} \cdot [\vec{p} - e\vec{A}(\vec{x}, t)] + \varsigma_3 \mathcal{P} m + eA^0(\vec{x}, t), \tag{5.12}$$

and stigma can be taken sharp[18]; i.e.,

$$H_B^{\text{I}} = \tfrac{1}{2}(1 + \varsigma_3) H_+^{\text{I}} + \tfrac{1}{2}(1 - \varsigma_3) H_-^{\text{I}}, \tag{5.13}$$

where

$$H_\pm^{\text{I}} = \vec{\sigma} \cdot [\vec{p} - e\vec{A}(\vec{x}, t)] \pm \mathcal{P} m + eA^0(\vec{x}, t). \tag{5.14}$$

As was the case for the corresponding operators (3.9a) in a vanishing electromagnetic field, neither H_+^{I} nor H_-^{I} is a Hamiltonian for a particle that experience *only* an electromagnetic force. One reason has already been given in the discussion following Eq. (3.14); i.e., H_+^{I} and H_-^{I} do not satisfy condition (a) there. Furthermore, if we make a gauge transformation (5.1) with $f(\vec{x}, t) \neq f(-\vec{x}, t)$, or if we perform a spatial translation or a boost to a new reference frame, we no longer have an even environment under spatial inversion with $A_-^0(\vec{x}, t) = 0$ and $\vec{A}_+(\vec{x}, t) = 0$ and cannot make the decomposition (5.13). [Since we do not have the free-particle situation here condition (b) following (3.14) does not apply. From the considerations of Sec. IV we know that neither H_+^{I} nor H_-^{I} in (5.14) can individually describe a particle under the influence of *only* electromagnetic forces. At the end of this section we shall answer the question as to whether our alternative interpretation of the BHV theory allows us to regard each of the operators in (5.14) as a Hamiltonian for a particle that is also under the influence of a potential of the type (4.1).]

For the special case of a Coulomb potential with $A^0 = -Ze/|\vec{x}|$ and $\vec{A} = 0$, (5.14) becomes

$$H_\pm^{\text{I}} = \sigma \cdot \vec{p} \pm \mathcal{P} m - \frac{Ze^2}{|\vec{x}|}. \tag{5.15}$$

As shown by BHV,[27] each of the operators H_+^{I} and H_-^{I} yields a nondegenerate spectrum and the two together give the same spectrum and the same degeneracy as the Dirac Hamiltonian. However, since according to the preceding paragraph only H_B^{I}, and not just H_+^{I} or H_-^{I}, describes a particle experiencing only an attractive Coulomb potential, the results are identical to those of the Dirac theory.

When we pass to the Schrödinger picture, the Dirac Hamiltonian (5.2) yields the equation

$$(\pi^0 - \rho_1 \vec{\sigma} \cdot \vec{\pi} - \rho_3 m)\Psi_D(\vec{x}, t) = 0, \tag{5.16}$$

where

$$\pi^0 = p^0 - eA^0(\vec{x}, t), \tag{5.17a}$$

$$\vec{\pi} = \vec{p} - e\vec{A}(\vec{x}, t). \tag{5.17b}$$

Similarly, the BHV Hamiltonians (5.11) lead to the equations

$$(\pi_B^0 \mp \vec{\sigma} \cdot \vec{\pi}_B - \mathcal{P}_{\text{ext}} m)\Psi_B^\pm(\vec{x}, t) = 0, \tag{5.18}$$

where

$$\pi_B^0 = p^0 - eA_B^0(\vec{x}, t), \tag{5.19a}$$

$$\vec{\pi}_B = \vec{p} - e\vec{A}_B(\vec{x}, t). \tag{5.19b}$$

From (5.10) and (3.3a) we obtain

$$\mathcal{P}_{\text{ext}} \pi_B^0 \mathcal{P}_{\text{ext}} = \pi_B^0, \tag{5.20a}$$

$$\mathcal{P}_{\text{ext}} \vec{\pi}_B \mathcal{P}_{\text{ext}} = -\vec{\pi}_B. \tag{5.20b}$$

This result leads to the first of the two identities

$$\{(\pi_B^0)^2 - (\vec{\sigma} \cdot \vec{\pi}_B)^2 - \rho_1[\pi_B^0, \vec{\sigma} \cdot \vec{\pi}_B] - m^2\} = (\pi_B^0 + \rho_1 \vec{\sigma} \cdot \vec{\pi}_B + \mathcal{P}_{\text{ext}} m)(\pi_B^0 - \rho_1 \vec{\sigma} \cdot \vec{\pi}_B - \mathcal{P}_{\text{ext}} m) \tag{5.21a}$$

$$= (\pi_B^0 + \rho_1 \vec{\sigma} \cdot \vec{\pi}_B + \rho_3 m)(\pi_B^0 - \rho_1 \vec{\sigma} \cdot \vec{\pi}_B - \rho_3 m). \tag{5.21b}$$

For obvious reason we shall refer to (5.21a) and (5.21b), respectively, as the BHV and Dirac factorizations of the operator on the left-hand side.

Both ρ_1 and ς_1 commute with the operator on the left-hand side of (5.21), and if the former is taken sharp ($\rho_1 \to \pm 1$) then only the BHV factorization

$$\{(\pi_B^0)^2 - (\vec{\sigma}\cdot\vec{\pi}_B)^2 \mp [\pi_B^0, \vec{\sigma}\cdot\vec{\pi}_B] - m^2\} = (\pi_B^0 \pm \vec{\sigma}\cdot\vec{\pi}_B + \mathcal{P}_{\text{ext}}m)(\pi_B^0 \mp \vec{\sigma}\cdot\vec{\pi}_B - \mathcal{P}_{\text{ext}}m) \tag{5.22}$$

exists. On the other hand, if ς_1 is taken sharp (e.g., $\varsigma_1 \to 1$) there exists only the Dirac factorization

$$\{(\pi^0)^2 - (\vec{\sigma}\cdot\vec{\pi})^2 - \rho_1[\pi^0, \vec{\sigma}\cdot\vec{\pi}] - m^2\} = (\pi^0 + \rho_1\vec{\sigma}\cdot\vec{\pi} + \rho_3 m)(\pi^0 - \rho_1\vec{\sigma}\cdot\vec{\pi} - \rho_3 m). \tag{5.23}$$

[Note that if $\varsigma_1 \to -1$, we obtain the analog of (5.23) with π^0 replaced by $p^0 - A^0(-\vec{x}, t)$ and with $\vec{\pi}$ replaced by $\vec{p} + \vec{A}(-\vec{x}, t)$.] The two choices of sign in (5.22) are equivalent under a transformation with $\mathcal{P}_{\text{ext}}$, and under a transformation with Γ in (3.21), (5.22) with the upper sign maps into

$$\{(\pi^0)^2 - (\vec{\sigma}\cdot\vec{\pi})^2 - \varsigma_1[\pi^0, \vec{\sigma}\cdot\vec{\pi}] - m^2\} = (\pi^0 + \varsigma_1\vec{\sigma}\cdot\vec{\pi} + \varsigma_3 m)(\pi^0 - \varsigma_1\vec{\sigma}\cdot\vec{\pi} - \varsigma_3 m), \tag{5.24}$$

which is just (5.23), since once more only the algebra satisfied by the operators is important. Therefore, (5.22) and (5.23) are equivalent.

By taking both ς_1 and ρ_1 sharp in (5.21) we obtain the two-component operator that appears in the Kramers equation[1,2]:

$$\{(\pi^0)^2 - (\vec{\sigma}\cdot\vec{\pi})^2 \mp [\pi^0, \vec{\sigma}\cdot\vec{\pi}] - m^2\}\Psi_K^{\pm} = 0, \tag{5.25}$$

but neither the Dirac nor the BHV factorization then directly exists. It was by formally regarding (5.22) as a two-component equation that BHV concluded that they had obtained a two-component factorization of the Kramers equation for an electromagnetic field.[4,5] The approach developed here shows that the Dirac and BHV operators have an equal right to be regarded as four-component factorizations of the two-component Kramers operator.

Can one extend the results of Sec. IV so that $H_\pm^{\text{I}}$ and $H_\pm^{\text{II}}$ in (3.9) become Hamiltonians describing a spin-$\frac{1}{2}$ particle under the influence of both an external electromagnetic field and the potential V_I in (4.1)? This cannot be accomplished in the usual way, since the substitution $\vec{p} \to \vec{p} - e\vec{A}(\vec{x}, t)$ and the addition of $eA^0(\vec{x}, t)$ to the individual two-component operators in (3.9) would give results depending on the particular gauge (5.1) chosen. Thus, the possibility of a two-component Hamiltonian with a Coulomb term, as in (5.15), is ruled out. Of course, as in the case with both the Dirac Hamiltonian and the four-component BHV Hamiltonian, one can always add terms depending on the components of the electromagnetic four-tensor to the individual operators in (3.9), but we shall make no attempt here to analyze the physical implications.

VI. SUMMARY

By working in the Heisenberg picture we have established that, if the Hamiltonian for a relativistic quantum theory of a free spin-J particle factorizes the Klein-Gordon operator, there is a doubling of the spin space, with a certain exception. The extra degrees of freedom are present because for each value of the canonical momentum and for each value of the canonical helicity, there are both positive- and negative-energy solutions. We have seen that the BHV Hamiltonian, like the Dirac and Sakata-Taketani spin-0 Hamiltonians, provides an example of such a theory.

The Dirac and the BHV theories are equivalent in the presence of an electromagnetic field, but, since the transformation (3.21) between the two is nonlocal, there may exist interactions that destroy this equivalence. If this is the case it is likely, as emphasized by BHV, to occur in the case of weak interactions, but the possibility has not yet been demonstrated. BHV have placed great emphasis on the importance of chirality,[5,28] which is the $\pm$ sign that appears in front of the $\vec{\sigma}\cdot[\vec{p} - e\vec{A}_B(\vec{x}, t)]$ term in (5.11). For chirality to be significant the interaction would have to be such that the operator $\mathcal{P}_{\text{ext}}$ no longer provides a mapping between the two alternatives as it does in (5.4).

ACKNOWLEDGMENTS

The authors wish to acknowledge very helpful discussions with Professor L. C. Biedenharn, Professor Ian Duck, and Professor George Trammell.

*Work supported in part by the U. S. Atomic Energy Commission.

†Present address: Oak Ridge National Laboratory, Oak Ridge, Tenn. 37830.

[1]H. A. Kramers, *Quantum Mechanics* (Interscience New York, 1957).

[2]L. M. Brown, in *Lectures in Theoretical Physics*, edited by W. E. Brittin, B. W. Downs, and J. Downs

(Interscience, New York, 1962), Vol. IV, p. 324.

[3]See, for example, R. P. Feynman and M. Gell-Mann, Phys. Rev. 109, 193 (1958).

[4]L. C. Biedenharn and M. Y. Han, Phys. Letters 36B, 475 (1971); L. C. Biedenharn, M. Y. Han, and H. van Dam, Phys. Rev. Letters 27, 1167 (1971); Lett. Nuovo Cimento 2, 730 (1971).

[5]L. C. Biedenharn, M. Y. Han, and H. van Dam, Phys. Rev. D 6, 500 (1972).

[6]E. Wigner, Ann. Math. 40, 39 (1939).

[7]Iu. M. Shirokov, Zh. Eksp. Teor. Fiz. 33, 861 (1957) [Sov. Phys. JETP 6, 664 (1958)]; 33, 1196 (1957) [6, 919 (1958)]; 33, 1208 (1957) [6, 929 (1958)].

[8]See, for example, S. Weinberg, Phys. Rev. 134, B882 (1964).

[9]S. Sakata and M. Taketani, Proc. Phys. Soc. (Japan) 22, 757 (1940).

[10]H. Feshbach and F. Villars, Rev. Mod. Phys. 30, 24 (1958). See also A. S. Davydov, *Quantum Mechanics*, translated by D. Ter Haar (Addison-Wesley, Reading, Mass., 1965), Chap. VIII.

[11]E. Guth, Ann. Phys. (N.Y.) 20, 309 (1962).

[12]Several other authors have also pointed out that the two-component BHV equations are not Poincaré-invariant: H. Grotch and E. Kazes, Phys. Rev. D 5, 3277 (1972); M. Flato, G. Lindblad, and B. Nagel, Physica Scripta 5, 153 (1972); R. Acharya and R. Madan, report, 1972 (unpublished).

[13]L. L. Foldy, Phys. Rev. 102, 568 (1956).

[14]P. A. M. Dirac, Proc. Roy. Soc. (London) A328, 1 (1972).

[15]E. Schrödinger, Sitzber. Preuss. Akad. Wiss. Physik Math. 24, 418 (1930).

[16]Iu. M. Shirokov, Zh. Eksp. Teor. Fiz. 34, 717 (1958) [Sov. Phys. JETP 7, 493 (1958)].

[17]Reference 5, p. 512.

[18]Reference 5, Eq. (93).

[19]This has been emphasized by R. H. Good, Phys. Rev. D 5, 1538 (1972).

[20]Reference 5, p. 517.

[21]Reference 5, p. 510, p. 511, and remarks following Eq. (136).

[22]Reference 5, Eq. (123).

[23]L. L. Foldy and S. A. Wouthuysen, Phys. Rev. 78, 29 (1950); S. Tani, Progr. Theoret. Phys. (Japan) 6, 267 (1951); R. Becker, Nachr. Akad. Wiss. Göttingen, Math. Physik, Kl. 1, 20 (1945).

[24]In the terminology of Good (Ref. 19) and also BHV (Refs. 4 and 5), the potentials $A^{\mu}_{B}(\vec{x},t)$ are such that one has an "even environment" with respect to $\mathcal{P}_{ext}$.

[25]Compare our Eqs. (5.9) and (5.11) to Eq. (54′) of Ref. 5.

[26]Again we adopt the terminology of Good and BHV, as in Ref. 24. In this particular case, the potentials $A^{\mu}(\vec{x},t)$ are such that one has an even environment with respect to $\mathcal{P}$.

[27]Reference 5, Sec. III.

[28]L. C. Biedenharn, M. Y. Han, and H. van Dam, Phys. Rev. D 5, 1539 (1972).

PHYSICAL REVIEW D VOLUME 7, NUMBER 4 15 FEBRUARY 1973

General Theory of Broken Local Symmetries*

Steven Weinberg
Laboratory for Nuclear Science and Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 4 October 1972)

A general formalism for theories with spontaneously broken local symmetries is developed in the unitarity gauge. The canonical quantization procedure is carried out, leading to a set of Lorentz-covariant Feynman rules. Various special topics are discussed.

I. INTRODUCTION AND SUMMARY

Models with spontaneously broken local symmetries have been suggested[1] as a solution to two of the major problems of elementary-particle theory:

(a) the unification[2] of the weak and electromagnetic interaction;

(b) the elimination of ultraviolet divergences appearing in higher-order effects of the weak interactions.

Recently there have been indications[3] that such models may also elucidate one other outstanding problem:

(c) the explanation of the weak breaking of intrinsic symmetries such as isospin.

The purpose of the present paper is to provide a formal foundation for general theories with spontaneously broken local symmetries.

The formalism described here is based on choice of a particular gauge, the "unitarity gauge," in which the absence of Goldstone bosons and the order-by-order unitarity of these theories is manifest. This is the gauge which was originally used to show that, instead of Goldstone bosons appearing when local symmetries are spontaneously brok-