Chapter 8

Nonrelativistic Limit of Dirac's Equation

In the nonrelativistic limit the Dirac theory should reproduce the Schrödinger-Pauli theory. Furthermore the Dirac theory must enable one to derive corrections to the Schrödinger-Pauli theory in the limit where relativistic effects are weak.

8.1 Large and small components

In the standard representation the upper two components of the Dirac wavefunction Ψ (which is a column matrix with four components) are much larger than the lower two components. This property enables one to derive the nonrelativistic limit in a simple way.

Reverting to normal units, the Dirac equation in the presence of an external field is

$$i\hbar\frac{\partial}{\partial t}\,\Psi = \left[c\boldsymbol{\alpha}\cdot(\hat{\boldsymbol{p}}+e\boldsymbol{A})+\beta mc^2-e\Phi\right]\Psi. \tag{8.1}$$

One appeals to the fact that the energy ε is close to mc^2 in the nonrelativistic limit. The positive energy part of the wavefunction varies with time as $\exp[-i\varepsilon t/\hbar]$, and this may be approximated by $\exp[-imc^2t/\hbar]$. The slower variation as $\exp[-i(\varepsilon-mc^2)t/\hbar]$ is retained in

$$\Psi' := \Psi\,\exp[imc^2t/\hbar]. \tag{8.2}$$

Next one writes

$$\Psi' = \begin{pmatrix}\phi'\\ \chi'\end{pmatrix} \tag{8.3}$$

where ϕ' and χ' are 2-spinors, i.e., column matrices with two components. Then (8.1) reduces to

$$\left(i\hbar\frac{\partial}{\partial t}+e\Phi\right)\phi' = c\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}}+e\boldsymbol{A})\chi', \tag{8.4}$$

$$(i\hbar\frac{\partial}{\partial t} + e\Phi + 2mc^2)\chi' = c\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}} + e\boldsymbol{A})\phi'. \tag{8.5}$$

Henceforth the primes on ϕ and χ are omitted.

Consider the left hand side of (8.5). The term $2mc^2$ is much larger than the other two terms in the nonrelativistic limit. Making the approximation in which only this dominant term is retained, one has

$$\chi' \approx \frac{1}{2mc}\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}} + e\boldsymbol{A})\phi'. \tag{8.6}$$

Substituting (8.6) into (8.4) gives

$$\left(i\hbar\frac{\partial}{\partial t} + e\Phi\right)\phi' \approx \frac{1}{2m}\left[\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}} + e\boldsymbol{A})\right]^2\phi'. \tag{8.7}$$

For arbitrary vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ one has

$$\boldsymbol{\sigma}\cdot\boldsymbol{a}\,\boldsymbol{\sigma}\cdot\boldsymbol{b} = \boldsymbol{a}\cdot\boldsymbol{b} + i\boldsymbol{\sigma}\cdot\boldsymbol{a}\times\boldsymbol{b}. \tag{8.8}$$

On applying this to the right hand side of (8.7) and using

$$[\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}} + e\boldsymbol{A})]\,[\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}} + e\boldsymbol{A})] = (\hat{\boldsymbol{p}} + e\boldsymbol{A})^2 + ie\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}}\times\boldsymbol{A} + \boldsymbol{A}\times\hat{\boldsymbol{p}}), \tag{8.9}$$

the final term reduces to $-ie\hbar\,\mathrm{curl}\boldsymbol{A} = -ie\hbar\boldsymbol{B}$, and equation (8.7) reduces to

$$i\hbar\frac{\partial}{\partial t}\phi = \left[\frac{1}{2m}(\hat{\boldsymbol{p}} + e\boldsymbol{A})^2 - e\Phi + \frac{e\hbar}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B}\right]\phi. \tag{8.10}$$

Equation (8.10) is just the Pauli-Schrödinger equation (1.20), as required.

8.2 Foldy-Wouthuysen transformation

There is a systematic way of generalizing this approach to derive a generalization of the Pauli-Schrödinger equation which includes relativistic corrections. The idea is to separate the Hamiltonian into *even* and *odd* components, denoted $\hat{\mathcal{O}}$ and $\hat{\mathcal{E}}$ respectively. These are defined to satisfy

$$\beta\hat{\mathcal{O}} = -\hat{\mathcal{O}}\beta, \qquad \beta\hat{\mathcal{E}} = \hat{\mathcal{E}}\beta. \tag{8.11}$$

One has

$$\hat{\mathcal{H}} = \beta mc^2 + \hat{\mathcal{O}} + \hat{\mathcal{E}}, \tag{8.12}$$

with

$$\hat{\mathcal{O}} = c\boldsymbol{\alpha}\cdot(\hat{\boldsymbol{p}} + e\boldsymbol{A}), \quad \hat{\mathcal{E}} = -e\Phi, \tag{8.13}$$

for the Hamiltonian in (8.1). Suppose a change of representation is made so that in the new representation $\hat{\mathcal{O}}$ is zero. Then the equations for the large and small components are decoupled, and the equations for the large component is the desired generalization of the Pauli-Schrödinger equation. The appropriate change of representation is called the *Foldy-Wouthuysen* transformation.

Let the (unitary) transformation operator be written as $\exp[i\hat{S}]$, where $\hat{S}$ is hermitian. Our objective is to find the appropriate $\hat{S}$ and hence the generalization of the Pauli-Schrödinger equation. The new time-dependent Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} \Psi' = \hat{\mathcal{H}}'\Psi', \tag{8.14}$$

with

$$\Psi' = \exp(i\hat{S})\, \Psi, \quad \hat{\mathcal{H}}' = \exp(i\hat{S})\, \hat{\mathcal{H}}\, \exp(-i\hat{S}) - i\hbar\, \exp(i\hat{S})\, \frac{\partial}{\partial t}\, \exp(-i\hat{S}). \tag{8.15}$$

The detailed evaluation of the new Hamiltonian is made using the identity

$$\exp(\hat{A})\, \hat{B}\, \exp(-\hat{A}) = \hat{B} + [\hat{A}, \hat{B}] + \tfrac{1}{2}[\hat{A}, [\hat{A}, \hat{B}]] + \cdots, \tag{8.16}$$

where the nth term contains n nested commutators and a numerical factor $(n!)^{-1}$. Thus one finds

$$\hat{\mathcal{H}}' = \hat{\mathcal{H}} + i[\hat{S}, \hat{\mathcal{H}}] - \tfrac{1}{2}[\hat{S}, [\hat{S}, \hat{\mathcal{H}}]] + \cdots - \hbar\, \frac{\partial}{\partial t}\, \hat{S} - i\tfrac{1}{2}\hbar\, [\hat{S}, \frac{\partial}{\partial t}\, \hat{S}] + \cdots. \tag{8.17}$$

Suppose that $\hat{S}$ contains a small parameter. This is necessarily the case if the nonrelativistic effects are sufficiently weak. A perturbation expansion involves expanding in this parameter, i.e., in powers of $\hat{S}$. (It is assumed that $\partial\hat{S}/\partial t$ is of higher order than $\hat{S}$.) To zeroth order the Hamiltonian is βmc^2. To first order in $\hat{S}$ one finds

$$\hat{\mathcal{H}}' = \beta mc^2 + \hat{\mathcal{O}} + \hat{\mathcal{E}} + i[\hat{S}, \beta mc^2]. \tag{8.18}$$

To this order the odd terms may be eliminated by requiring

$$\hat{\mathcal{O}} + i[\hat{S}, \beta mc^2] = 0. \tag{8.19}$$

Using (8.11) the solution of (8.19) is

$$\hat{S} = -i\, \frac{\beta\hat{\mathcal{O}}}{2mc^2}\, . \tag{8.20}$$

To find the Hamiltonian to this order one inserts (8.20) in (8.17) and retains only the second order terms. The result is the Pauli-Schrödinger equation.

8.3 Corrections to Pauli-Schrödinger equation

The expansion (7.17) to next order contains odd terms which may be eliminated in the same way. That is one chooses a correction to (8.20) that causes the odd terms to also vanish to next highest order. This generalization of (8.20) is then used to evaluated the next highest order terms, and so find relativisitic corrections to the Pauli-Schrödinger equation.

Omitting the details of the calculation, the result is

$$\begin{aligned}\hat{\mathcal{H}} = \beta mc^2 + \frac{1}{2m}(\hat{\boldsymbol{p}}+e\boldsymbol{A})^2 - e\Phi + \frac{e\hbar}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B} \\ - \frac{1}{8m^3c^2}\left\{(\hat{\boldsymbol{p}}+e\boldsymbol{A})^2 + e(\boldsymbol{\sigma}\cdot\boldsymbol{B})\right\}^2 + \frac{ie\hbar^2}{8m^2c^2}\boldsymbol{\sigma}\cdot\mathrm{curl}\boldsymbol{E} \\ + \frac{e\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot\boldsymbol{E}\times(\hat{\boldsymbol{p}}+e\boldsymbol{A}) + \frac{e\hbar^2}{8m^2c^2}\mathrm{div}\boldsymbol{E}.\end{aligned} \tag{8.21}$$

The first term on the right hand side of (8.21) corresponds to the rest energy for electrons, when only the leading 2 × 2 elements are retained. The second, third and fourth terms on the right hand side of (8.21) are just the Pauli-Schrödinger terms. The fourth term may be interpreted as a relativisitic correction from the expansion of $\sqrt{m^2c^4 + \hat{\boldsymbol{p}}^2c^2} - mc^2$ to second order in $\hat{\boldsymbol{p}}^2$. The fifth term is rarely important; Maxwell's equations allow one to re-express curl$\boldsymbol{E}$ as $-\partial\boldsymbol{B}/\partial t$, and hence this term is important only when the time derivative of the magnetic field is important. The sixth term is the spin orbit coupling term which is introduced is an artificial way into the nonrelativistic theory. The final term is called the *Darwin* term.

The Darwin term may be attributed to the Zitterbewegung. An electron fluctuates in position very rapidly over a distance of order its Compton wavelength $\hbar/mc$. The effective potential that it experiences is the potential Φ at its mean position plus a correction due to the average over these rapid fluctuations. Let the correction be $\Delta\Phi$. On making a Taylor series expansion about the mean position $\boldsymbol{x}_0$, with the fluctuations over a distance $\Delta\boldsymbol{x}$, one has

$$\Phi(\boldsymbol{x}) = \Phi(\boldsymbol{x}_0) + \Delta\boldsymbol{x}\cdot\frac{\partial}{\partial\boldsymbol{x}_0}\Phi(\boldsymbol{x}_0) + \tfrac{1}{2}\Delta\boldsymbol{x}_i\Delta\boldsymbol{x}_j\frac{\partial^2}{\partial\boldsymbol{x}_{0i}\partial\boldsymbol{x}_{0j}}\Phi(\boldsymbol{x}_0) + \cdots. \tag{8.22}$$

On averaging over the position of the electron, the term linear in $\Delta\boldsymbol{x}$ averages to zero, and the second order term gives

$$\left\langle \tfrac{1}{2}\Delta\boldsymbol{x}_i\Delta\boldsymbol{x}_j\frac{\partial^2}{\partial\boldsymbol{x}_{0i}\partial\boldsymbol{x}_{0j}}\Phi(\boldsymbol{x}_0)\right\rangle \approx \left(\frac{\hbar}{mc}\right)^2\nabla^2\Phi, \tag{8.23}$$

where the angular brackets denote an average over position, and where $\langle\Delta\boldsymbol{x}^2\rangle$ is assumed to be of order $(\hbar/mc)^2$. Apart from the numerical factor, (8.23) with $\nabla^2\Phi = -\mathrm{div}\boldsymbol{E}$ reproduces the Darwin term.