From: Bjorken + Drell,
Relativistic Quantum Mechanics
McGraw Hill, 1964

1.1 Formulation of a Relativistic Quantum Theory

Since the principles of special relativity are generally accepted at this time, a correct quantum theory should satisfy the requirement of relativity: laws of motion valid in one inertial system must be true in all inertial systems. Stated mathematically, relativistic quantum theory must be formulated in a Lorentz covariant form.

In making the transition from nonrelativistic to relativistic quantum mechanics, we shall endeavor to retain the principles underlying the nonrelativistic theory. We review them briefly:[1]

1. For a given physical system there exists a state function Φ that summarizes all that we can know about the system. In our initial development of the relativistic one-particle theory, we usually deal directly with a coordinate realization of the state function, the wave function $\psi(q_i \cdots, s_i \cdots, t)$. $\psi(q,s,t)$ is a complex function of all the classical degrees of freedom, $q_1 \cdots q_n$, of the time t and of any additional degrees of freedom, such as spin s_i, which are intrinsically quantum-mechanical. The wave function has no direct physical interpretation; however, $|\psi(q_1 \cdots q_n, s_1 \cdots s_n, t)|^2 \geq 0$ is interpreted as the probability of the system having values $(q_1 \cdots s_n)$ at time t. Evidently this probability interpretation requires that the sum of positive contributions $|\psi|^2$ for all values of $q_1 \cdots s_n$ at time t be finite for all physically acceptable wave functions ψ.

2. Every physical observable is represented by a linear hermitian operator. In particular, for the canonical momentum p_i the operator correspondence in a coordinate realization is

$$p_i \rightarrow \frac{\hbar}{i}\frac{\partial}{\partial q_i}$$

def'n, not postulate

3. A physical system is in an eigenstate of the operator Ω if

$$\Omega\Phi_n = \omega_n\Phi_n \tag{1.1}$$

where Φ_n is the nth eigenstate corresponding to the eigenvalue ω_n. For a hermitian operator, ω_n is real. In a coordinate realization the equation corresponding to (1.1) is

$$\Omega(q,s,t)\psi_n(q,s,t) = \omega_n\psi_n(q,s,t)$$

[1] See, for example, W. Pauli, "Handbuch der Physik," 2d ed., vol. 24, p. 1, J. Springer, Berlin, 1933. L. I. Schiff, "Quantum Mechanics," 2d ed., McGraw-Hill Book Company, Inc., New York, 1955. P. A. M. Dirac, "The Principles of Quantum Mechanics," 4th ed., Oxford University Press, London, 1958.

4. The expansion postulate states that an arbitrary wave function, or state function, for a physical system can be expanded in a complete orthonormal set of eigenfunctions ψ_n of a complete set of commuting operators (Ω_n). We write, then,

$$\psi = \sum_n a_n\psi_n$$

where the statement of orthonormality is

$$\sum_s \int (dq_1 \cdots)\psi_n^*(q_1 \cdots, s \cdots, t)\psi_m(q_1 \cdots, s \cdots, t) = \delta_{nm}$$

$|a_n|^2$ records the probability that the system is in the nth eigenstate.

5. The result of a measurement of a physical observable is any one of its eigenvalues. In particular, for a physical system described by the wave function $\psi = \Sigma a_n\psi_n$, with $\Omega\psi_n = \omega_n\psi_n$, measurement of a physical observable Ω results in the eigenvalue ω_n with a probability $|a_n|^2$. The average of many measurements of the observable Ω on identically prepared systems is given by

$$\langle\Omega\rangle_\psi \equiv \sum_s \int \psi^*(q_1 \cdots, s \cdots, t)\Omega\psi(q_1 \cdots, s \cdots, t)(dq_1 \cdots)$$
$$= \sum_n |a_n|^2\omega_n$$

6. The time development of a physical system is expressed by the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi \tag{1.2}$$

where the hamiltonian H is a linear hermitian operator. It has no explicit time dependence for a closed physical system, that is, $\partial H/\partial t = 0$, in which case its eigenvalues are the possible stationary states of the system. A superposition principle follows from the linearity of H and a statement of conservation of probability from the hermitian property of H:

$$\frac{d}{dt}\sum_s \int \psi^*\psi(dq_1 \cdots) = \frac{i}{\hbar}\sum_s \int (dq_1 \cdots)[(H\psi)^*\psi - \psi^*(H\psi)]$$
$$= 0 \tag{1.3}$$

We strive to maintain these familiar six principles as underpinnings of a relativistic quantum theory.

1.2 Early Attempts

The simplest physical system is that of an isolated free particle, for which the nonrelativistic hamiltonian is

$$H = \frac{p^2}{2m} \tag{1.4}$$

The transition to quantum mechanics is achieved with the transcription

$$H \to i\hbar \frac{\partial}{\partial t} \tag{1.5}$$

$$\mathbf{p} \to \frac{\hbar}{i} \nabla$$

which leads to the nonrelativistic Schrödinger equation

$$i\hbar \frac{\partial\psi(q,t)}{\partial t} = \frac{-\hbar^2\nabla^2}{2m} \psi(q,t) \tag{1.6}$$

Equations (1.4) and (1.6) are noncovariant and therefore unsatisfactory. The left- and right-hand sides transform differently under Lorentz transformations. According to the theory of special relativity, the total energy E and momenta (p_x,p_y,p_z) transform as components of a contravariant four-vector

$$p^\mu = (p^0,p^1,p^2,p^3) = \left(\frac{E}{c}, p_x,p_y,p_z\right)$$

of invariant length

$$\sum_{\mu=0}^{3} p_\mu p^\mu \equiv p_\mu p^\mu = \frac{E^2}{c^2} - \mathbf{p}\cdot\mathbf{p} \equiv m^2c^2 \tag{1.7}$$

m is the rest mass of the particle and c the velocity of light in vacuo. The covariant notation used throughout this book is discussed in more detail in Appendix A. Here we only note that the operator transcription (1.5) is Lorentz covariant, since it is a correspondence between two contravariant four-vectors[1] $p^\mu \to i\hbar\, \partial/\partial x_\mu$.

Following this it is natural to take as the hamiltonian of a relativistic free particle

$$H = \sqrt{p^2c^2 + m^2c^4} \tag{1.8}$$

[1] We define $x^\mu = (ct,\mathbf{x})$ and $\nabla^\mu = \partial/\partial x_\mu$.

and to write for a relativistic quantum analogue of (1.6)

$$i\hbar \frac{\partial\psi}{\partial t} = \sqrt{-\hbar^2c^2\nabla^2 + m^2c^4}\, \psi \tag{1.9}$$

Immediately we are faced with the problem of interpreting the square-root operator on the right in Eq. (1.9). If we expand it, we obtain an equation containing all powers of the derivative operator and thereby a nonlocal theory. Such theories are very difficult to handle and present an unattractive version of the Schrödinger equation in which the space and time coordinates appear in unsymmetrical form.

In the interest of mathematical simplicity (though perhaps with a lack of complete physical cogency) we remove the square-root operator in (1.9), writing

$$H^2 = p^2c^2 + m^2c^4 \tag{1.10}$$

Equivalently, iterating (1.9) and using the fact that[1] if $[A,B] = 0$, $A\psi = B\psi$ implies $A^2\psi = B^2\psi$, we have

$$-\hbar^2 \frac{\partial^2}{\partial t^2} \psi = (-\hbar^2\nabla^2c^2 + m^2c^4)\psi$$

This is recognized as the classical wave equation

$$\left[\Box + \left(\frac{mc}{\hbar}\right)^2\right]\psi = 0$$

where

$$\Box \equiv \frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x^\mu} \tag{1.11}$$

Before looking further into (1.11), we note first that in squaring the energy relation we have introduced an extraneous negative-energy root

$$H = -\sqrt{p^2c^2 + m^2c^4}$$

In order to gain a simple equation, we have sacrificed positive definite energy and introduced the difficulty of "extra" negative-energy solutions. This difficulty is eventually surmounted (as we shall study in Chap. 5), and the negative-energy solutions prove capable of physical interpretation. In particular, they are associated with antiparticles. and the existence of antiparticles in nature lends strong experimental support for this procedure. So let us for a moment consider Eq. (1.10) and the inferred wave equation (1.11). Our first task is to construct a conserved current, since (1.11) is a second-order

[1] Throughout, we use the notation $[A,B] \equiv AB - BA$ for commutator brackets and $\{A,B\} \equiv AB + BA$ for anticommutator brackets.

wave equation and is altered from the Schrödinger form (1.2) upon which the probability interpretation in the nonrelativistic theory is based. This we do in analogy with the Schrödinger equation, taking ψ^* times (1.11), ψ times the complex conjugate equation, and subtracting:

$$\psi^*\left[\Box + \left(\frac{mc}{\hbar}\right)^2\right]\psi - \psi\left[\Box + \left(\frac{mc}{\hbar}\right)^2\right]\psi^* = 0$$

$$\nabla^\mu(\psi^*\nabla_\mu\psi - \psi\nabla_\mu\psi^*) = 0$$

or

$$\frac{\partial}{\partial t}\left[\frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right)\right] + \operatorname{div}\frac{\hbar}{2im}[\psi^*(\nabla\psi) - \psi(\nabla\psi^*)] = 0 \qquad (1.12)$$

We would like to interpret $(i\hbar/2mc^2)\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right)$ as a probability density ρ. However, this is impossible, since it is not a positive definite expression. For this reason we follow the path of history[1] and temporarily discard Eq. (1.11) in the hope of finding an equation of first order in the time derivative which admits a straightforward probability interpretation as in the Schrödinger case. We shall return to (1.11), however. Although we shall find a first-order equation, it still proves impossible to retain a positive definite probability density for a single particle while at the same time providing a physical interpretation of the negative-energy root of (1.10). Therefore Eq. (1.11), also referred to frequently as the Klein-Gordon equation, remains an equally strong candidate for a relativistic quantum mechanics as the one which we now discuss.

1.3 The Dirac Equation

We follow the historic path taken in 1928 by Dirac[2] in seeking a relativistically covariant equation of the form (1.2) with positive definite probability density. Since such an equation is linear in the time derivative, it is natural to attempt to form a hamiltonian linear in the space derivatives as well. Such an equation might assume a form

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{\hbar c}{i}\left(\alpha_1\frac{\partial\psi}{\partial x^1} + \alpha_2\frac{\partial\psi}{\partial x^2} + \alpha_3\frac{\partial\psi}{\partial x^3}\right) + \beta mc^2\psi \equiv H\psi \qquad (1.13)$$

[1] E. Schrödinger, *Ann. Physik*, **81**, 109 (1926); W. Gordon, *Z. Physik*, **40**, 117 (1926); O. Klein, *Z. Physik*, **41**, 407 (1927).

[2] P. A. M. Dirac, *Proc. Roy. Soc. (London)*, **A117**, 610 (1928); *ibid.*, **A118**, 351 (1928); "The Principles of Quantum Mechanics," *op. cit.*

The coefficients α_i here cannot simply be numbers, since the equatio would not be invariant even under a spatial rotation. Also, if w wish to proceed at this point within the framework stated in Sec. 1.1 the wave function ψ cannot be a simple scalar. In fact, the proba bility density $\rho = \psi^*\psi$ should be the time component of a conserve four-vector if its integral over all space, at fixed t, is to be an invarian

To free (1.13) from these limitations, Dirac proposed that it b considered as a matrix equation. The wave function ψ, in analog with the spin wave functions of nonrelativistic quantum mechanic is written as a column matrix with N components

$$\psi = \begin{bmatrix}\psi_1\\ \cdot\\ \cdot\\ \cdot\\ \psi_N\end{bmatrix}$$

and the constant coefficients α_i, β are $N \times N$ matrices. In effe then, Eq. (1.13) is replaced by N coupled first-order equations

$$i\hbar\frac{\partial\psi_\sigma}{\partial t} = \frac{\hbar c}{i}\sum_{\tau=1}^{N}\left(\alpha_1\frac{\partial}{\partial x^1} + \alpha_2\frac{\partial}{\partial x^2} + \alpha_3\frac{\partial}{\partial x^3}\right)_{\sigma\tau}\psi_\tau + \sum_{\tau=1}^{N}\beta_{\sigma\tau}mc^2\psi_\tau$$
$$= \sum_{\tau=1}^{N}H_{\sigma\tau}\psi_\tau \qquad (1.1$$

Hereafter we adopt matrix notation and drop summation indic in which case Eq. (1.14) appears as (1.13), to be now interpreted a matrix equation.

If this equation is to serve as a satisfactory point of departu first, it must give the correct energy-momentum relation

$$E^2 = p^2c^2 + m^2c^4$$

for a free particle, second, it must allow a continuity equation an probability interpretation for the wave function ψ, and third, it m be Lorentz covariant. We now discuss the first two of th requirements.

In order that the correct energy-momentum relation emerge fr Eq. (1.13), each component ψ_σ of ψ must satisfy the Klein-Gor second-order equation, or

$$-\hbar^2\frac{\partial^2\psi_\sigma}{\partial t^2} = (-\hbar^2c^2\nabla^2 + m^2c^4)\psi_\sigma \qquad (1.$$

Iterating Eq. (1.13), we find

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = -\hbar^2 c^2 \sum_{i,j=1}^{3} \frac{\alpha_j \alpha_i + \alpha_i \alpha_j}{2} \frac{\partial^2 \psi}{\partial x^i \, \partial x^j} + \frac{\hbar m c^3}{i} \sum_{i=1}^{3} (\alpha_i \beta + \beta \alpha_i) \frac{\partial \psi}{\partial x^i} + \beta^2 m^2 c^4 \psi$$

We may resurrect (1.15) if the four matrices α_i, β obey the algebra:

$$\begin{aligned} \alpha_i \alpha_k + \alpha_k \alpha_i &= 2\delta_{ik} \\ \alpha_i \beta + \beta \alpha_i &= 0 \\ \alpha_i^2 = \beta^2 &= 1 \end{aligned} \tag{1.16}$$

What other properties do we require of these four matrices α_i, β, and can we explicitly construct them? The α_i and β must be hermitian matrices in order that the hamiltonian $H_{\sigma\tau}$ in (1.14) be a hermitian operator as desired according to the postulates of Sec. 1.1. Since, by (1.16), $\alpha_i^2 = \beta^2 = 1$, the eigenvalues of α_i and β are ± 1. Also, it follows from their anticommutation properties that the trace, that is, the sum of the diagonal elements, of each α_i and β is zero. For example,

$$\alpha_i = -\beta \alpha_i \beta$$

and by the cyclic property of the trace

$$\operatorname{Tr} AB = \operatorname{Tr} BA$$

one has

$$\operatorname{Tr} \alpha_i = + \operatorname{Tr} \beta^2 \alpha_i = + \operatorname{Tr} \beta \alpha_i \beta = - \operatorname{Tr} \alpha_i = 0$$

Since the trace is just the sum of eigenvalues, the number of positive and negative eigenvalues ± 1 must be equal, and the α_i and β must therefore be even-dimensional matrices. The smallest even dimension, $N = 2$, is ruled out, since it can accommodate only the three mutually anticommuting Pauli matrices σ_i plus a unit matrix. The smallest dimension in which the α_i and β can be realized is $N = 4$, and that is the case we shall study. In a particular explicit representation the matrices are

$$\alpha_i = \begin{bmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{bmatrix} \qquad \beta = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1.17}$$

where the σ_i are the familiar 2×2 Pauli matrices and the unit entries in β stand for 2×2 unit matrices.

To construct the differential law of current conservation, we first introduce the hermitian conjugate wave functions $\psi^\dagger = (\psi_1^* \cdots \psi_4^*)$

and left-multiply (1.13) by $\psi^\dagger$:

$$i\hbar \psi^\dagger \frac{\partial \psi}{\partial t} = \frac{\hbar c}{i} \sum_{k=1}^{3} \psi^\dagger \alpha_k \frac{\partial}{\partial x^k} \psi + m c^2 \psi^\dagger \beta \psi \tag{1.18}$$

Next we form the hermitian conjugate of (1.13) and right-multiply by ψ:

$$-i\hbar \frac{\partial \psi^\dagger}{\partial t} \psi = -\frac{\hbar c}{i} \sum_{k=1}^{3} \frac{\partial \psi^\dagger}{\partial x^k} \alpha_k \psi + m c^2 \psi^\dagger \beta \psi \tag{1.19}$$

where $\alpha_i^\dagger = \alpha_i$, $\beta_i^\dagger = \beta$. Subtracting (1.19) from (1.18), we find

$$i\hbar \frac{\partial}{\partial t} \psi^\dagger \psi = \sum_{k=1}^{3} \frac{\hbar c}{i} \frac{\partial}{\partial x^k} (\psi^\dagger \alpha^k \psi)$$

or

$$\frac{\partial}{\partial t} \rho + \operatorname{div} \mathbf{j} = 0 \tag{1.20}$$

where we make the identification of probability density

$$\rho = \psi^\dagger \psi = \sum_{\sigma=1}^{4} \psi_\sigma^* \psi_\sigma \tag{1.21}$$

and of a probability current with three components

$$j^k = c \psi^\dagger \alpha^k \psi \tag{1.22}$$

Integrating (1.20) over all space and using Green's theorem, we find

$$\frac{\partial}{\partial t} \int d^3x \, \psi^\dagger \psi = 0 \tag{1.23}$$

which encourages the tentative interpretation of $\rho = \psi^\dagger \psi$ as a positive definite probability density.

The notation (1.20) anticipates that the probability current $\mathbf{j}$ forms a vector if (1.22) is to be invariant under three-dimensional space rotations. We must actually show much more than this. The density and current in (1.20) must form a four-vector under Lorentz transformations in order to ensure the covariance of the continuity equation and of the probability interpretation. Also, the Dirac equation (1.13) must be shown to be Lorentz covariant before we may regard it as satisfactory.

1.4 Nonrelativistic Correspondence

Before delving into the problem of establishing Lorentz invariance of the Dirac theory, it is perhaps more urgent to see first that the equation makes sense physically.

We may start simply by considering a free electron and counting the number of solutions corresponding to an electron at rest. Equation (1.13) then reduces to

$$i\hbar \frac{\partial \psi}{\partial t} = \beta m c^2 \psi$$

since the de Broglie wavelength is infinitely large and the wave function is uniform over all space. In the specific representation of Eq. (1.17) for β, we can write down by inspection four solutions:

$$\psi^1 = e^{-(imc^2/\hbar)t}\begin{bmatrix}1\\0\\0\\0\end{bmatrix} \qquad \psi^2 = e^{-(imc^2/\hbar)t}\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$
$$\psi^3 = e^{+(imc^2/\hbar)t}\begin{bmatrix}0\\0\\1\\0\end{bmatrix} \qquad \psi^4 = e^{+(imc^2/\hbar)t}\begin{bmatrix}0\\0\\0\\1\end{bmatrix} \tag{1.24}$$

the first two of which correspond to positive energy, and the second two to negative energy. The extraneous negative-energy solutions which result from the quadratic form of $H^2 = p^2c^2 + m^2c^4$ are a major difficulty, but one for which the resolution leads to an important triumph in the form of antiparticles. We come to this point in Chap. 5. Here we confine ourselves to the "acceptable" positive-energy solutions. In particular, we wish to show that they have a sensible nonrelativistic reduction to the two-component Pauli spin theory. To this end we introduce an interaction with an external electromagnetic field described by a four-potential

$$A^\mu : (\Phi, \mathbf{A})$$

The coupling is most simply introduced by means of the gauge-invariant substitution

$$p^\mu \to p^\mu - \frac{e}{c}A^\mu \tag{1.25}$$

 thus, for an electron, e = -|e|

made in classical relativistic mechanics to describe the interactio point charge e with an applied field. In the present case

$$p^\mu \to i\hbar\, \partial/\partial x_\mu \equiv \mathsf{p}^\mu$$

according to (1.5), and (1.25) takes the Dirac equation (1.13) t

$$i\hbar \frac{\partial \psi}{\partial t} = \left(c\boldsymbol{\alpha}\cdot\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right) + \beta mc^2 + e\Phi\right)\psi$$

Equation (1.26) expresses the "minimal" interaction of a particle, considered to be a point charge, with an applied e magnetic field. To emphasize its classical parallel, we write in $H = H_0 + H'$, with $H' = -e\boldsymbol{\alpha}\cdot\mathbf{A} + e\Phi$. The matrix $c\boldsymbol{\alpha}$ a here as the operator transcription of the velocity operator classical expression for the interaction energy of a point charge

$$H'_{\text{classical}} = -\frac{e}{c}\mathbf{v}\cdot\mathbf{A} + e\Phi$$

This operator correspondence $\mathbf{v}_{\text{op}} = c\boldsymbol{\alpha}$ is again evident in Eq. for the probability current. It also follows if we make the relat extension of the Ehrenfest relations:[1]

$$\frac{d}{dt}\mathbf{r} = \frac{i}{\hbar}[H,\mathbf{r}] = c\boldsymbol{\alpha} \equiv \mathbf{v}_{\text{op}}$$

and

$$\frac{d}{dt}(\boldsymbol{\pi}) = \frac{i}{\hbar}[H,\boldsymbol{\pi}] - \frac{e}{c}\frac{\partial}{\partial t}\mathbf{A}$$
$$\frac{d}{dt}(\boldsymbol{\pi}) = e\left[\mathbf{E} + \frac{1}{c}\mathbf{v}_{\text{op}} \times \mathbf{B}\right]$$

with $\boldsymbol{\pi} \equiv \mathbf{p} - (e/c)\mathbf{A}$ the operator corresponding to the l momentum and

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \nabla\Phi \qquad \text{and} \qquad \mathbf{B} = \text{curl}\,\mathbf{A}$$

the field strengths. Equation (1.27) is the operator equat motion for a point charge e. More general couplings in (1.26) lead to specific dipole and higher multipole terms in analogy wi classical development.

In taking the nonrelativistic limit of Eq. (1.26), it is conv to work in the specific representation of Eq. (1.17) and to expre

[1] Pauli, Schiff, and Dirac, *op. cit.*

wave function in terms of two-component column matrices $\tilde{\varphi}$ and $\tilde{\chi}$:

$$\psi = \begin{bmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{bmatrix} \tag{1.28}$$

We then obtain for (1.26)

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{bmatrix} = c\boldsymbol{\sigma} \cdot \boldsymbol{\pi} \begin{bmatrix} \tilde{\chi} \\ \tilde{\varphi} \end{bmatrix} + e\Phi \begin{bmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{bmatrix} + mc^2 \begin{bmatrix} \tilde{\varphi} \\ -\tilde{\chi} \end{bmatrix}$$

In the nonrelativistic limit the rest energy mc^2 is the largest energy in the problem and we write

$$\begin{bmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{bmatrix} = e^{-(imc^2/\hbar)t} \begin{bmatrix} \varphi \\ \chi \end{bmatrix} \tag{1.29}$$

where now φ and χ are relatively slowly varying functions of time which are solutions of the coupled equations

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \varphi \\ \chi \end{bmatrix} = c\boldsymbol{\sigma} \cdot \boldsymbol{\pi} \begin{bmatrix} \chi \\ \varphi \end{bmatrix} + e\Phi \begin{bmatrix} \varphi \\ \chi \end{bmatrix} - 2mc^2 \begin{bmatrix} 0 \\ \chi \end{bmatrix} \tag{1.30}$$

The second of Eqs. (1.30) may be approximated, for kinetic energies and field interaction energies small in comparison with mc^2, to

dropping ... and $e\Phi$ term in (1.30)

$$\chi = \frac{\boldsymbol{\sigma} \cdot \boldsymbol{\pi}}{2mc} \varphi \tag{1.31}$$

$\Rightarrow \varphi >> \chi$ ← large number

Equation (1.31) reveals χ as the "small" components of the wave function ψ in comparison with the "large" components φ. Relative to φ, χ is reduced by $\sim v/c \ll 1$ in the nonrelativistic approximation. Inserting (1.31) into the first of Eqs. (1.30), we obtain a two-component spinor equation

$$i\hbar \frac{\partial \varphi}{\partial t} = \left(\frac{\boldsymbol{\sigma} \cdot \boldsymbol{\pi}\, \boldsymbol{\sigma} \cdot \boldsymbol{\pi}}{2m} + e\Phi \right) \varphi \tag{1.32}$$

This is further reduced by the identity for Pauli spin matrices

$$\boldsymbol{\sigma} \cdot \mathbf{a}\, \boldsymbol{\sigma} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{b} + i\boldsymbol{\sigma} \cdot \mathbf{a} \times \mathbf{b}$$

(follows from: $\sigma_i \sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$)

or, here,

$$\boldsymbol{\sigma} \cdot \boldsymbol{\pi}\, \boldsymbol{\sigma} \cdot \boldsymbol{\pi} = \boldsymbol{\pi}^2 + i\boldsymbol{\sigma} \cdot \boldsymbol{\pi} \times \boldsymbol{\pi}$$
$$= \boldsymbol{\pi}^2 - \frac{e\hbar}{c} \boldsymbol{\sigma} \cdot \mathbf{B} \tag{1.33}$$

Then we have

$$i\hbar \frac{\partial \varphi}{\partial t} = \left[\frac{(\mathbf{p} - (e/c)\mathbf{A})^2}{2m} - \frac{e\hbar}{2mc} \boldsymbol{\sigma} \cdot \mathbf{B} + e\Phi \right] \varphi \tag{1.34}$$

which is recognized[1] as the Pauli equation. Equation (1.34) gives us confidence that we are on the right track in accepting Eqs. (1.13) and (1.26) as a starting point in constructing a relativistic electron theory. The two components of φ suffice to accommodate the two spin degrees of freedom of a spin one-half electron; and the correct magnetic moment of the electron, corresponding to the gyromagnetic ratio $g = 2$, automatically emerges. To see this explicitly, we reduce (1.34) further, keeping only first-order terms in the interaction with a weak uniform magnetic field $\mathbf{B} = \text{curl}\,\mathbf{A}$; $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$:

$$i\hbar \frac{\partial \varphi}{\partial t} = \left[\frac{\mathbf{p}^2}{2m} - \frac{e}{2mc} (\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B} \right] \varphi \tag{1.35}$$

Here $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is the orbital angular momentum, $\mathbf{S} = \frac{1}{2}\hbar\boldsymbol{\sigma}$ is the electron spin, with eigenvalues $\pm\hbar/2$, and the coefficient of the interaction of the spin with $\mathbf{B}$ field gives the correct magnetic moment of the electron corresponding to a g value of 2.

Fortified by this successful nonrelativistic reduction of the Dirac equation, we go on and establish the Lorentz covariance of the Dirac theory, as required by special relativity. Next we must investigate further physical consequences of this theory; especially we must interpret those "negative-energy" solutions.

Problems

1. Write the Maxwell equations in Dirac form (1.13) in terms of a six-component field amplitude. What are the matrices corresponding to α and β? [See H. E. Moses, *Phys. Rev.*, **113**, 1670 (1959).]
2. Verify that the matrices (1.17) satisfy the algebra of (1.16). ✓
3. Verify (1.33). ✓ $a_i b_j \sigma_i \sigma_j = a_i b_j (\delta_{ij} + i\varepsilon_{ijk}\sigma_k) = \vec{a}\cdot\vec{b} + i\varepsilon_{ijk} a_i b_j \sigma_k$ $= i\vec{\sigma}\cdot(\vec{a}\times\vec{b})$
4. Verify (1.27). ✓

[1] *Ibid.*

2.1 Covariant Form of the Dirac Equation

It is necessary that the Dirac equation and the continuity equation upon which its physical interpretation rests be covariant under Lorentz transformations. Let us first review what is meant by a Lorentz transformation.[1] Two observers O and O' who are in different inertial reference frames will describe the same physical event with the different space-time coordinates. The rule which relates the coordinates x^μ with which observer O describes the event to the coordinates $(x^\mu)'$ used by observer O' to describe the same event is given by the Lorentz transformation between the two sets of coordinates:

$$(x^\nu)' = \sum_{\mu=0}^{3} a^\nu{}_\mu x^\mu \equiv a^\nu{}_\mu x^\mu \tag{2.1}$$

It is a linear homogeneous transformation, and the coefficients $a^\nu{}_\mu$ depend only upon the relative velocities and spatial orientations of the two reference frames of O and O'. The basic invariant of the Lorentz transformation is the proper time interval

$$ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu = dx^\mu\, dx_\mu \tag{2.2}$$

This is derived from the physical observation that the velocity of light in vacuo is the same in all Lorentz frames. Equations (2.1) and (2.2) lead to the relation on the transformation coefficients

$$a_\mu{}^\nu a^\mu{}_\sigma = \delta^\nu{}_\sigma \tag{2.3}$$

Equations (2.1) and (2.3) serve as defining relations for both proper and improper Lorentz transformations. In the former case the determinant of the transformation coefficients satisfies the relation

$$\det |a| = +1$$

Not required but $\det a = 1$ and $a_{00} > 0$.

Proper Lorentz transformations can be built up by an infinite succession of infinitesimal transformations. They include transformations to coordinates in relative motion along any spatial direction as well as ordinary three-dimensional rotations. The improper Lorentz transformations are the discrete transformations of space inversion and of time inversion. They cannot be built up from a succession of infinitesimal ones. Their transformation coefficients satisfy the

[1] W. Pauli, "Theory of Relativity," Pergamon Press, New York, 1958. "The Principle of Relativity," collected papers of H. A. Lorentz, A. Einstein, H. Minkowski, and H. Weyl, Dover Publications, Inc., New York, 1923 reissue.

relation

$$\det |a| = -1$$

in both cases.

Our task is to construct a correspondence relating a given set of observations of a Dirac particle made by observers O and O' in their respective reference frames. In other words, we seek a transformation law relating the wave functions $\psi(x)$ and $\psi'(x')$ used by observers O and O', respectively. This transformation law is a rule which allows O' to compute $\psi'(x')$ if given $\psi(x)$. According to the requirement of Lorentz covariance, this transformation law must lead to wave functions which are solutions of Dirac equations of the same form in the primed as well as unprimed reference frame. This form invariance of the Dirac equation expresses the Lorentz invariance of the underlying energy-momentum connection

$$p_\mu p^\mu = m^2c^2$$

upon which the considerations of Chap. 1 were based.

In discussing covariance it is desirable to express the Dirac equation in a four-dimensional notation which preserves the symmetry between ct and x^i. To this end we multiply (1.13) by β/c and introduce the notation

$$\gamma^0 = \beta \qquad \gamma^i = \beta\alpha_i \qquad i = 1, 2, 3$$

This gives

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi = i\hbar\left(\gamma^0 \frac{\partial}{\partial x^0} + \gamma^1 \frac{\partial}{\partial x^1} + \gamma^2 \frac{\partial}{\partial x^2} + \gamma^3 \frac{\partial}{\partial x^3}\right)\psi - mc\psi = 0 \tag{2.4}$$

$(\gamma^\mu p_\mu - mc)\psi$

The new matrices γ^μ provide an elegant restatement of the commutation relations (1.16)

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}1 \tag{2.5}$$

where 1 is the 4×4 unit matrix and hereafter will not be explicitly indicated. It is clear from their definition that the γ^i are antihermitian, with $(\gamma^i)^2 = -1$, and that γ^0 is hermitian. In the representation (1.17) they have the form

$$\gamma^i = \begin{bmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{bmatrix} \qquad \gamma^0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.6}$$

It is convenient to introduce the Feynman dagger, or slash, notation:

$$A\!\!\!/ = \gamma^\mu A_\mu = g_{\mu\nu}\gamma^\mu A^\nu = \gamma^0 A^0 - \boldsymbol{\gamma}\cdot\mathbf{A} = \beta A^0 - \beta\vec{\alpha}\cdot\vec{A}$$

(In Weyl's matrix $\gamma_\mu A_\mu = A\!\!\!/ = \gamma_4 A_4 - i\beta\vec{\alpha}\cdot\vec{A} = i\beta(\phi - \vec{\alpha}\cdot\vec{A})$