4

The

Foldy-Wouthuysen

Transformation

T

4.1 Introduction

Aside from the negative-energy problem, the Dirac equation appears to provide a suitable description of the electron. It has a sensible nonrelativistic limit, and it automatically yields the correct magnetic moment. We now investigate the interaction of the Dirac electron with prescribed external potentials. In particular, we shall be primarily interested in low-energy properties, avoiding the difficulties associated with the as yet uninterpreted negative-energy solutions, which are an essentially relativistic feature. We anticipate from our discussions of the packet in the preceding chapter that in practice they play a very minor role in a problem such as the hydrogen atom, which finds the electron localized in Bohr orbits of radius[1] $1/\alpha m \gg 1/m$.

We shall see, in fact, that the stationary energy levels deduced from the Dirac equation for the hydrogen atom are in exceedingly close agreement with the observed eigenvalues. However, before indicating the solution to the eigenvalue problem in the Coulomb potential, it is instructive to cast the Dirac theory in a form which displays the different interaction terms between the electron and an applied field in a nonrelativistic and easily interpretable form.

We consider, then, a systematic procedure developed by Foldy and Wouthuysen,[2] namely, a canonical transformation which decouples the Dirac equation into two two-component equations: one reduces to the Pauli description in the nonrelativistic limit; the other describes the negative-energy states.

4.2 Free-particle Transformation

As a first illustration of the Foldy-Wouthuysen transformation we consider the Dirac equation for a free particle, most conveniently—for this purpose—written in hamiltonian form and with the α matrices in the representation introduced in Eq. (1.17). We search for a unitary transformation U_F which will remove from the equation all operators such as α which couple the large to the small components. We call

---

[1] Henceforth we set $\hbar = c = 1$. The Compton wavelength of the electron is $1/m = 3.86 \times 10^{-11}$ cm, and the rest energy $m = 0.511$ MeV. The dimensionless fine-structure constant is $\alpha = e^2/4\pi \cong \frac{1}{137}$.

$$0.511 \text{ MeV} = \frac{10^{11}}{3.86} \text{ cm}^{-1} = \frac{10^{21}}{1.29} \text{ sec}^{-1} = m$$

in these units.

[2] L. L. Foldy and S. A. Wouthuysen, *Phys. Rev.*, **78**, 29 (1950).

46

any such operator "odd"; operators which do not couple large and small components are "even"; thus α, γ, γ_5, etc., are odd, and 1, β, δ, etc., are even.

Writing $U_F = e^{+iS}$ with S hermitian and not explicitly time-dependent, the unitary transformation is

$$\psi' = e^{+iS}\psi$$

and

$$i\frac{\partial\psi'}{\partial t} = e^{+iS}H\psi = e^{+iS}He^{-iS}\psi' = H'\psi'$$

H' is to contain no odd operators by construction.

Since $H = \alpha \cdot p + \beta m$ with $\{\alpha,\beta\} = 0$, our problem is quite analogous to that of attempting to find a unitary transformation which changes a two-component spin hamiltonian $\mathcal{H} = \sigma_x B_x + \sigma_z B_z$ into a form which contains only even operators (that is, 1 and σ_z). Such a transformation is simply a rotation about the y axis and the operator is $e^{+(i/2)\sigma_y\theta_0} = e^{+\frac{1}{2}\sigma_x\sigma_z\theta_0}$, with $\tan\theta_0 = B_x/B_z$. This suggests that a good operator to try in our case would be

$$e^{iS} = e^{\beta\alpha \cdot p\theta(p)} = \cos|p|\theta + \frac{\beta\alpha \cdot p}{|p|}\sin|p|\theta$$

where the right-hand side is established by expansion of the exponential in powers of θ.

With this choice H' becomes:

$$H' = \left(\cos|p|\theta(p) + \frac{\beta\alpha \cdot p}{|p|}\sin|p|\theta(p)\right)(\alpha \cdot p + \beta m)\left(\cos|p|\theta - \frac{\beta\alpha \cdot p}{|p|}\sin|p|\theta\right)$$

$$= (\alpha \cdot p + \beta m)\left(\cos|p|\theta - \frac{\beta\alpha \cdot p}{|p|}\sin|p|\theta\right)^2$$

$$= (\alpha \cdot p + \beta m)\exp\left(-2\beta\alpha \cdot p\theta\right)$$

$$= \alpha \cdot p(\cos 2|p|\theta - \frac{m}{|p|}\sin 2|p|\theta) + \beta(m\cos 2|p|\theta + |p|\sin 2|p|\theta)$$

In order to eliminate the odd operator, we choose

$$\tan 2|p|\theta = \frac{|p|}{m}$$

and the transformed hamiltonian is

$$H' = \beta \sqrt{m^2 + p^2} \tag{4.1}$$

as may be verified with the aid of the triangle construction of Fig. 4.1. The new hamiltonian is just the one rejected in Chap. 1, with the important change that now the negative energies are also accepted. The negative energies and four-component wave functions are the price we must pay in order to have a factorization of H' in (4.1) into a linear Dirac equation.

4.3 The General Transformation

We turn now to the more general case of an electron in a prescribed external electromagnetic field and search for the corresponding transformation S. The hamiltonian is

$$H = \boldsymbol{\alpha} \cdot (\boldsymbol{p} - e\mathbf{A}) + \beta m + e\Phi$$
$$= \beta m + \mathcal{O} + \mathcal{E} \tag{4.2}$$

with $\mathcal{O} = \boldsymbol{\alpha} \cdot (\boldsymbol{p} - e\mathbf{A})$ and $\mathcal{E} = e\Phi$; as before, $\beta\mathcal{O} = -\mathcal{O}\beta$ and

$$\beta\mathcal{E} = +\mathcal{E}\beta$$

The fields appearing in (4.2) and hence the hamiltonian itself may be time-dependent. In the general case the transformation S is also time-dependent and it is not possible to construct an S which removes the odd operators from H' to all orders, as was achieved in (4.1). Therefore, we content ourselves with a nonrelativistic expansion of the transformed hamiltonian in a power series in $1/m$, keeping terms only through order (kinetic energy/m)^3 and (kinetic energy)(field energy)/m^2.
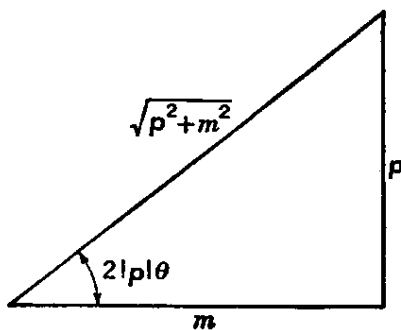

Fig. 4-1   Foldy-Wouthuysen triangle construction.

The Foldy-Wouthuysen transformation   **49**

Again we introduce the transformation by

$$\psi' = e^{iS}\psi$$

finding

$$i\frac{\partial}{\partial t}e^{-iS}\psi' = H\psi = He^{-iS}\psi' = e^{-iS}\left(i\frac{\partial\psi'}{\partial t}\right) + \left(i\frac{\partial}{\partial t}e^{-iS}\right)\psi'$$

Thus

$$i\frac{\partial\psi'}{\partial t} = \left[e^{iS}\left(H - i\frac{\partial}{\partial t}\right)e^{-iS}\right]\psi' = H'\psi'$$

Since S is expanded in powers of $1/m$ and is therefore "small" in the nonrelativistic limit, we expand the quantity in brackets in a series of multiple commutators, using the relation[1]

$$e^{+iS}He^{-iS} = H + i[S,H] + \frac{(i)^2}{2!}[S,[S,H]] + \cdots$$

$$+ \frac{(i)^n}{n!}[S,[S,\cdots,[S,H]\cdots]] + \cdots$$

Since $S = O(1/m)$, to the desired order of accuracy we have

$$H' = H + i[S,H] - \frac{1}{2}[S,[S,H]] - \frac{i}{6}[S,[S,[S,H]]]$$

$$+ \frac{1}{24}[S,[S,[S,[S,\beta m]]]] - \dot{S} - \frac{i}{2}[S,\dot{S}] + \frac{1}{6}[S,[S,\dot{S}]]$$

To start constructing S, we consider just the terms through order unity:

$$H' = \beta m + \mathcal{E} + \mathcal{O} + i[S,\beta]m \qquad (4.3)$$

We require that the odd term in (4.3) vanish; and taking our cue from the behavior in the free-particle case, we choose $S = -i\beta\mathcal{O}/2m$.

---

[1] This may be verified by considering

$$F(\lambda) = e^{i\lambda S}He^{-i\lambda S} = \sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\left(\frac{\partial^n F}{\partial\lambda^n}\right)_{\lambda=0} \qquad (a)$$

It follows that

$$\frac{\partial F}{\partial\lambda} = e^{i\lambda S}i[S,H]e^{-i\lambda S}$$

and thus

$$\frac{\partial^n F}{\partial\lambda^n} = e^{i\lambda S}i^n[S,[S,\cdots,[S,H]\cdots]]e^{-i\lambda S}$$

from which the identity follows upon setting $\lambda = 1$ in (a).

We then find, to the order of accuracy desired,

$$i[S,H] = -\mathcal{O} + \frac{\beta}{2m}[\mathcal{O},\mathcal{E}] + \frac{1}{m}\beta\mathcal{O}^2$$

$$\frac{i^2}{2}[S,[S,H]] = -\frac{\beta\mathcal{O}^2}{2m} - \frac{1}{8m^2}[\mathcal{O},[\mathcal{O},\mathcal{E}]] - \frac{1}{2m^2}\mathcal{O}^3$$

$$\frac{i^3}{3!}[S,[S,[S,H]]] = \frac{\mathcal{O}^3}{6m^2} - \frac{1}{6m^3}\beta\mathcal{O}^4$$

$$\frac{i^4}{4!}[S,[S,[S,[S,H]]]] = \frac{\beta\mathcal{O}^4}{24m^3}$$

$$-\dot{S} = +\frac{i\beta\dot{\mathcal{O}}}{2m}$$

$$-\frac{i}{2}[S,\dot{S}] = -\frac{i}{8m^2}[\mathcal{O},\dot{\mathcal{O}}]$$

Collecting everything together,

$$H' = \beta\left(m + \frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^3}\right) + \mathcal{E} - \frac{1}{8m^2}[\mathcal{O},[\mathcal{O},\mathcal{E}]] - \frac{i}{8m^2}[\mathcal{O},\dot{\mathcal{O}}]$$

$$+ \frac{\beta}{2m}[\mathcal{O},\mathcal{E}] - \frac{\mathcal{O}^3}{3m^2} + \frac{i\beta\dot{\mathcal{O}}}{2m} = \beta m + \mathcal{E}' + \mathcal{O}' \quad (4.4)$$

The odd terms now appear in (4.4) only in order $1/m$. To reduce them further, we apply a second Foldy-Wouthuysen transformation using the same prescription:

$$S' = \frac{-i\beta}{2m}\mathcal{O}' = \frac{-i\beta}{2m}\left(\frac{\beta}{2m}[\mathcal{O},\mathcal{E}] - \frac{\mathcal{O}^3}{3m^2} + \frac{i\beta\dot{\mathcal{O}}}{2m}\right)$$

Under this transformation we find

$$H'' = e^{iS'}\left(H' - i\frac{\partial}{\partial t}\right)e^{-iS'} = \beta m + \mathcal{E}' + \frac{\beta}{2m}[\mathcal{O}',\mathcal{E}'] + \frac{i\beta\dot{\mathcal{O}}'}{2m}$$

$$= \beta m + \mathcal{E}' + \mathcal{O}''$$

where $\mathcal{O}''$ is now $O(1/m^2)$. Finally, by a third canonical transformation

$$S'' = \frac{-i\beta\mathcal{O}''}{2m}$$

cs

The Foldy-Wouthuysen transformation 51

this term may also be eliminated in the same way, the end result being

$$H''' = e^{iS''}\left(H'' - i\frac{\partial}{\partial t}\right)e^{-iS''} = \beta m + \mathcal{E}'$$

$$= \beta\left(m + \frac{\Theta^2}{2m} - \frac{\Theta^4}{8m^3}\right) + \mathcal{E} - \frac{1}{8m^2}[\Theta,[\Theta,\mathcal{E}]] - \frac{i}{8m^2}[\Theta,\dot{\Theta}]$$

Evaluating the operator products to the desired order of accuracy, we find

$$\frac{\Theta^2}{2m} = \frac{(\boldsymbol{\alpha}\cdot(\mathbf{p}-e\mathbf{A}))^2}{2m} = \frac{(\mathbf{p}-e\mathbf{A})^2}{2m} - \frac{e}{2m}\boldsymbol{\sigma}\cdot\mathbf{B}$$

$$\frac{1}{8m^2}([\Theta,\mathcal{E}] + i\dot{\Theta}) = \frac{e}{8m^2}(-i\boldsymbol{\alpha}\cdot\nabla\Phi - i\boldsymbol{\alpha}\cdot\dot{\mathbf{A}}) = \frac{ie}{8m^2}\boldsymbol{\alpha}\cdot\mathbf{E}$$

$$\left[\Theta, \frac{ie}{8m^2}\boldsymbol{\alpha}\cdot\mathbf{E}\right] = \frac{ie}{8m^2}[\boldsymbol{\alpha}\cdot\mathbf{p},\boldsymbol{\alpha}\cdot\mathbf{E}]$$

$$= \frac{ie}{8m^2}\sum_{ij}\alpha_i\alpha_j\left(-i\frac{\partial E^j}{\partial x^i}\right) + \frac{e}{4m^2}\boldsymbol{\sigma}\cdot\mathbf{E}\times\mathbf{p}$$

$$= \frac{e}{8m^2}\,\mathrm{div}\,\mathbf{E} + \frac{ie}{8m^2}\boldsymbol{\sigma}\cdot\mathrm{curl}\,\mathbf{E} + \frac{e}{4m^2}\boldsymbol{\sigma}\cdot\mathbf{E}\times\mathbf{p}$$

and thus the reduced hamiltonian is to this order

$$H''' = \beta\left(m + \frac{(\mathbf{p}-e\mathbf{A})^2}{2m} - \frac{\mathbf{p}^4}{8m^3}\right) + e\Phi - e\frac{1}{2m}\beta\boldsymbol{\sigma}\cdot\mathbf{B}$$

$$- \frac{ie}{8m^2}\boldsymbol{\sigma}\cdot\mathrm{curl}\,\mathbf{E} - \frac{e}{4m^2}\boldsymbol{\sigma}\cdot\mathbf{E}\times\mathbf{p} - \frac{e}{8m^2}\,\mathrm{div}\,\mathbf{E} \quad (4.5)$$

The individual terms in (4.5) have a direct physical interpretation. The terms in the first bracket give the expansion of

$$\sqrt{(\mathbf{p}-e\mathbf{A})^2 + m^2}$$

to the desired order, showing the relativistic mass increase. The second and third terms are the electrostatic and magnetic dipole energies. The next pair of terms which taken together are hermitian comprise the spin-orbit energy, and they have a very familiar form in a spherically symmetric static potential. In this case curl $\mathbf{E} = 0$,

$$\boldsymbol{\sigma}\cdot\mathbf{E}\times\mathbf{p} = -\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{\sigma}\cdot\mathbf{r}\times\mathbf{p} = -\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{\sigma}\cdot\mathbf{L}$$

and this term reduces to

$$H_{\text{spin-orbit}} = \frac{e}{4m^2}\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{\sigma}\cdot\mathbf{L} \quad (4.6)$$

Equation (4.6) is in agreement with the classical result obtained by considering the magnetic field $\mathbf{B}' = -\mathbf{v} \times \mathbf{E}$ experienced by the moving electron. The interaction energy one would expect is thus

$$-\frac{e}{2m}\, \mathbf{\delta} \cdot \mathbf{B}' = \frac{e}{2m^2}\, \mathbf{\delta} \cdot (\mathbf{p} \times \mathbf{E})$$

However, this is reduced by a factor of 2 owing to the Thomas precession effect and indicates that the orbital moment of the electron has the standard gyromagnetic ratio of $g_e = 1$.

The last term—known as the Darwin term—may be attributed to the zitterbewegung. Because the electron coordinate fluctuates over distances $\delta r \simeq 1/m$, it sees a somewhat smeared out Coulomb potential; the correction is

$$\langle \delta V \rangle = \langle V(\mathbf{r} + \delta\mathbf{r}) \rangle - \langle V(\mathbf{r}) \rangle = \left\langle \delta r \frac{\partial V}{\partial r} + \frac{1}{2} \sum_{ij} \delta r_i\, \delta r_j\, \frac{\partial^2 V}{\partial r_i\, \partial r_j} \right\rangle$$

$$\cong \frac{1}{6}\, \delta r^2\, \nabla^2 V \simeq \frac{1}{6m^2}\, \nabla^2 V \qquad (4.7)$$

in qualitative accord with the sign, form, and magnitude of the Darwin term.

4.4 The Hydrogen Atom[1]

We turn to a discussion of the bound-state solutions of the Dirac equation, considering in particular the energy levels of the electron in a Coulomb field. For this problem the Dirac equation is

$$H\psi = [\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m + V(r)]\psi = E\psi \qquad (4.8)$$

with $V = -Z\alpha/r$. In order to separate variables, we take advantage of the fact that the angular momentum of a particle in a central field is conserved. Evidently $\mathbf{J} = \mathbf{L} + \mathbf{S} = \mathbf{r} \times \mathbf{p} + \frac{1}{2}\mathbf{\delta}$ commutes with the hamiltonian (4.8) and therefore we may construct simultaneous eigenfunctions of H, J^2, and J_z. To do this, we call on experience with the Pauli matrices, observing that in the representation of

[1] The eigensolutions in the Coulomb potential were first given by C. G. Darwin, *Proc. Roy. Soc. (London)*, **A118**, 654 (1928), and W. Gordon, *Z. Physik*, **48**, 11 (1928). For a complete discussion and references of the atomic applications of the Dirac equation see H. A. Bethe and E. E. Salpeter, "Quantum Mechanics of One- and Two-electron Atoms," Academic Press Inc., New York, 1957, and M. E. Rose, "Relativistic Electron Theory," John Wiley & Sons, Inc., New York, 1961.

Chap. 3

$$\mathbf{\delta} = \begin{bmatrix} \mathbf{\delta} & 0 \\ 0 & \mathbf{\delta} \end{bmatrix}$$

is diagonal in terms of 2×2 Pauli spin matrices. Therefore, if we construct ψ in terms of two-component spinors

$$\psi = \begin{bmatrix} \varphi \\ \chi \end{bmatrix}$$

the angular separation for the solutions of φ and χ is precisely that of the Pauli two-component theory. The two-component angular solutions are eigenfunctions of J^2, J_z, L^2, and S^2 and are of two types:
For $j = l + \frac{1}{2}$

$$\varphi_{j,m}^{(+)} = \begin{bmatrix} \sqrt{\dfrac{l + \frac{1}{2} + m}{2l + 1}}\, Y_l^{m-\frac{1}{2}} \\ \sqrt{\dfrac{l + \frac{1}{2} - m}{2l + 1}}\, Y_l^{m+\frac{1}{2}} \end{bmatrix} \tag{4.9a}$$

For $j = l - \frac{1}{2}$

$$\varphi_{j,m}^{(-)} = \begin{bmatrix} \sqrt{\dfrac{l + \frac{1}{2} - m}{2l + 1}}\, Y_l^{m-\frac{1}{2}} \\ -\sqrt{\dfrac{l + \frac{1}{2} + m}{2l + 1}}\, Y_l^{m+\frac{1}{2}} \end{bmatrix} \tag{4.9b}$$

The spherical harmonics here are written with the convention $Y_{l,m}^* = (-)^m Y_{l,-m}$, and the solution $\varphi^{(-)}$ exists only for $l > 0$. The two solutions above satisfy the eigenvalue equations

$$J^2 \varphi_{jm}^{(\pm)} = j(j + 1)\varphi_{jm}$$

and

$$\mathbf{L} \cdot \mathbf{\delta} \varphi_{jm}^{(\pm)} = (J^2 - L^2 - \tfrac{3}{4})\varphi_{jm}^{(\pm)}$$
$$= -(1 + \kappa)\varphi_{jm}^{(\pm)}$$

with

$$\kappa = \begin{cases} -(l + 1) = -(j + \frac{1}{2}) & j = l + \frac{1}{2} \\ +l = +(j + \frac{1}{2}) & j = l - \frac{1}{2} \end{cases}$$

For a given j they are of opposite parity, since their l values differ by 1, and can be formed from each other by a scalar operator of odd parity. This operator will be a linear combination of $Y_1^m(\theta,\varphi)$ since it must change the l value by 1, and is therefore proportional to $\mathbf{r}$. Dotting with $\mathbf{\delta}$, the only pseudovector at our disposal, we form the

pseudoscalar $\mathbf{d} \cdot \mathbf{r}/r$ and find with the above sign convention

$$\varphi_{jm}^{(+)} = \frac{\mathbf{d} \cdot \mathbf{r}}{r} \varphi_{jm}^{(-)} \qquad (4.10)$$

The general solution to the central field problem for a given jm is

$$\psi_{jm} = \begin{pmatrix} \dfrac{iG_j^+}{r} \varphi_{jm}^{(+)} + \dfrac{iG_j^-}{r} \varphi_{jm}^{(-)} \\[2ex] \dfrac{F_j^+}{r} \varphi_{jm}^{(-)} + \dfrac{F_j^-}{r} \varphi_{jm}^{(+)} \end{pmatrix}$$

We may finally break this down into two solutions each of definite parity. Since $V(r)$ is invariant under reflection of coordinates, we know that the energy eigenfunctions can be classified into parity eigenstates along with (j,m); and therefore we form the even and odd solutions, which have the property under the transformation $\mathbf{x}' = -_d\mathbf{x}$

$$\psi'(x') = \beta\psi(x) = \pm\psi(x') \qquad (4.11)$$

These are given by

$$\psi_{jm}^l = \begin{bmatrix} \dfrac{iG_{lj}}{r} \varphi_{jm}^l \\[2ex] \dfrac{F_{lj}}{r} \dfrac{\mathbf{d} \cdot \mathbf{r}}{r} \varphi_{jm}^l \end{bmatrix} \qquad (4.12)$$

where as a common notation we have introduced

$$G_{lj} = \begin{cases} G_j^+ & j = l + \tfrac{1}{2} \\ G_j^- & j = l - \tfrac{1}{2} \end{cases} \qquad\qquad F_{lj} = \begin{cases} F_j^+ & j = l + \tfrac{1}{2} \\ F_j^- & j = l - \tfrac{1}{2} \end{cases}$$

$$\varphi_{jm}^l = \begin{cases} \varphi_{jm}^+ & j = l + \tfrac{1}{2} \\ \varphi_{jm}^- & j = l - \tfrac{1}{2} \end{cases}$$

and have made use of (4.10). The parity of these solutions is $(-)^l$ by the convention (4.11). With the aid of the following identities we can now find the radial equations following from (4.8):

$$\mathbf{d} \cdot \mathbf{p} \frac{f(r)}{r} \varphi_{jm}^l = \frac{\mathbf{d} \cdot \mathbf{r}}{r^2} (\mathbf{d} \cdot \mathbf{r}\, \mathbf{d} \cdot \mathbf{p}) \frac{f(r)}{r} \varphi_{jm}^l$$

$$= \frac{\mathbf{d} \cdot \mathbf{r}}{r^2} \left(\frac{1}{i} r \frac{\partial}{\partial r} + i\, \mathbf{d} \cdot \mathbf{L} \right) \frac{f(r)}{r} \varphi_{jm}^l$$

$$= \left[\frac{1}{i} \frac{d}{dr} \frac{f(r)}{r} - i(1 + \kappa) \frac{f(r)}{r^2} \right] \left(\frac{\mathbf{d} \cdot \mathbf{r}}{r} \right) \varphi_{jm}^l$$

The radial equations are then

$$\left(E - m + \frac{Z\alpha}{r}\right) G_{lj}(r) = -\frac{dF_{lj}(r)}{dr} + \frac{\kappa}{r} F_{lj}(r)$$

$$\left(E + m + \frac{Z\alpha}{r}\right) F_{lj}(r) = +\frac{dG_{lj}(r)}{dr} + \frac{\kappa}{r} G_{lj}(r) \tag{4.13}$$

The bound-state solutions of these equations may be found by standard methods;[1] we quote only some of the results.

The energy eigenvalues are

$$E_n = m \left[1 + \left(\frac{Z\alpha}{n - (j + \frac{1}{2}) + \sqrt{(j + \frac{1}{2})^2 - Z^2\alpha^2}}\right)^2\right]^{-\frac{1}{2}} \tag{4.14}$$

where the quantum number $n = 1, 2, \ldots, \infty$ is a positive integer and the angular-momentum eigenvalues range from 0 to $j + \frac{1}{2} \le n$, with the restriction $0 \le l \le n - 1$. Expanding (4.14) in powers of $(Z\alpha)^2$, we see that n corresponds to the principal quantum number of the nonrelativistic theory

$$E_n \approx m \left\{1 - \frac{1}{2}\frac{Z^2\alpha^2}{n^2}\left[1 + \frac{(Z\alpha)^2}{n}\left(\frac{1}{j + \frac{1}{2}} - \frac{3}{4n}\right)\right] + 0((Z\alpha)^6)\right\} \tag{4.15}$$

The ground-state energy is, with $n = 1$, $j = \frac{1}{2}$,

$$E_g = m\sqrt{1 - Z^2\alpha^2} \cong m - \frac{1}{2}Z^2\alpha^2 m - \frac{1}{8}Z^4\alpha^4 m + \cdots$$

The corresponding spin-up and spin-down normalized eigenfunctions are

$$\psi_{n=1, j=\frac{1}{2}, \uparrow}(r, \theta, \varphi)$$

$$= \frac{(2mZ\alpha)^{\frac{3}{2}}}{\sqrt{4\pi}} \sqrt{\frac{1 + \gamma}{2\Gamma(1 + 2\gamma)}} (2mZ\alpha r)^{\gamma - 1} e^{-mZ\alpha r} \begin{bmatrix} 1 \\ 0 \\ \dfrac{i(1 - \gamma)}{Z\alpha}\cos\theta \\ \dfrac{i(1 - \gamma)}{Z\alpha}\sin\theta e^{i\varphi} \end{bmatrix}$$

$$\psi_{n=1, j=\frac{1}{2}, \downarrow}(r, \theta, \varphi)$$

$$= \frac{(2mZ\alpha)^{\frac{3}{2}}}{\sqrt{4\pi}} \sqrt{\frac{1 + \gamma}{2\Gamma(1 + 2\gamma)}} (2mZ\alpha r)^{\gamma - 1} e^{-mZ\alpha r} \begin{bmatrix} 0 \\ 1 \\ \dfrac{i(1 - \gamma)}{Z\alpha}\sin\theta e^{-i\varphi} \\ \dfrac{-i(1 - \gamma)}{Z\alpha}\cos\theta \end{bmatrix}$$

[1] Darwin, Gordon, Bethe and Salpeter, and Rose, *op. cit.*

with $\gamma = \sqrt{1 - Z^2\alpha^2}$. In the nonrelativistic limit $\gamma \to 1$ and $(1 - \gamma)/Z\alpha \to 0$, and they reduce to the Schrödinger wave functions multiplied by two-component Pauli spinors. In the relativistic case we see that as $r \to 0$, ψ exhibits a mild singularity of order $(2mZ\alpha r)^{-(Z^2\alpha^2)/2}$ which becomes important only at distances

$$r \sim \frac{1}{2mZ\alpha}\, e^{-2/Z^2\alpha^2}$$

For $Z\alpha \geq 1$, γ is imaginary and the solutions exhibit an oscillatory behavior reminiscent of that found in the Klein paradox. In this case there is no longer a gap between the positive- and negative-energy spectra, and again we lack a physical interpretation of the solution.

In classifying the energy levels (4.14) it is customary to denote them by their nonrelativistic labels, that is, by the orbital angular momentum l appearing in ψ^l_{jm} and by the total j. In the following table we list a few of the first terms:

	n	l	j	E_{nj}
$1S_{1/2}$	1	0	$\tfrac{1}{2}$	$m\sqrt{1 - Z^2\alpha^2}$
$2S_{1/2}$	2	0	$\tfrac{1}{2}$	$m\sqrt{\dfrac{1 + \sqrt{1 - Z^2\alpha^2}}{2}}$
$2P_{1/2}$	2	1	$\tfrac{1}{2}$	$m\sqrt{\dfrac{1 + \sqrt{1 - Z^2\alpha^2}}{2}}$
$2P_{3/2}$	2	1	$\tfrac{3}{2}$	$\dfrac{m}{2}\sqrt{4 - Z^2\alpha^2}$

The $2S_{1/2}$ and $2P_{1/2}$ states are degenerate, being the two eigenstates of opposite parity corresponding to the same n and j. The $2P_{3/2}$ state is higher in energy than the $2P_{1/2}$ state; the energy difference, $[m(Z\alpha)^4/32](1 + 0(Z\alpha)^2 + \cdot \cdot \cdot)$, is the fine-structure splitting due to the spin-orbit interaction, (4.6). In general, the state of larger j, for a given n, lies higher in energy according to (4.15).

How do these predictions agree with observations for the H atom? Prior to 1947 the agreement was completely satisfactory after the above predictions were modified to take into account the hyperfine splitting of each level due to coupling between the electron and proton
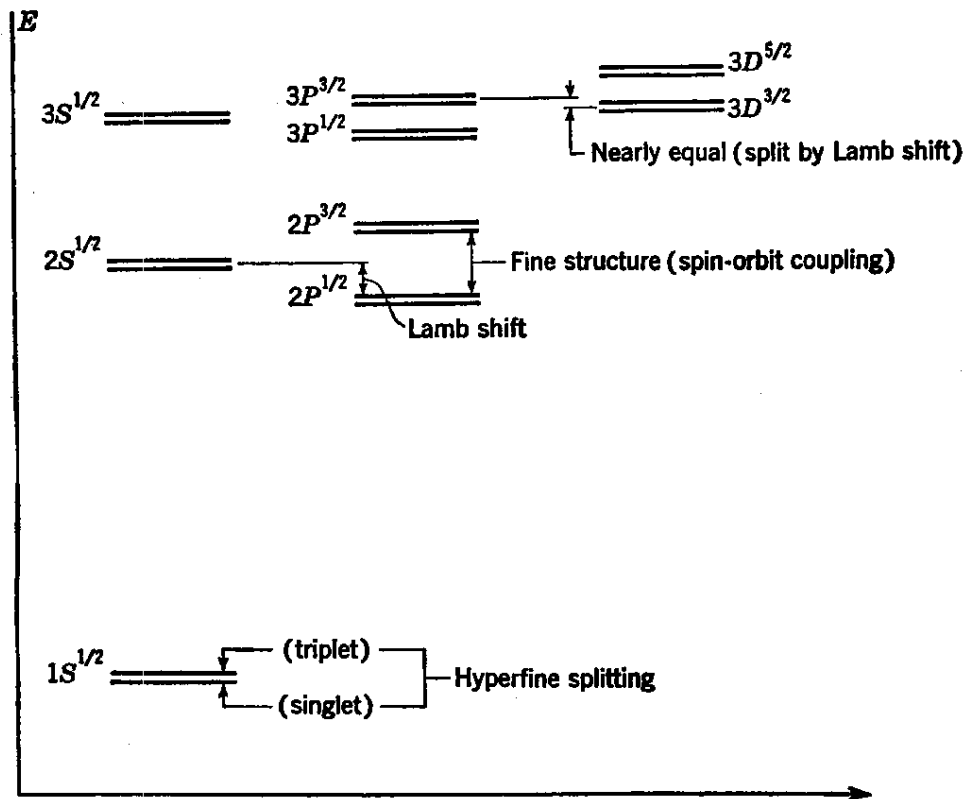
Fig. 4-2 Low-lying energy levels of atomic hydrogen. The diagram is not drawn to scale.

spins.   In 1947 the Lamb-Retherford measurements[1] of the H-atom fine structure confirmed an earlier suspicion of a shift of the $2S_{1/2}$ levels upward relative to the $2P_{1/2}$ lines.   This "Lamb shift," breaking the degeneracy of levels with the same n and j but differing l, arises from the interaction of the electrons with the fluctuations of the quantized radiation field.   Both the hyperfine structure splitting and the Lamb shift have been measured and calculated to a very high precision with good agreement.[2]

The hyperfine structure results from the interaction of the proton with the electron magnetic moment.[3]   This has the effect of splitting all lines into doublets corresponding to the two possible states of total angular momentum compounded from the j of the electron

[1] W. E. Lamb, Jr., and R. C. Retherford, *Phys. Rev.*, **72**, 241 (1957).   For references to subsequent work see Bethe and Salpeter, *op. cit.*; see also W. E. Lamb, Jr., *Repts. Progr. Phys.*, **14**, 19 (1951).

[2] For a review of the current situation, see R. P. Feynman, *Proc.* 1961 *Solvay Conf.*, Interscience, New York, 1962.

[3] E. Fermi, *Z. Physik*, **60**, 320 (1930); see also Bethe and Salpeter, *op. cit.*, p. 163.

system and the half-integer spin of the proton. Let us compute the magnitude of this effect for s states. For our purpose a nonrelativistic description of the electron suffices. The interaction is of the form

$$H' = + \frac{|e|}{2m} \, \boldsymbol{\sigma} \cdot \mathbf{B}$$

and

$$\mathbf{B} = \frac{g_p e}{2M_p} \int d^3r' \, \rho(r') \nabla \times (\mathbf{I} \times \nabla) \frac{1}{4\pi |\mathbf{r} - \mathbf{r'}|}$$

Here **I** is the proton spin operator ($I_z = \pm \frac{1}{2}$) and $\rho(r')$ is the magnetic moment density of the proton, owing to the fact it is not a point particle. Using the relations $\nabla \times (\mathbf{I} \times \nabla) = \mathbf{I} \nabla^2 - (\mathbf{I} \cdot \nabla) \nabla$ and taking the angular average for spherically symmetric wave functions so that

$$\nabla_i \nabla_j \Rightarrow \tfrac{1}{3} \delta_{ij} \nabla^2$$

we find

$$\mathbf{B} = \frac{2}{3} g_p \frac{e}{2M_p} \int d^3r' \, \rho(r') \mathbf{I} \, \nabla^2 \left(\frac{1}{4\pi |\mathbf{r} - \mathbf{r'}|} \right) = \frac{2}{3} g_p \frac{e}{2M_p} \mathbf{I} \, \rho(r)$$

The energy shift is then given, in nonrelativistic theory, by

$$\Delta E_n = (\psi_n H' \psi_n) = \frac{2}{3} \frac{g_p e^2}{4m M_p} \boldsymbol{\sigma} \cdot \mathbf{I} \int d^3r \, \psi_n^*(r) \rho(r) \psi_n(r)$$

$$\approx \frac{1}{6} \frac{g_p e^2}{m M_p} \boldsymbol{\sigma} \cdot \mathbf{I} |\psi_n(0)|^2$$

$$= \frac{1}{2} m\alpha^2 \left[\frac{4}{3} g_p \frac{Z^3 \alpha^2}{n^3} \left(\frac{m}{M_p} \right) \boldsymbol{\sigma} \cdot \mathbf{I} \right]$$

with

$$\boldsymbol{\sigma} \cdot \mathbf{I} = \begin{cases} +\frac{1}{2} & \text{triplet states} \\ -\frac{3}{2} & \text{singlet state} \end{cases}$$

The splitting δ_n of the nth s-state level is thus

$$\delta_n = \frac{1}{2} m\alpha^2 \left[\frac{8}{3} g_p \frac{Z^3 \alpha^2}{n^3} \left(\frac{m}{M_p} \right) \right]$$

and is reduced by the mass ratio m/M_p relative to the fine structure.

Welton[1] has given a simple qualitative description of the Lamb shift by considering the interaction of an electron, treated nonrelativistically, with the vacuum fluctuations of the electromagnetic field. Since the dynamics of a normal mode of the electromagnetic field is equivalent to that of a harmonic oscillator, each mode upon quantization acquires a zero-point energy of $\omega/2$. As a result of this quantum

[1] T. A. Welton, *Phys. Rev.*, **74**, 1157 (1948).

cs

ıe

ıc

effect there are now fluctuating electromagnetic fields even when no external fields are applied. Although the average field strengths are zero, their mean-square values are nonvanishing, and this leads to a mean-square fluctuation in the electron's position coordinate due to its coupling with the field. It is the amplitude of this jiggling of a bound electron in the hydrogen atom that we estimate. It implies, as we saw in our discussion of the origin of the Darwin term, (4.7), an additional interaction energy $\frac{1}{6}\langle(\delta r)^2\rangle\nabla^2 V$ from the smearing out of the Coulomb potential $V(r)$ seen by the electron. To lowest order, the change in the energy level for the electron due to this is then

g-
a
)∇
ıc-

$$\Delta E_n(\text{Lamb}) = \frac{1}{6}\langle(\delta r)^2\rangle\int\psi_n^*\nabla^2 V(r)\psi_n \, d^3r$$

$$= \frac{2\pi}{3}Z\alpha\langle(\delta r)^2\rangle|\psi_n(0)|^2 \tag{4.16}$$

To estimate $\langle(\delta r)^2\rangle$, we treat the electron classically and non-relativistically as a charged particle. Its equation of motion for oscillation about its equilibrium coordinate in the atom is $\delta\ddot{r} = \frac{e}{m}\mathbf{E}$, where $\mathbf{E}$ is the fluctuating electromagnetic field. For the ωth Fourier amplitude we have

$$\delta r_\omega = -\frac{eE_\omega}{m\omega^2}$$

and hence for its mean-square amplitude

$$\langle(\delta r_\omega)^2\rangle = \frac{e^2\langle(E_\omega)^2\rangle}{m^2\omega^4}$$

re.
nb
la-
ld.
is
za-
ım

and

$$\langle(\delta r)^2\rangle = \frac{e^2}{m^2}\int\frac{d\omega}{\omega^4}\langle(E_\omega)^2\rangle \tag{4.17}$$

To calculate the mean-square field strength, we consider the total vacuum field energy

$$\frac{1}{2}\int d^3x\,(E^2 + B^2) = \sum_{\lambda=1}^{2}\sum_{k}\frac{1}{2}\omega$$

where the two values of λ refer to the two states of transverse polarization and the sum extends over all modes in a large box of volume

$$L^3 = \int d^3x \qquad \sum_{k} \to \frac{L^3}{(2\pi)^3}\int d^3k$$

Since $\int d^3x\,E^2 = \int d^3x\,B^2$ and $\omega = |\mathbf{k}|$ for free electromagnetic waves,

the mean-square field strength in vacuo is

$$\langle E^2 \rangle = \frac{1}{L^3} \int E^2 \, d^3x = 2 \frac{1}{(2\pi)^3} \int d^3k \frac{\omega}{2} = \frac{1}{2\pi^2} \int \omega^3 \, d\omega = \int d\omega \, \langle E_\omega^2 \rangle$$

Inserting in (4.17), we find

$$\langle (\delta r^2) \rangle = \frac{e^2}{2\pi^2 m^2} \int \frac{d\omega}{\omega} \tag{4.18}$$

where the frequency integral extends from 0 to ∞. Because of the crudity of our approximate treatment of the electron, the integral diverges at both ends. This is not the case for an accurate relativistic treatment of the electron localized in a hydrogen atom. Wavelengths larger than the Bohr radius $\sim (Z\alpha m)^{-1}$ will not be effective, since there must be a minimum frequency for the induced oscillations corresponding to this typical atomic size; therefore, $\omega_{min} \sim mZ\alpha$. There is also a high-frequency cutoff at distances $\sim$ the electron Compton wavelength $1/m$ coming from the relativistic structure of the electron. This structure corresponding to the zitterbewegung amplitude suggests that frequencies higher than $\omega_{max} \sim m$ will not be effective in jiggling the electron. Hence we approximate $\int d\omega/\omega \sim \ln (1/Z\alpha)$ and find for the mean-square amplitude of the oscillations in the vacuum field, by (4.18),

$$\langle (\delta r)^2 \rangle = \left(\frac{2\alpha}{\pi} \ln \frac{1}{Z\alpha} \right) \left(\frac{1}{m} \right)^2 \tag{4.19}$$

The resulting energy shift is by (4.16)

$$\Delta E_n = \frac{4Z\alpha^2}{3} \left(\frac{1}{m} \right)^2 \ln \frac{1}{Z\alpha} |\psi_n(0)|^2$$

$$= \left[\frac{8}{3\pi} \frac{Z^4\alpha^3}{n^3} \left(\ln \frac{1}{Z\alpha} \right) \right] (\tfrac{1}{2}\alpha^2 m) \delta_{l0}$$

$$\cong 1{,}000 \text{ mc/sec for } n = 2, Z = 1, l = 0$$

This accounts for most of the measured shift of the $2S_{\frac{1}{2}}$ level in the hydrogen atom; for the p and higher l states the shifts are not precisely zero but are much smaller because the wave functions at the origin are zero. By way of comparison with the ordinary fine structure we see by looking back at the hamiltonian (4.5) that the ratio of the Lamb term to the Darwin term is $(8\alpha/3\pi)[\ln (1/Z\alpha)]$ corresponding to the ratio of the mean-square fluctuation amplitude (4.19) to the zitterbewegung structure $\simeq (1/m)^2$.

The Foldy-Wouthuysen transformation   61

Problems

1. Derive (4.10).

2. The Dirac equation describing the interaction of a proton or neutron with an
applied electromagnetic field will have an additional magnetic moment interaction
representing their observed anomalous magnetic moments:

$$\left(i\overline{\nabla} - e_i A + \frac{\kappa_i e}{4M_i}\, \sigma_{\mu\nu} F^{\mu\nu} - M_i \right) \psi(x) = 0$$

where

$$F^{\mu\nu} = \frac{\partial}{\partial x_\nu}\, A^\mu - \frac{\partial}{\partial x_\mu}\, A^\nu$$

represents the field strengths as defined in Appendix 1.

a. For the proton, $i = p$, $e_p = |e|$; for the neutron $i = n$, $e_n = 0$. Verify that the
choice of $\kappa_p = 1.79$ and $\kappa_n = -1.91$ corresponds to the observed magnetic moments
and check that the additional interaction does not disturb the Lorentz covariance
of the equation. Check also that the Dirac hamiltonian is hermitian and that
probability is conserved in the presence of the additional interaction.

b. Make a Foldy-Wouthuysen transformation for the neutron, keeping terms to
the accuracy of (4.5), and give a physical interpretation of the individual terms.
Calculate the cross section for the scattering of a slow neutron by an applied
electrostatic field. How might this be measured?   [See L. L. Foldy, *Rev. Mod.
Phys.*, **30**, 471 (1958).]

3. Solve for the exact energy eigenvalues and eigenfunctions of a Dirac electron in
a uniform static magnetic field. [See L. D. Huff, *Phys. Rev.*, **38**, 501 (1931);
M. H. Johnson and B. A. Lippman, *Phys. Rev.*, **77**, 702 (1950).]

4. Calculate to lowest order in α^2 the first-order Zeeman effect for an electron in a
hydrogen atom.   If the electron gyromagnetic ratio differs from $g = 2$, how are the
Zeeman levels altered (to first order in the difference $g - 2$)?

5. Discuss the precession of the spin of a charged Dirac particle with an anomalous
magnetic moment κ in an applied static magnetic field.   Show in particular that
the difference in the spin and orbital precession frequencies is proportional to
$g - 2$, or κ.   How does it depend upon the mass of the particle?   See:

H. A. Tolhoek and S. R. de Groot, *Physica*, **17**, 17 (1951).
K. M. Case, *Phys. Rev.*, **106**, 173L (1957).
H. Mendlowitz and K. M. Case, *Phys. Rev.*, **97**, 33 (1955).
M. Carrassi, *Nuovo Cimento*, **7**, 524 (1958).
V. Bargmann, L. Michel, and V. L. Telegdi, *Phys. Rev. Letters*, **2**, 435 (1959).
Louisell, Pidd, and Crane, *Phys. Rev.*, **94**, 7 (1954).
Schupp, Pidd, and Crane, *Phys. Rev.*, **121**, 1 (1961).
Charpak, Farley, Garwin, Muller, Sens, Telegdi, and Zichichi, *Phys. Rev. Letters*, **6**,
128 (1961.)

6. Construct an additional interaction term to represent a possible anomalous electric dipole moment of a Dirac particle. What happens to the parity transformation? What is the effect of such a term on the hydrogen-atom energy levels? [See G. Feinberg, *Phys. Rev.*, **112**, 1637 (1958); E. E. Salpeter, *Phys. Rev.*, **112**, 1642 (1950).]

7. Owing to meson effects (discussed in Chap. 10), the proton charge is distributed over a small region of spatial extent $\sim 10^{-13}$ cm. Compute the effect on the hydrogen-atom energy levels of such a charge distribution with mean square radius $r \cong 0.8 \times 10^{-13}$ cm. Compare the result with the Lamb shift.