Potential scattering of Dirac particles

1981 J. Phys. A: Math. Gen. 14 3067

(http://iopscience.iop.org/0305-4470/14/11/028)

J. Phys. A: Math. Gen. **14** (1981) 3067–3083. Printed in Great Britain

Potential scattering of Dirac particles

B Thaller†

Institut für Theoretische Physik, Universität Graz, A-8010 Graz, Austria

Received 3 March 1981, in final form 30 April 1981

Abstract. In this paper we consider the Dirac equation with external electromagnetic fields and give a careful formulation of the Stückelberg–Feynman interpretation (i.e. the negative energy waves are directly interpreted as positron solutions). This electron–positron theory may be applied whenever electron and positron states can be separated, which is possible for a wide class of time-independent potentials vanishing at infinity. One then may introduce a system of observables with a position operator $\boldsymbol{X}(t)$ which is free of Zitterbewegung and yields the usual relativistic connection between momentum and velocity. The Stückelberg–Feynman interpretation is usually associated with a scattering theory (propagator formalism) which differs from the one of a one-particle interpretation (formalism imitating non-relativistic techniques). For the above-mentioned potentials there is a simple relationship between these physically different formalisms. There are, however, more general potentials for which the Stückelberg–Feynman interpretation may still be applied asymptotically. With the help of Klein's paradox we show that the Dirac equation may then have unphysical solutions not conserving the total charge (in that case a unitary Feynman scattering operator does not exist).

1. Introduction

In this paper we shall discuss the quantum mechanical interpretation of the Dirac equation (2.1) for particles in external electromagnetic potentials. We are interested in the following questions: what are the external fields for which the Dirac equation (treated as a quantum mechanical wave equation) yields reasonable results (such fields exist; remember the hydrogen atom), and what happens if one tries to calculate more general situations? Our main problem is that up to now no particular quantum mechanical interpretation is generally accepted. Throughout this paper we shall apply the Stückelberg–Feynman interpretation (Stückelberg 1942, Feynman 1949). It is intermediate between a one-particle theory and Dirac's hole theory, because it claims that the Dirac equation is able to describe two kinds of particles, namely electrons and positrons (but not their interaction; negative energy states are directly observed as positrons with positive energy). Our theory is formulated in the language of wave-packets (i.e. in a Hilbert-space context) and does not rely on unobservable objects like the Dirac sea.

In § 2 the separation of electron and positron states is carried through for a wide class of external fields (even if bound states appear with both signs of energy). These potentials, which are time independent and vanish at infinity, allow no transitions to states with different energy, and therefore it depends only on the initial conditions, whether the wavefunction describes an electron or a positron. This viewpoint has

† Supported by Fonds zur Förderung der Wissenschaftlichen Forschung in Österreich, project No 3764.

interesting consequences. It is possible to define electron and positron observables by restricting the observables of one-particle theories to the electron and positron subspaces. One then obtains a position operator with non-commuting components whose commutator with the energy yields the usual relativistic connection between momentum and velocity $v = p/E$ without 'Zitterbewegung'.

The Klein paradox (i.e. the scattering by very strong step-potentials; § 3) is an example of a more general situation, because the potential does not vanish at infinity. A separation of electrons and positrons is only possible for the asymptotic states. It turns out that there are transitions from electron to positron states, but that these solutions do not conserve total charge. (Because of time reversal invariance it is, however, possible to find charge conserving solutions for the scattering at static potential steps.)

Finally, we turn to the general theory of potential scattering. The Stückelberg–Feynman interpretation is basic to Feynman's theory of potential scattering (Feynman 1949), which is presented in every modern textbook as a simple application of the propagator formalism (cf Bjorken and Drell 1964 ch 6, Pilkuhn 1979, ch 2.6, Scadron 1979, ch 7). Prosser (1963) developed for the first time a mathematically rigorous theory which originates in a one-particle interpretation (see also Eckhardt 1974, Pearson 1977, Bongaarts and Ruijsenaars 1977). This formalism 'imitates' non-relativistic time-dependent scattering theory by substituting Dirac for Schrödinger operators. In § 4 it is our intention to give a rigorous Hilbert-space formulation of the heuristic Feynman theory and to compare it with the theory of Prosser. We obtained the following results. The two formalisms are mathematically equivalent for the potentials allowing a separation of electron and positron states. For more general situations like the Klein paradox, the two formalisms are no longer equivalent. It is possible to find a unitary scattering operator $\tilde{S}$ in the sense of Prosser (cf Bongaarts and Rujsenaars 1976, 1977), but the existence of unphysical solutions prevents the definition of a unitary operator S in the sense of Feynman. The same difficulty arises with strongly time-dependent potentials: conservation of norm prevents any change in the number of particles, whereas transitions from electron to positron states give rise to solutions that do not conserve total charge (i.e. $\tilde{S}$ may exist and be unitary, but a unitary S cannot be found). We thus conclude that pair creation, which may actually happen for sufficiently high scattering or bound state energies, is not described by the Dirac equation without using explicitly quantum field theory.

2. The quantum mechanical interpretation

A one-particle interpretation of the Dirac wavefunction ($|\psi(x)|^2$ = position probability density for electrons, $e|\psi(x)|^2$ = charge density) leads to some well known difficulties.

(1) A careful study of the expectation value of $\boldsymbol{X}(t)$ reveals the phenomenon of 'Zitterbewegung' resulting from an interference between the positive and negative energy part of the wavefunction. It has some peculiar properties: for plane waves, Zitterbewegung is a steady-state phenomenon, but it is transient for wavepackets (i.e. the oscillatory motion is damped, and asymptotically the wavepackets move according to Newton's first law). For special initial conditions (e.g. only positive or negative energies) it is totally absent (cf Lock 1979, Barut and Bracken 1981).

(2) Another drawback is the occurrence of bound states in repulsive potentials. The electron Dirac equation with a repulsive Coulomb potential Ze^2/r gives bound states at $-E_n$, if E_n are the hydrogen energy levels (attractive Coulomb potential $-Ze^2/r$); cf

Brysk and Zweifel (1981), Messiah (1962). A square well may even draw energy levels beyond $E = 0$ (cf Rafelski *et al* 1978).

It is the purpose of this section to present another quantum mechanical interpretation which associates a position probability density with the wavefunctions, but which is free of Zitterbewegung and gives a simple explanation for the bound states in repulsive potentials.

2.1. The Dirac equation

The Dirac equation in Hamiltonian form reads

$$\mathrm{i}\frac{\partial}{\partial t}\psi = H(g)\psi, \qquad \psi \in \mathscr{L}^2(\mathbb{R}^3)^4, \qquad \hbar = c = 1, \qquad g \in R,$$

$$H(g) = H_0 + gV, \qquad H_0 = -\mathrm{i}\gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{\nabla} + \gamma^0 m, \qquad V = \gamma^0\gamma^\mu A_\mu. \tag{2.1}$$

(In the following we shall concentrate on the electron equation: $g = e = -|e|$; the discussion is completely analogous for positrons $g = -e$. For the metric and notation of γ matrices see Bjorken and Drell (1964).)

In order to interpret the wavefunction we have to impose some conditions on the external potential. Let V be time independent. Assume $H(g)$ to be self-adjoint and its eigenvalues continuous and strictly monotone in g (the eigenfunctions should also depend continuously on g; cf the definition (III) of § 2.2). Furthermore assume the following decomposition of the spectrum:

$$\begin{aligned} &\sigma(H) = \sigma_{\mathrm{p}}(H) \cup \sigma_{\mathrm{ac}}(H), \\ &\sigma_{\mathrm{ac}}(H) = (-\infty, -m] \cup [m, \infty), \qquad \sigma_{\mathrm{p}}(H) \subseteq [-m, +m]. \end{aligned} \tag{2.2}$$

(The point spectrum $\sigma_{\mathrm{p}}(H)$ is the set of all eigenvalues of H, and $\sigma_{\mathrm{ac}}(H)$ is the absolutely continuous spectrum of scattering energies.) These conditions are very restrictive: neglecting mathematical details (one has to exclude some pathological potentials), it is required that each component of the 4×4 matrix $V(\boldsymbol{x})$ vanishes at infinity (e.g. like $(1/|\boldsymbol{x}|)$ and is less singular than the Coulomb potential with $Z = 137$ in any finite region. We shall see in the next section what happens if only weaker conditions are imposed. (More details concerning mathematical problems can be found in the following references. A review of the self-adjointness problem is given e.g. in Jörgens (1973); for the absence of the singular-continuous spectrum see Weidmann (1971), Weder (1973), Pearson (1977) and Simon (1979); a treatment including Coulomb forces can be found in Nenciu (1976), Wüst (1977), Klaus and Wüst (1979); see also Landgren and Reito (1979), Landgren *et al* (1980); the absence of eigenvalues embedded in the continuum is considered in Kalf (1976, 1980); the dependence of the point spectrum on g is investigated in Klaus (1980).)

The Dirac equation is form invariant under Poincaré transformations; the requirement of time independence of the external potential destroys, however, the covariance of the theory, because it gives preference to a special Lorentz frame. It would be very difficult to admit time-dependent potentials, because then the decomposition of the Hilbert space generally would be impossible.

2.2. Decomposition of the Hilbert space

We consider first the simpler problem of interpreting scattering states. According to

(2.2) any scattering wavefunction $\psi \in \mathscr{H}_{ac}(H)$ can be written uniquely as $\psi = \psi^+ + \psi^-$, where the positive and negative energy parts are orthogonal (i.e. the subspace of scattering wavefunctions may be decomposed further into two orthogonal Hilbert spaces; we write $\mathscr{H}_{ac}(H) = \mathscr{H}^+_{ac}(H) \oplus \mathscr{H}^-_{ac}(H)$). This is very important, because now we can interpret the two parts separately.

(I) The wavepacket $\psi^+ \in \mathscr{H}^+_{ac}(H)$ describes an electron with positive energy moving in the external potential V.

(II) The wavepacket $\psi^- \in \mathscr{H}^-_{ac}(H)$ describes a positron with positive energy moving in the same external potential V.

The identification of positive energy states as electrons is obvious, but (II) needs further justification.

(a) Note that (II) is in accordance with the Stückelberg–Feynman interpretation, where the scattering states ψ^- are treated as electrons travelling backwards in time. In the formalism they appear as electrons with negative energies, but experimentally they are observed as positrons with positive energies (Stückelberg 1942, Feynman 1949).

(b) ψ^- is a solution of (2.1) and not of the positron Dirac equation (with $H(-e)$), but we can construct positron wavefunctions via charge conjugation: $(\mathscr{C}\psi^-)(\boldsymbol{x}, t) := i\gamma^2\overline{\psi^-(\boldsymbol{x}, t)}$ is a solution of the positron equation ($e \to -e$, but same external potential V) with positive energy, whenever $\psi^-(\boldsymbol{x}, t)$ is a solution of the electron equation (2.1) with negative energy. If we interpret $|\psi^{\pm}(\boldsymbol{x}, t)|^2 \equiv \sum_{i=1}^4 |\psi_i^{\pm}(\boldsymbol{x}, t)|^2$ as position probability density, our definition (II) will be justified by

$$|(\mathscr{C}\psi^-)(\boldsymbol{x}, t)|^2 = |\psi^-(\boldsymbol{x}, t)|^2. \tag{2.3}$$

The motion of a negative energy wavepacket is observed as the motion of a positron with positive energy.

For bound states the sign of the energy $E \in \sigma_p(H)$ does not distinguish between electron and positron wavefunctions (a square well, repulsive or attractive, may have eigenvalues with both signs in $[-m, +m]$). So we use a more refined definition to separate bound states into electron and positron parts†.

(III) A bound state with energy $E \in \sigma_p(H)$ is called an electron (positron) bound state if the eigenvalue decreases (increases) with increasing coupling strength $|g|$. Position probabilities are defined as in (I) and (II).

Here we have used the monotony of the eigenvalues. This is indeed a decomposition of the subspace of bound states into two orthogonal parts. Usually, electron and positron bound states will have different energies and are therefore orthogonal, but for a certain value of g the energies may be equal (if an increasing positron-eigenvalue crosses a decreasing electron-eigenvalue; because of the strict monotony this is only possible for discrete values of g). Then the continuity with respect to g is sufficient to prove orthogonality for all g.

The definitions (I), (II) and (III) may be summarised as follows: if all the conditions on the external potential V are fulfilled, it is possible to decompose the total Hilbert space $\mathscr{L}^2(\mathbb{R}^3)^4$ into two orthogonal parts, namely into an electron and a positron Hilbert space:

$$\mathscr{L}^2(\mathbb{R}^3)^4 = \mathscr{H}^+(H) \oplus \mathscr{H}^-(H). \tag{2.4}$$

(Each may be decomposed further into a subspace of scattering states and a subspace of

† This definition also holds for bound states in scalar potentials (cf Dosch *et al* 1971) and Landau levels in magnetic fields (Sokolov and Ternov 1968).

bound states: $\mathscr{H}^{\pm}(H) = \mathscr{H}^{\pm}_{\mathrm{p}}(H) \oplus \mathscr{H}^{\pm}_{\mathrm{ac}}(H)$.) We write $P^{\pm}(H)$ for the orthogonal projectors on $\mathscr{H}^{\pm}(H)$. For $H = H_0$, the free Hamiltonian, we simply have

$$P^{\pm}(H_0) \equiv P^{\pm}_0 = \theta(\pm H_0) = \tfrac{1}{2}(1 \pm H_0/|H_0|), \qquad |H_0| = (-\Delta + m^2)^{1/2}. \tag{2.5}$$

Equation (2.4) implies an analogous decomposition of wavefunctions: $\psi = \psi^+ + \psi^-$, $\psi^{\pm} \equiv P^{\pm}(H)\psi \in \mathscr{H}^{\pm}(H)$. Having interpreted $|\psi^{\pm}(\boldsymbol{x}, t)|^2$ as position probability densities, the conserved current is consequently defined by

$$\begin{aligned} j^{\mu +} &= e(\psi^+)^* \gamma^0 \gamma^\mu \psi^+ \qquad \text{for electrons,} \\ j^{\mu -} &= -e(\psi^-)^* \gamma^0 \gamma^\mu \psi^- \qquad \text{for positrons.} \end{aligned} \tag{2.6}$$

In a one-particle theory one would have $j^\mu = e\psi^* \gamma^0 \gamma^\mu \psi$, with $\psi = a\psi^+ + b\psi^-$. Here we want to exclude such expressions by introducing a *superselection rule* (cf for example Streater and Wightman (1964, ch 1.1)). One has never observed particles being a superposition of electrons and positrons. We therefore demand that only the states belonging to one of the 'coherent subspaces' $\mathscr{H}^{\pm}(H)$ represent physical states. This charge superselection rule is only meaningful as long as time evolution leaves $\mathscr{H}^{\pm}(H)$ invariant. Then the initial state $\psi^+ \in \mathscr{H}^+(H)$ ($\psi^- \in \mathscr{H}^-(H)$) will remain an electron (positron) state for all times. This condition is fulfilled by our time-independent Hamiltonian of § 2.1. So it depends only on the initial conditions, whether we describe an electron or a positron. (The Klein paradox represents a completely different situation, cf § 3.)

2.3. Electron and positron observables

We still need a rule for obtaining values for measurable quantities (= expectation values of self-adjoint operators), i.e. we have to define the observables of the theory. Because of the superselection rule (§ 2.2) not every self-adjoint operator A can be an observable; only operators leaving $\mathscr{H}^{\pm}(H)$ invariant are admissible. One can obtain such operators by restricting A to $\mathscr{H}^{\pm}(H)$ (cf Pryce 1948), but as can be seen from below, only an additional charge conjugation of the positron expectation values would yield reasonable results. Thus we redefine positron observables according to

$$(\mathscr{C}\psi^-, A\mathscr{C}\psi^-) = (\psi^-, \mathscr{C}^* A^* \mathscr{C}\psi^-). \tag{2.7}$$

(* denotes the adjoint operator; we have $\mathscr{C}^* = \mathscr{C}^{-1}$ and for observables $A^* = A$. For the charge conjugation of expectation values see Messiah (1962, ch 20.6.1).)

(IV) Let $A(e)$ be self-adjoint in $\mathscr{L}^2(\mathbb{R}^3)^4$. An electron observable is represented by the self-adjoint operator

$$A^+ := P^+ A(e) P^+,$$

a positron observable by

$$A^- := P^- \mathscr{C}^{-1} A(-e) \mathscr{C} P^-.$$

Consider as a first example the energy ($\mathscr{C}^{-1} H(-e) \mathscr{C} = -H(e)$)

$$\begin{aligned} H^+ &= P^+ H(e) P^+, & H^+\psi^+ &= H(e)\psi^+, \\ H^- &= P^- \mathscr{C}^{-1} H(-e) \mathscr{C} P^-, & H^-\psi^- &= -H(e)\psi^-. \end{aligned} \tag{2.8}$$

Particles in a scattering state always have positive energies. The bound states in

potentials which are repulsive for electrons are of course positron bound states. The Dirac equation with the 'repulsive' potential $+Ze/r$ has eigenvalues at $-E_n$, but the actually observed positron energies are $+E_n$.

By definition, electron observables always commute with positron observables. A^+ and A^- may have properties completely different from those of A. For example one generally has $[A^+, B^+] \neq [A, B]P^+$ unless A and B commute with P^+. Note further that the definition of observables depends on the interaction.

Now we turn to the free-particle observables. The generators of the Poincaré group are represented by ten self-adjoint operators in $\mathscr{L}^2(\mathbb{R}^3, \mathrm{d}^3x)^4$, namely

$$\begin{aligned} \boldsymbol{P} &:= -\mathrm{i}\boldsymbol{\nabla}, & H_0 &:= \gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{P} + \gamma^0 m, \\ \boldsymbol{J} &:= \boldsymbol{X}\times\boldsymbol{P} + \boldsymbol{S}, & \boldsymbol{K} &:= \tfrac{1}{2}(\boldsymbol{X}H_0 + H_0\boldsymbol{X}) - t\boldsymbol{P}. \end{aligned} \tag{2.9}$$

With X_i = 'multiplication by x_i' and

$$S_i := \tfrac{1}{4}\mathrm{i}\varepsilon_{ijk}[\gamma_j, \gamma_k] \equiv \tfrac{1}{2}\varepsilon_{ijk}\sigma^{jk} \tag{2.10}$$

we obtain a realisation of the well known Poincaré algebra. Both $\boldsymbol{X}$ and $\boldsymbol{S}$ do not leave invariant $\mathscr{H}^\pm(H_0) \equiv \mathscr{H}_0^\pm$, but, if G is a Poincaré generator, then

$$[G, H_0/|H_0|] = 0. \tag{2.11}$$

$\mathscr{H}_0^\pm$ are therefore invariant with respect to Poincaré transformations; electrons and positrons are characterised in an invariant manner. G^+ is also a representation of the Poincaré algebra, and G^- is a representation with complex conjugate structure constants, because $\mathscr{C}$ is antilinear and $\mathscr{C}^{-1}G\mathscr{C} = -G$ for all G. In the following we give the most important formal properties of the electron and positron observables.

Momentum:

$$\boldsymbol{P}^\pm\psi^\pm = \mp\mathrm{i}\boldsymbol{\nabla}\psi^\pm. \tag{2.12}$$

Energy: From $H_0 = |H_0|P_0^+ - |H_0|P_0^-$ and $\mathscr{C}^{-1}H_0\mathscr{C} = -H_0$ we obtain

$$H_0^\pm\psi^\pm = [(\boldsymbol{P}^\pm)^2 + m^2]^{1/2}\psi^\pm = (-\Delta + m^2)^{1/2}\psi^\pm. \tag{2.13}$$

Position and velocity: $(\mathscr{C}^{-1}\boldsymbol{X}\mathscr{C} = \boldsymbol{X})$

$$\boldsymbol{X}^\pm\psi^\pm = [\boldsymbol{X} \mp (\mathrm{i}/2|H_0|)(\gamma^0\boldsymbol{\gamma} - \boldsymbol{P}/H_0)]\psi^\pm. \tag{2.14}$$

$\boldsymbol{X}^\pm$ is invariant with respect to the Wigner time reversal, and fulfils canonical commutation relations

$$[X_i^\pm, P_j^\pm] = \pm\mathrm{i}\delta_{ij} \tag{2.15}$$

and rotates as a vector

$$[X_i^\pm, J_j^\pm] = \mp\mathrm{i}\varepsilon_{ijk}X_k^\pm. \tag{2.16}$$

Its commutator with $H_0^\pm$ yields for the velocity

$$\mathrm{d}\boldsymbol{X}^\pm/\mathrm{d}t = \pm\mathrm{i}[H_0^\pm, \boldsymbol{X}^\pm] = \boldsymbol{P}^\pm/|H_0| \tag{2.17}$$

i.e. the observed velocity always points in the direction of the observed momentum. As expected, there is no term describing 'Zitterbewegung'. $\boldsymbol{X}^\pm$ is nonlocal in the sense that its components do not commute:

$$[X_i^\pm, X_j^\pm] = -(\mathrm{i}/2|H_0|^2)P_0^\pm\sigma_{ij}P_0^\pm \tag{2.18}$$

with σ_{ij} defined by (2.10). Thus, considering the Dirac equation in its standard representation, we are led to a position operator with the unfamiliar property (2.18). Starting within the Foldy–Wouthuysen representation (which is only known for free particles; Foldy and Wouthuysen (1950)), one is led to the Newton–Wigner operator by similar considerations (Newton and Wigner 1949). This operator satisfies properties analogous to (2.15)–(2.17) but has commuting components. Accepting the Newton–Wigner operator as a position operator, it would not be possible to interpret $|\psi^{\pm}(\boldsymbol{x})|^2$. But the mean position according to $\boldsymbol{X}^{\pm}$ is simply

$$(\psi^{\pm}, \boldsymbol{X}^{\pm}\psi^{\pm}) = \int \boldsymbol{x}|\psi^{\pm}(\boldsymbol{x})|^2 \, \mathrm{d}^3x \tag{2.19}$$

so that $|\psi^{\pm}(\boldsymbol{x})|^2$ may be interpreted as a position probability density. Both the Newton–Wigner operator and $\boldsymbol{X}^{\pm}$ are not parts of a four-vector. (Within Dirac's theory time is treated as a parameter and not as a variable like $\boldsymbol{x}$. It is only possible to obtain position four-vectors within proper-time theories, cf Horwitz and Lavie (1980).)

Angular momentum:

$$\begin{aligned} &\boldsymbol{J}^{\pm} = \boldsymbol{L}^{\pm} + \boldsymbol{S}^{\pm}, \qquad \boldsymbol{L}^{\pm} = \boldsymbol{X}^{\pm} \times \boldsymbol{P}^{\pm}, \\ &S_i^{\pm} = \pm\tfrac{1}{2}\varepsilon_{ijk} P_0^{\pm} \sigma^{jk} P_0^{\pm} = \pm\tfrac{1}{2}\varepsilon_{ijk}\left(\sigma^{jk} \pm \frac{\mathrm{i}\gamma^0(\gamma^j P^k - \gamma^k P^j)}{|H_0|}\right) P_0^{\pm}. \end{aligned} \tag{2.20}$$

Orbital and spin angular momentum are conserved separately (because of Zitterbewegung this is not the case for the operators $\boldsymbol{L}$ and $\boldsymbol{S}$).

We conclude with some remarks on the representations of symmetry transformations within the coherent subspaces $\mathscr{H}^{\pm}$. The commutation relation (2.15) requires for the representation of finite translations

$$\begin{aligned} &[\exp(-\mathrm{i}\boldsymbol{a}\cdot\boldsymbol{P}^{+})\psi^{+}](\boldsymbol{x}) = \psi^{+}(\boldsymbol{x}-\boldsymbol{a}) \qquad \text{for electrons,} \\ &[\exp(+\mathrm{i}\boldsymbol{a}\cdot\boldsymbol{P}^{-})\psi^{-}](\boldsymbol{x}) = \psi^{-}(\boldsymbol{x}-\boldsymbol{a}) \qquad \text{for positrons.} \end{aligned} \tag{2.21}$$

This is indeed very reasonable. Consider a unitary operator U leaving $\mathscr{H}^{\pm}$ invariant (e.g. finite Poincaré transformations). The measurable quantities concerned with U are transition probabilities. As was the case for expectation values of observables, one should perform a charge conjugation with the probabilities for transitions between negative energy states in order to obtain observable values for positrons:

$$|(\mathscr{C}\psi^{-}, U(-e)\mathscr{C}\varphi^{-})|^2 = |(\psi^{-}, \mathscr{C}^{-1}U(-e)\mathscr{C}\varphi^{-})|^2. \tag{2.22}$$

So, the symmetry transformation under consideration has to be represented by $U(e)P^{+}$ for electrons, and by $\mathscr{C}^{-1}U(-e)\mathscr{C}P^{-}$ for positrons. This would lead to observable transition probabilities. Because of

$$\mathscr{C}^{-1}\exp(-\mathrm{i}\boldsymbol{a}\cdot\boldsymbol{P})\mathscr{C}P_0^{-} = \exp(+\mathrm{i}\boldsymbol{a}\cdot\boldsymbol{P}^{-})P_0^{-}$$

this is in agreement with (2.21).

The interpretation presented here is not a one-particle theory, because we need electrons and positrons to interpret all $\mathscr{L}^2(\mathbb{R}^3)^4$ solutions (however, within $\mathscr{H}^{+(-)}$ the theory behaves perfectly as a one-particle theory of electrons (positrons)). It is also not identical with hole theory, where the positrons are 'holes' in the sea of negative energy solutions, because we use the negative energy solutions themselves to describe positrons. We need no quantum field formalism to establish a connection between negative

energy states and positrons. Our interpretation is in fact a slightly generalised version (to include bound states with both signs of energy) of the Stückelberg–Feynman theory. Its relation to the interpretation of positrons as 'electrons propagating backwards in time' will become clear in § 4.

It is possible to generalise our interpretation to potentials which do not fulfil (2.2) as long as there remains a gap in the spectrum of scattering energies which permits a separation into electron and positron states. In § 3 we consider a problem where this condition is not fulfilled.

3. Klein paradox

In § 2 we have specified a class of potentials for which the electron–positron interpretation may be applied. More general potentials would lead to a phenomenon usually called Klein's paradox. It is the purpose of this section to give a formulation of the problem on the basis of our 'two-particle interpretation'. It is not possible to solve this paradox within the (first quantised) Dirac theory, but our viewpoint is different from the usual one and leads to the consideration of new solutions (cf (3.18), (3.19)).

The best known example of Klein's paradox is a one-dimensional potential step higher than twice the rest mass of the particle:

$$V(x) = e\Phi(x) = \begin{cases} V_0, & x > a, \\ \text{monotone}, & -a \leqslant x \leqslant +a \quad (a \geqslant 0,\ V_0 > m), \\ -V_0, & x < -a, \end{cases} \tag{3.1}$$

(our choice of a 'symmetric step' will stress the symmetry between electron and positron states). Solutions for $V_0 < m$ may be found for example in Bongaarts and Ruijsenaars (1977); potential wells ($V(x)$ vanishes as $|x| \to \infty$) are considered in Dosch *et al* (1971). Some recent investigations using field theory indicate pair creation, cf Rafelski (1977), Aoyama and Kobayashi (1980), Hansen and Ravndal (1980).

The conditions (2.2) are not fulfilled by (3.1), because V does not vanish at infinity. Instead we have $\sigma(H) = \sigma_{ac}(H) = (-\infty, +\infty)$. The superselection rule of § 2.2 makes no sense: one cannot define the 'coherent subspaces' $\mathscr{H}^{\pm}(H)$, because there is no gap in the spectrum of scattering energies. A separation of electron and positron states will only be possible in the asymptotic region $|x| > a$, and may be performed as follows.

First note that a constant potential $V(x) \equiv V_0$ shifts the whole spectrum by V_0, but has no observable effect on the motion of particles (no force):

$$\sigma(H_0 + V_0) = (-\infty, -m + V_0] \cup [m + V_0, +\infty) \equiv \Sigma^-(H_0 + V_0) \cup \Sigma^+(H_0 + V_0). \tag{3.2}$$

As in § 2, we interpret $\Sigma^{\pm}$ as the electron resp. positron part of the spectrum. The step potential (3.1) is constant in the asymptotic region and it is possible to decide whether a given asymptotic state represents an electron or a positron. Throughout this section we are interested in wavepackets with energies in the interval $(m - V_0, -m + V_0)$. For $x \to +\infty$ they always represent positrons, $E \in (-\infty, -m + V_0] = \Sigma^-(H_0 + V_0)$, and for $x \to -\infty$ electrons, $E \in [m - V_0, \infty) = \Sigma^+(H_0 - V_0)$.

Now we turn to the solutions of the one-dimensional Dirac equation with constant potential:

$$\mathrm{i}\frac{\partial}{\partial t}\psi(x, t) = (H_0 + V)\psi(x, t) = \left(-\mathrm{i}\alpha\frac{\partial}{\partial x} + \beta m + V(x)\right)\psi(x, t). \tag{3.3}$$

For simplicity, we use a two-component form with 2×2 matrices α and β (we neglect spin effects since there is no spin-flip for step potentials, cf Bjorken and Drell (1964)):

$$\alpha=\alpha^*, \qquad \beta=\beta^*, \qquad \alpha^2=\beta^2=1, \qquad \alpha\beta+\beta\alpha=0, \tag{3.4}$$

for example

$$\alpha=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad \beta=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{3.5}$$

The operators for charge conjugation and time reversal are now given by

$$(\mathscr{C}\psi)(x,t) := \alpha\beta\overline{\psi(x,t)}, \qquad (\mathscr{T}\psi)(x,t) := \beta\overline{\psi(x,-t)}. \tag{3.6}$$

We further introduce eigenfunctions of the free Hamiltonian H_0:

$$\begin{aligned}&\psi_0^{(j)}(x,E) := \frac{1}{4\pi}\begin{pmatrix}\eta(E)\\ (-1)^{j+1}/\eta(E)\end{pmatrix}\exp(ip^{(j)}(E)x), \qquad j=1,2,\\ &\eta(E) := \left(\frac{E+m}{E-m}\right)^{1/4}, \qquad p^{(j)}(E) := (-1)^{j+1}\operatorname{sgn}(E)(E^2-m^2)^{1/2}.\end{aligned} \tag{3.7}$$

They form a complete orthonormal set and every electron resp. positron solution of (3.3) with constant potential $V(x)\equiv V_0$ is a linear combination of

$$f^{j\pm}(x,t)=\int_{\Sigma^\pm} dE\, \hat{f}^{(j)}(E)\psi_0^{(j)}(x,E-V_0)\,e^{-iEt}, \qquad j=1,2, \tag{3.8}$$

with $\int dE\,|\hat{f}^{(j)}(E)|^2<\infty$, and $\Sigma^\pm$ given by (3.2).

In the following we summarise some properties of the wavepackets (3.8) which will be useful later.

(1) With the normalisation of (3.7) the scalar product becomes

$$\begin{aligned}&(f^{j\pm}, g^{i\pm})=\delta_{ij}\int_{\Sigma^\pm} dE\,\overline{\hat{f}^{(j)}(E)}\hat{g}^{(i)}(E),\\ &(f^{j+}, g^{j-})=0.\end{aligned} \tag{3.9}$$

(2) f^{j+}, $j=1, 2$, always represents an electron. f^{j-}, $j=1, 2$, always represents a positron.

(3) The group velocity is given by (cf (2.17))

$$\left(f^{j\pm}, \frac{P^\pm}{|H_0|}f^{j\pm}\right)=(-1)^{j+1}\int_{\Sigma^\pm} dE\,|\hat{f}^{(j)}(E)|^2\frac{[(E-V_0)^2-m^2]^{1/2}}{|E-V_0|}. \tag{3.10}$$

$j=1$ always represents a particle moving from left to right, $j=2$ a particle moving from right to left.

(4) Time reversal is an invariance transformation for all static potentials. We have

$$(\mathscr{T}f^{\frac{1}{2}\pm})(x,t)=f^{\frac{2}{1}\pm}(x,t) \tag{3.11}$$

i.e. the direction of motion is reversed.

Now we are prepared for a discussion of the electrostatic step-potential (3.1). The physical situation is not changed qualitatively, if we concentrate for simplicity on the rectangular step $a=0$, which has the advantage that the problem can be solved exactly. The stationary solutions coincide with $\psi_0^{(j)}(x, E\pm V_0)$ in the region $x\gtrless 0$ and have to be continuous at $x=0$. A solution for the rectangular step with $V_0>m$ and $E\in$

$(m - V_0, -m + V_0)$ is given by

$$\psi_{\mathrm{I}}(x, E) = \begin{cases} \psi_0^{(1)}(x, E + V_0) + R(E)\psi_0^{(2)}(x, E + V_0), & x < 0, \\ \mathcal{T}(E)\psi_0^{(1)}(x, E - V_0), & x > 0, \end{cases} \tag{3.12}$$

(R(E) and $T(E)$ are often called reflection and transmission coefficients, see below) and with the help of time reversal $\mathcal{T}$ (i.e. interchanging the indices (1) and (2)) we obtain immediately another solution

$$\psi_{\mathrm{II}}(x, E) = (\mathcal{T}\psi_{\mathrm{I}})(x, E). \tag{3.13}$$

ψ_{I} is the solution orginally given by Klein (1929); Bjorken and Drell investigated ψ_{II} for the first time. A review of the different interpretations of these solutions by various authors can be found in Bongaarts and Ruijsenaars (1976). Requiring continuity at $x = 0$, we obtain explicitly

$$1 + R(E) = \frac{\eta(E - V_0)}{\eta(E + V_0)} T(E), \qquad 1 - R(E) = \frac{\eta(E + V_0)}{\eta(E - V_0)} T(E), \tag{3.14}$$

and therefore $R(E)$, $T(E)$ are real in our energy region. We always have $R^2(E) + T^2(E) = 1$, which is reasonable for reflection and transmission coefficients. Wavepackets may be formed according to

$$f_{\mathrm{I,II}}(x, t) = \int_{m - V_0}^{-m + V_0} \mathrm{d}E\, \hat{f}_{\mathrm{I,II}}(E)\psi_{\mathrm{I,II}}(x, E)\, \mathrm{e}^{-\mathrm{i}Et}. \tag{3.15}$$

The norm of these wavepackets is constant during the scattering process. Remembering that $j = 1$ resp. 2 corresponds to particles moving to the right resp. left, and that for the energy region under consideration the wavepackets in $x < 0$ are electrons, in $x > 0$ positrons (cf the properties (1)–(4) of wavepackets), we conclude the following. For $t \to -\infty$ the wavepacket $f_{\mathrm{I}}(x, t)$ (formed with the help of (3.12)) describes an incoming electron, moving to the right. For $t \to +\infty$ the wavepacket splits up into a reflected electron part and a transmitted positron part. $f_{\mathrm{II}}(x, t) = (Tf_{\mathrm{I}})(x, t)$ describes the time reversed situation, i.e. incoming electron and positron (from the left resp. right) and outgoing electron wavepacket, moving to the left. Both solutions are paradoxical in the same respect: the total charge is not conserved in time. The incoming electron may be detected as an outgoing positron on the right, as $t \to +\infty$.

If we exchange the roles of electrons and positrons, we obtain two other solutions which do not conserve charge: $E \in (m - V_0, -m + V_0)$,

$$\psi_{\mathrm{III}}(x, E) = \begin{cases} \tilde{T}(E)\psi_0^{(2)}(x, E + V_0), & x < 0, \\ \psi_0^{(2)}(x, E - V_0) + \tilde{R}(E)\psi_0^{(1)}(x, E - V_0), & x > 0, \end{cases} \tag{3.16}$$

$$\psi_{\mathrm{IV}}(x, E) = (\mathcal{T}\psi_{\mathrm{III}})(x, E). \tag{3.17}$$

Continuity at $x = 0$ implies $\tilde{T}(E) = T(E)$, $\tilde{R}(E) = -R(E)$. Equation (3.16) describes a positron, coming from the right and splitting up into a positron and electron part; equation (3.17) is the time reversed process.

We can use the four 'basic' solutions ψ_{I}–ψ_{IV} (only two of them are linearly independent) to construct charge conserving solutions by linear combination. There

exist two orthogonal eigenfunctions with the required property, namely:

$$\frac{1}{N}[\psi_{\mathrm{I}}(x,E)+\psi_{\mathrm{II}}(x,E)]=A\begin{cases}\psi_0^{(2)}(x,E+V_0)+\psi_0^{(1)}(x,E+V_0), & x<0,\\ B[\psi_0^{(2)}(x,E-V_0)+\psi_0^{(1)}(x,E-V_0)], & x>0,\end{cases} \tag{3.18}$$

and

$$\frac{1}{N}[\psi_{\mathrm{III}}(x,E)-\psi_{\mathrm{IV}}(x,E)]=A\begin{cases}B[\psi_0^{(2)}(x,E+V_0)-\psi_0^{(1)}(x,E+V_0)], & x<0,\\ \psi_0^{(2)}(x,E-V_0)-\psi_0^{(1)}(x,E-V_0), & x>0,\end{cases} \tag{3.19}$$

with

$$N:=[2R(E)(R(E)+1)]^{1/2},\qquad A:=[(R(E)+1)/2R(E)]^{1/2},\qquad B:=\frac{\eta(E+V_0)}{\eta(E-V_0)}.$$

The corresponding wavepackets with energies in $(m-V_0,-m+V_0)$ conserve total charge and consist of electron and positron parts being simultaneously totally reflected at the potential step. In the usual sense (i.e. that the step becomes transparent if it is higher than $2m$; cf Klein (1929)) there is no Klein paradox at all! In our language the paradox consists in the existence of solutions violating charge conservation. Note that within this interpretation it is not possible to obtain physically reasonable solutions corresponding to the situation where electrons alone were sent towards the step. The only charge conserving solutions describe electrons coming in from the left and positrons coming in from the right. With our normalisation constant N equation (3.18) yields wavepackets with charge e, equation (3.19) with $-e$. Note that the electron and positron parts cannot be treated independently.

Finally, let us summarise the results of this section. The rectangular step is important for the following reasons: equation (2.1) can be solved exactly and it shows what kinds of difficulties arise, if the condition (2.2) on the invariance of the essential spectrum is relaxed. For $V_0<m$ there are no difficulties because there still remains a gap between electron and positron states. For $V_0>m$ this gap vanishes, transitions from electron to positron states occur and consequently there are solutions which do not conserve total charge. The invariance with respect to time reversal (if ψ_{I} is a solution then $\mathcal{T}\psi_{\mathrm{I}}=\psi_{\mathrm{II}}$ is a solution) made it possible to find charge conserving solutions (e.g. $\psi_{\mathrm{I}}+\psi_{\mathrm{II}}$) in this case. Qualitatively, we then have the following situation. For $V_0<m$, total charge is always conserved. The solution with charge e in the energy interval $I\subset(m-V_0,V_0)$ consists of an electron wavepacket being totally reflected at the barrier, and (at the moment of scattering) an exponentially decreasing tail in the region $x>0$. Here we have the complete analogue to corresponding solutions of the Schrödinger equation. (For solutions in other energy regions cf Bongaarts and Ruijsenaars (1977).) Increasing V_0 beyond m and keeping I fixed, the exponential tail becomes smaller and smaller. To conserve total charge, we have to form a positron wavepacket in $x>0$. For sufficiently large $V_0>m$, the interval I lies in $(m-V_0,-m+V_0)$. Then the electron wavepacket is still totally reflected, but there is no exponentially decreasing part in $x>0$. Instead, we have a propagating positron, which is simultaneously totally reflected at the barrier. If the total charge is required to be e as before, the norm of the electron wavepacket now has to be greater than it was for $V_0<m$, because it has to compensate for the norm of the positron part. Roughly speaking, if $V_0>m$, an additional electron–positron pair is necessary to obtain a solution with the same conserved charge. This, however, is not the description of a creation process. For

$V_0 > m$ the additional pair has been there for all times, so we have only another static situation.

For time-dependent potentials the situation becomes even worse. If we consider a potential step increasing with time, then the unitarity of time evolution prevents any increase in the norm. Furthermore, there is no time reversal invariance that helps to construct charge conserving solutions. The general situation for time-dependent potentials, allowing a separation of asymptotic states into electron and positron solutions, is the following. The sum of the norms $\|\psi^+(t)\|^2 + \|\psi^-(t)\|^2$ remains constant, but not their difference $e(\|\psi^+(t)\|^2 - \|\psi^-(t)\|^2)$, i.e. the total charge.

4. Scattering theory

Assume that the conditions of § 2.1 are fulfilled, i.e. that we can separate the solutions with the help of orthogonal projections $P^{\pm}(H)$ into electron and positron solutions for all times. The time development of electrons and positrons is then given by the Dirac equation ((2.1) with $g = e$)

$$\exp(-\mathrm{i}H(e)t) = \exp(-\mathrm{i}H^+t)P^+ + \exp(+\mathrm{i}H^-t)P^-$$

($\exp(+\mathrm{i}H^-t)P^-$ is the positron time evolution according to (2.21)). If one formally applies (2.7) to obtain the positron time evolution, one obtains

$$[\exp(-\mathrm{i}H(e)t)]^- = P^-\mathscr{C}^{-1}\exp(+\mathrm{i}H(-e)t)\mathscr{C}P^- = \exp(-\mathrm{i}H^-t)P^-.$$

This is a formal contradiction. Obviously (2.7) can only be applied to yield the correct positron operators (not only observables, but also unitary transformations) if it is somehow possible to transfer the wrong sign to the time parameter, i.e. to reinterpret the solutions $\psi^- \in \mathscr{H}^-$. Feynman interpreted the negative energy solutions as electrons propagating backwards in time, thereby adopting the viewpoint of proper-time theories for the Dirac equation. We pick up this idea by using the 'Feynman operators' (4.1) instead of the unitary groups $\exp(-\mathrm{i}Ht)$ and $\exp(-\mathrm{i}H_0t)$ to describe the time evolution:

$$\begin{aligned} S^{\pm}(t) &:= \theta(t)\exp(-\mathrm{i}Ht)P^{\pm} - \theta(-t)\exp(-\mathrm{i}Ht)P^{\mp}, \\ S_0^{\pm}(t) &:= \theta(t)\exp(-\mathrm{i}H_0t)P_0^{\pm} - \theta(-t)\exp(-\mathrm{i}H_0t)P_0^{\mp}, \end{aligned} \tag{4.1}$$

with orthogonal projectors $P^{\pm}$ on $\mathscr{H}^{\pm}(H)$ and $P_0^{\pm}$ on $\mathscr{H}^{\pm}(H_0) \equiv \mathscr{H}_0^{\pm}$ (cf (2.5)). Note that the integral kernel of $\mathrm{i}S_0^+(t-t')\gamma^0$ in configuration space is just the Feynman propagator $S_{\mathrm{F}}(x-x')$ (notation of Bjorken and Drell (1964)), which is a fundamental solution of (2.1):

$$(\mathrm{i}\gamma^0\partial/\partial t - \gamma^0 H_0)S_{\mathrm{F}}(x-x') = (\mathrm{i}\gamma^\mu\partial_\mu - m)S_{\mathrm{F}}(x-x') = \mathrm{i}\delta^4(x-x'). \tag{4.2}$$

A solution of the Dirac equation for $t_1 < t < t_2$ is for example given by $S^+(t-t')\psi(t')$, if an initial condition for $(P^+\psi)(t')$ at $t' = t_1$ and a final condition for $(P^-\psi)(t')$ at $t' = t_2$ are imposed. A sharp distinction between electrons and positrons is made, taking into account that negative energy states are considered as electrons moving backwards in time. With the Feynman operators we obtain a description close to so-called proper-time theories (cf Horwitz and Lavie (1980)).

Next we show that we obtain a rigorous formulation of Feynman's approach to relativistic potential scattering by the following simple statement. Take the Feynman operators instead of the unitary groups $\exp(-\mathrm{i}Ht)$ and $\exp(-\mathrm{i}H_0t)$ when defining asymptotic states.

Definition. Let $\psi^{\text{in}}(t)$ and $\psi^{\text{out}}(t)$ be solutions of the free Dirac equation and $\psi(t)$ a solution of (2.1). ψ^{in} resp. ψ^{out} are called asymptotic states for ψ: $\leftrightarrow$

$$\lim_{t\to\mp\infty} \|S^{-}(t)\psi - S_0^{-}(t)\psi^{\text{in}}\| = 0 \tag{4.3}$$

resp.

$$\lim_{t\to\pm\infty} \|S^{+}(t)\psi - S_0^{+}(t)\psi^{\text{out}}\| = 0. \tag{4.4}$$

Incoming electrons and outgoing positrons are now called in-asymptotes, incoming positrons and outgoing electrons are called out-asymptotes (i.e. the terms 'in' and 'out' are to be understood in the sense of proper time). Inserting definition (4.1), one can show by standard procedures that

$$\psi(t) = \Omega^{\text{in}}\psi^{\text{in}}(t) = \Omega^{\text{out}}\psi^{\text{out}}(t) \tag{4.5}$$

with

$$\begin{aligned} \Omega^{\text{in}} &:= \Omega^{-}P_0^{+} + \Omega^{+}P_0^{-}, \\ \Omega^{\text{out}} &:= \Omega^{+}P_0^{+} + \Omega^{-}P_0^{-}, \end{aligned} \tag{4.6}$$

and where

$$\Omega^{\pm} := s\text{-}\lim_{t\to\pm\infty} \mathrm{e}^{\mathrm{i}Ht}\,\mathrm{e}^{-\mathrm{i}H_0 t} \tag{4.7}$$

are the Möller operators introduced by Prosser (1963). Therefore the problem of existence and completeness of $\Omega^{\text{in}}_{\text{out}}$ reduces to that of $\Omega^{\pm}$. A very powerful method of proving asymptotic completeness for non-relativistic scattering systems has been developed by Enss (1978, 1979) (an introductory survey is given in Enss (1980)). This method can also be applied to the Dirac equation, cf Simon (1979). Roughly speaking, it is only required that $H = H_0 + eV$ be self-adjoint and each component of V falls off faster than $1/|x|$ at infinity (conditions (2.2) are fulfilled in that case). Assuming asymptotic completeness for $\Omega^{\pm}$, we obtain the following. The isometric operators Ω^{in} and Ω^{out} exist and

$$R(\Omega^{\text{in}}) = R(\Omega^{\text{out}}) = \mathscr{H}_{\text{ac}}(H) \equiv P_{\text{ac}}\mathscr{L}^2(\mathbb{R}^3)^4 \tag{4.8}$$

(the range of $\Omega^{\text{in}}_{\text{out}}$ is the subspace of scattering states). The adjoint operators are given by

$$\begin{aligned} (\Omega^{\text{in}})^* &= P_0^{+}(\Omega^{-})^* + P_0^{-}(\Omega^{+})^*, \\ (\Omega^{\text{out}})^* &= P_0^{+}(\Omega^{+})^* + P_0^{-}(\Omega^{-})^*, \end{aligned} \tag{4.9}$$

with

$$(\Omega^{\pm})^* = s\text{-}\lim_{t\to\pm\infty} \mathrm{e}^{\mathrm{i}H_0 t}\,\mathrm{e}^{-\mathrm{i}Ht}P_{\text{ac}}. \tag{4.10}$$

All the Möller operators fulfil intertwining relations

$$f(H)\Omega = \Omega f(H_0), \qquad f(H_0)\Omega^* = \Omega^* f(H), \tag{4.11}$$

and, since $\Omega P_0^{\pm} = \Omega\theta(\pm H_0) = \theta(\pm H)\Omega = P^{\pm}\Omega$ $(P^{\pm} = P^{\pm}_{\text{ac}} = \theta(\pm H)$ on $\text{Ran}(\Omega) = H_{\text{ac}}(H))$, we have

$$S^{\pm}(t)\Omega = \Omega S_0^{\pm}(t), \qquad S_0^{\pm}(t)\Omega^* = \Omega^* S^{\pm}(t). \tag{4.12}$$

Writing the Ω's as integrals of their derivatives, e.g.

$$(\Omega^+)^* = P_{ac} - ie \int_0^\infty dt\, e^{iH_0 t} V e^{-iHt} P_{ac}$$

one can easily deduce the integral representations

$$(\Omega_{out}^{in})^* = P_{ac} + ie \int_{-\infty}^{+\infty} dt\, S_0^{\pm}(-t)\gamma^0\gamma^\mu A_\mu e^{-iHt} P_{ac}. \tag{4.13}$$

The representation of $(\Omega^{in})^*$ implies for any wavefunction $\psi(t) \in \mathcal{H}_{ac}(H)$:

$$\psi(x) = \psi^{in}(x) + e \int d^4y\, S_F(x-y)\gamma^\mu A_\mu(y)\psi(y), \tag{4.14}$$

$(x = (t, \boldsymbol{x}))$. Since $\psi(t) \xrightarrow[t\to\infty]{} (\psi^{out})^+(t) + (\psi^{in})^-(t)$ and $\psi(t) - \psi^{in}(t) \xrightarrow[t\to+\infty]{} (\psi^{out})^+(t) - (\psi^{in})^+(t)$, we see that

$$\lim_{t\to\pm\infty} \|P^{\mp}(\psi(t) - \psi^{in}(t))\| = 0. \tag{4.15}$$

The 'scattered wave' $\psi - \psi^{in}$ consists only of negative energies in the distant past and only of positive energies in the distant future. These are just the Feynman constraints. A comparison of (4.14) and (4.15) with formulae (6.53)–(6.55) of Bjorken and Drell (1964) shows that we have obtained a rigorous formulation of the basic principles of Feynman's propagator formalism for potential scattering.

We may now define the unitary scattering operator

$$S := (\Omega^{out})^* \Omega^{in} = \tilde{S} P_0^+ + \tilde{S}^{-1} P_0^- \tag{4.16}$$

where

$$\tilde{S} := (\Omega^+)^* \Omega^- = s - \lim_{t\to\infty} e^{iH_0 t} e^{-2iH(e)t} e^{iH_0 t} \equiv \tilde{S}(e) \tag{4.17}$$

is the unitary scattering operator of Prosser (1963).

Equation (4.16) describes the scattering process in terms of electrons moving backwards or forwards in time. The operator providing us with a description in terms of electrons and positrons is given via (2.7) by $P_0^+ S(e) P_0^+ = \tilde{S}(e) P_0^+$ for electrons, and by $P_0^- \mathscr{C}^{-1} S^*(-e) \mathscr{C} P_0^- = \tilde{S}(e) P_0^-$ for positrons. $\tilde{S} = \tilde{S} P_0^+ + \tilde{S} P_0^-$, the scattering operator of Prosser, yields the same result. It is in this sense that the two scattering formalisms are equivalent. Note, however, the 'physical' difference: S is a mapping from 'in' to 'out' (i.e. in the direction of proper time) and $\tilde{S}$ is a mapping from '$(t \to -\infty)$' to '$(t \to +\infty)$' (i.e. in the direction of ordinary time). S simulates the scattering operators of proper-time quantum mechanics (cf Horwitz and Lavie 1980).

The quantity $((\phi^{out})^+, S(\psi^{in})^-)$ should describe spontaneous pair creation (cf Bjorken and Drell 1964, figure 6.7): a particle with negative energy moving to the past is scattered to positive energies and travels back to the future. In the present case the intertwining relations (4.11) and (4.12) imply in the usual manner $[S, H_0] = 0$, i.e. conservation of energy during the scattering process. We find, that for static potentials there is no transition to states with different energy if the S operator exists and is unitary. So, if we initially have one electron (positron), we finally have one electron (positron), which is in accordance with the remarks at the end of § 2.2. Consequently, the above matrix element is zero and there is no creation process.

Now we turn to more general situations.

(1) Without modification our scattering formalism is not applicable to Coulomb forces: the S operator does not exist in that case. The reason for this is quite clear and is the same as in non-relativistic theories. The definition of $\tilde{S}$ via the Möller operators $\Omega(H, H_0)$ neglects the fact that Coulomb particles never become asymptotically free, because of the long range of the $1/r$ potential (this is also the reason for many divergencies, e.g. in Born expansions). Since the Coulomb potential satisfies the conditions of § 2 ($Z < 137$) the solutions may be separated into electrons and positrons and S may be obtained as in (4.16), if a Coulomb-modified $\tilde{S}$ is known. For the correct definition of $\tilde{S}$ for relativistic Coulomb scattering see Dollard and Velo (1966).

(2) An additional constant potential V_0 has no influence on the motion of particles. In the formalism one has only to replace the asymptotic evolution generated by H_0 by the asymptotic evolution generated by $H_0 + V_0$.

(3) Consider one-dimensional step potentials (3.1) with $V_0 < m$. There remains still a gap in the spectrum of scattering energies, and the solutions may be separated into electrons and positrons. If the different spatial asymptotics to the right and to the left is taken into account for the definition of $\tilde{S}$ (cf Bongaarts and Ruijsenaars 1977, Davies and Simon 1978) then S may be defined with little modification as in (4.16).

In examples (1)–(3) electron and positron states can be separated for all times. Quantum mechanical methods (like scattering theory) may be applied consistently for such potentials, and the different scattering formalisms of Feynman and Prosser are mathematically equivalent. From § 3 we know that for more general potentials our interpretation cannot be carried through in full generality, although it would often be possible to separate electron and positron states at least asymptotically. The Dirac equation may then have solutions violating the principle of charge conservation. In the following examples we sketch briefly how these facts are reflected by scattering theory.

(4) Consider step potentials with $V_0 > m$ (cf § 3). There is no gap between electron and positron states and (4.16) cannot be applied. Consider the solutions given in § 3. In our terminology the outgoing positron wavepackets moving to the right are in-states, incoming positrons are out-states. A short study of the various solutions shows that the total charge of in- and out-states is always the same (i.e. we have charge conservation in the sense of proper time), whereas they may have different norms (which is the case for $\psi_{\mathrm{I}}, \ldots, \psi_{\mathrm{IV}}$). We see that there is no unitary S operator (which would be norm conserving with respect to proper time), and the reason for this is the existence of solutions which do not conserve charge. On the other hand, the operator $\tilde{S}$ introduced by Bongaarts and Ruijsenaars (1977) is unitary, because it maps asymptotes for $t \to -\infty$ on asymptotes for $t \to +\infty$ and the norm conservation is guaranteed by the unitarity of time evolution. We therefore see that the Feynman scattering theory is very sensitive to situations which are 'not allowed'.

(5) Consider strongly time-dependent potentials $V(t)$, for which our interpretation may be applied to the asymptotic states (e.g. potentials vanishing sufficiently for $t \to \pm\infty$). If there are solutions which do not conserve total charge (cf the remarks at the end of § 3), we see from the example above that no unitary S can be found, even if $\tilde{S}$ may still be defined and is unitary.

So we have obtained the following results. For a wide class of potentials the Feynman theory of potential scattering can be formulated rigorously. As long as S, defined by (4.16), exists and is unitary, there are no unphysical solutions of the Dirac equation. For more general situations (see items (4) and (5) above) the Dirac equation has solutions which do not conserve total charge. This cannot be seen from the

properties of $\tilde{S}$, but is reflected by the fact that a unitary Feynman scattering-operator S does not exist. Pair creation is not described consistently by the Dirac equation within our interpretation, because the conservation of norm prevents any change in the number of particles.

5. Conclusions

We have shown that a consequent development of the Stückelberg–Feynman theory into a probabilistic interpretation of the Dirac equation (with external fields satisfying the conditions of § 2) corrects some prejudices concerning negative energy states, Zitterbewegung, and bound states in repulsive potentials, and yields the connection between propagator theory and 'usual' scattering theory. However, an application of our interpretation to more general potentials leads to difficulties, because the Dirac equation then has physically unrealistic solutions which do not conserve charge (which is clearly reflected by Feynman's scattering theory). So, we have also obtained the limits of the Dirac equation, considered as a wave mechanical equation. The range of applicability of this theory is being left, whenever pair creation or annihilation actually happens. This cannot be described by the Dirac equation, because the norm (and not the charge) of wavepackets is always constant in time. The situation is completely different within proper-time quantum mechanics (Horwitz and Lavie (1980); a theory for particles with spin is given in Piron and Reuse (1978)), where the norm is conserved with respect to proper time, but may change with ordinary time (this has, of course, consequences for the description of the Klein paradox, cf Thaller (1981)). The coupled Maxwell–Dirac equations should be able to describe interactions between electron and positron states (as in QED); cf Glassey and Strauß (1979). It is then necessary to consider the behaviour of $\{\psi, A^\mu\}$ with the help of a formalism that conserves the total energy (and not the norm $\|\psi\|$ of particles). $\|\psi\|$ need not be constant, because the electromagnetic field A^μ may take over a part of the energy of the particles. We suppose that in such a framework it should be possible to describe pair creation and annihilation in terms of ordinary quantum mechanics (without using quantum field theory).

Acknowledgments

I would like to thank F Gesztesy, L Pittner and P Vogl for many stimulating discussions and critical remarks. I am also grateful to Professor H Mitter for giving me the opportunity to perform this work.

References

Aoyama H and Kobayashi M 1980 *Prog. Theor. Phys.* **64** 1045–57
Barut A O and Bracken A J 1981 *Phys. Rev.* D **23** 2454–63
Bjorken J D and Drell S D 1964 *Relativistic Quantum Mechanics* (New York: McGraw-Hill)
Bongaarts P J M and Ruijsenaars S N M 1976 *Ann. Phys., NY* **101** 289–318
—— 1977 *Ann. Inst. Henri Poincaré* **A26** 1–17
Brysk H and Zweifel P F 1981 *Phys. Rev.* C **23** 968–72
Davies E B and Simon B 1978 *Commun. Math. Phys.* **63** 277–301

Dollard J and Velo G 1966 *Nuovo Cimento* **A45** 801–12
Dosch H G, Jensen J H D and Müller V F 1971 *Phys. Norwegia* **5** 151–62
Eckhardt K J 1974 *Math. Z.* **139** 105–31
Enss V 1978 *Commun. Math. Phys.* **61** 285–91
—— 1979 *Ann. Phys., NY* **119** 117–32
—— 1980 in *Rigorous Atomic and Molecular Physics, Proc. Erice School in Mathematical Physics 1980* ed G Velo and A S Wightman (New York: Plenum)
Feynman R P 1949 *Phys. Rev.* **76** 749–59
Foldy L L and Wouthuysen S A 1950 *Phys. Rev.* **78** 29–36
Glassey R T and Strauß W A 1979 *J. Math. Phys.* **20** 454–8
Hansen A and Ravndal F 1980 *Phys. Scr.* **23** 1036–42
Horwitz L P and Lavie Y 1980 *Tel Aviv University Preprint TAUP 833-80*
Jörgens K 1973 in *Lecture Notes in Math.* (Berlin: Springer) **280** 87–102
Kalf H 1976 *J. Funct. Anal.* **21** 389–96
—— 1980 *TH Darmstadt Preprint No 549*
Klein O 1929 *Z. Phys.* **53** 157–65
Klaus M 1980 *Helv. Phys. Acta* **53** 453–62
Klaus M and Wüst R 1979 *Commun. Math. Phys.* **64** 171–6
Landgren J J and Reito P A 1979 *J. Math. Phys.* **20** 2204–11
Landgren J J , Reito P A and Klaus M 1980 *J. Math. Phys.* **21** 1210–7
Lock J A 1979 *Am. J. Phys.* **47** 797–802
Messiah A 1962 *Quantum Mechanics* vol 2 (Amsterdam: North-Holland)
Nenciu G 1976 *Commun. Math. Phys.* **48** 235–47
Newton T D and Wigner E P 1949 *Rev. Mod. Phys.* **21** 400–6
Pearson D P 1977 *Commun. Math. Phys.* **57** 117–33
Pilkuhn H M 1979 *Relativistic Particle Physics* (New York: Springer)
Piron C and Reuse F 1978 *Helv. Phys. Acta* **51** 146–56
Prosser R T 1963 *J. Math. Phys.* **48** 235–47
Pryce M H L 1948 *Proc. R. Soc.* A **195** 62–81
Rafelski J 1977 *Ref. TH. 2414 CERN* (Lectures presented at the Advanced Study Institute on Non-linear Problems in Physics and Mathematics, Istanbul)
Rafelski J, Fulcher L P and Klein A 1978 *Phys. Rep.* **38** 227–361
Scadron M D 1979 *Advanced Quantum Theory* (New York: Springer)
Simon B 1979 *Duke Math. J.* **46** 119–68
Sokolov A A and Ternov I M 1968 *Synchrotron Radiation* (Berlin: Akademie)
Streater R F and Wightman A S 1964 *PCT, Spin and Statistics, and all that* (New York: Benjamin)
Stückelberg E C C 1942 *Helv. Phys. Acta* **15** 23–37
Thaller B 1981 *Lett. Nuovo Cimento* **31** 439–44
Weder R A 1973 *Ann. Soc. Sci. Bruxelles* **87** 341–55
Weidmann J 1971 *Math. Z.* **119** 349–73
Wüst R 1977 *Math. Z.* **152** 259–71