14 The Dirac equation

In this section we will discuss the celebrated Dirac equation. Traditionally, the quantum mechanics II course begins with the "derivation" of this equation from the requirement that the wave equation be both, relativistically invariant and linear in time derivative (as originally noted by Dirac, the latter requirement is necessary for the standard probabilistic interpretation of the wave function, that is, for the existence of a positive definite probability density satisfying the continuity equation). Then one spends a lot of time fighting with obvious defects of the equation such as negative energy states or the so-called Klein paradox and explaining strange notions like the Dirac sea, particles moving backward in time etc. Finally, one formulates the Feynman rules either by using hand-waving arguments or by starting anew and quantizing fermionic fields (or the Dirac wave function as if it were a physical field). You can find all this e.g. in the monographs of Bjorken and Drell and Schiff. In this course we followed the other way: we have constructed a fully relativistic theory of interacting particles first and only now we will show where is the place of the Dirac equation in this construction. This seems more logical and moreover stresses the fact that the Dirac equation is only an effective description useful in some special limiting situations. Its defects are therefore neither surprising nor dangerous - they arise because it is not possible to construct a relativistic theory for a single interacting particle: consistent relativistic approach requires that there be processes in which the number of particles is not conserved.

14.1 Free Dirac equation

For spin 1/2 particles we have constructed in the previous section the field operators of the form

$$\begin{aligned}\psi_\alpha(x) &= \int d\Gamma_{\mathbf{p}} \sum_\sigma \left[u_\alpha(\mathbf{p},\sigma)e^{-ip\cdot x}b(\mathbf{p},\sigma) + v_\alpha(\mathbf{p},\sigma)e^{+ip\cdot x}d^\dagger(\mathbf{p},\sigma)\right] \\ \bar{\psi}_\alpha(x) &= \int d\Gamma_{\mathbf{p}} \sum_\sigma \left[\bar{v}_\alpha(\mathbf{p},\sigma)e^{-ip\cdot x}d(\mathbf{p},\sigma) + \bar{u}_\alpha(\mathbf{p},\sigma)e^{+ip\cdot x}b^\dagger(\mathbf{p},\sigma)\right] \quad (14.1)\end{aligned}$$

We have also found that the functions $u_\alpha(\mathbf{p},\sigma)$ and $v_\alpha(\mathbf{p},\sigma)$ satisfy the equations

$$(\not{p} - m)\, u(\mathbf{p},\sigma) = 0, \qquad \bar{u}(\mathbf{p},\sigma)\,(\not{p} - m) = 0,$$

$$(\not{p} + m)\, v(\mathbf{p}, \sigma) = 0, \qquad \bar{v}(\mathbf{p}, \sigma)\, (\not{p} + m) = 0,$$

Due to them the field operators $\psi(x)$ and $\bar{\psi}(x)$ satisfy the Dirac equations

$$(i\, \not{\partial} - m)\, \psi(x) = 0\ , \qquad \bar{\psi}(x) \left(i\, \overleftarrow{\not{\partial}} + m\right) = 0\ . \tag{14.2}$$

We can now compare the structure of the field operator $\psi(x)$ in eq. (14.1) with the field operator $a(\mathbf{x}, t)$ in the nonrelativistic second quantization formalism. In the interaction picture we have

$$a(\mathbf{x}, t) = e^{iH_0 t/\hbar} a(\mathbf{x}) e^{-iH_0 t/\hbar} = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\ e^{-iE_{\mathbf{p}} t/\hbar + i\mathbf{p}\cdot\mathbf{x}}\ a(\mathbf{p}) \tag{14.3}$$

where $E_{\mathbf{p}} = \hbar^2 \mathbf{p}^2/2m$. The second quantization operator $a(\mathbf{x}, t)$ satisfies the 1-particle free Schrödinger equation

$$\left(i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \boldsymbol{\nabla}^2\right) a(\mathbf{x}, t) = 0 \tag{14.4}$$

because the function multiplying $a(\mathbf{p})$ under the integral is the plane wave solution to the 1-particle free Schrödinger equation. In the same way, we can interpret the functions

$$u_\alpha(\mathbf{p}, \sigma)\ e^{-ip\cdot x}, \qquad v_\alpha(\mathbf{p}, \sigma)\ e^{+ip\cdot x}\ , \tag{14.5}$$

as the plane wave solutions of the wave equation which we may try to interpret as the equation defining the relativistic 1-particle Hamiltonian quantum mechanics.[86] In what follows it will be important to remember that in the functions (14.5) $p^0 = E_{\mathbf{p}}$ (where $E_{\mathbf{p}}$ is always positive: $E_{\mathbf{p}} \equiv +\sqrt{\mathbf{p}^2 + m^2}$).

Much in the same way we can interpret the coefficient of $a(\mathbf{p})$ in $\phi^{(+)}(x)$ as the wave function of a relativistic spin 0 particle having the three-momentum $\mathbf{p}$. In other words, the wave function can be identified with the matrix element

$$\phi_{\mathbf{p}}(\mathbf{x}, t) \equiv \langle\Omega|\phi(x) a^\dagger(\mathbf{p})|\Omega\rangle\ . \tag{14.6}$$

[86] For spin 0 particles the equation satisfied by the field operators is the Klein-Gordon equation which does not define a *Hamiltonian* quantum mechanics. In this sense, at low energies there emerges a conceptual distinction between matter spin $1/2$ particles, for which the *Hamiltonian* quantum mechanics can be constructed, and integer spin particles which in this limit are replaced by classical fields.

As the field operator $\phi(x)$ satisfies the Klein-Gordon equation

$$(\partial_\mu \partial^\mu + m^2)\phi(x) = 0 \tag{14.7}$$

so does the wave function $\phi_{\mathbf{p}}(\mathbf{x}, t)$.

We will therefore consider the Dirac equation as a "good" relativistic wave equation for the wave function $\psi(x)$ (now $\psi(x)$ is the ordinary c-number function, not an operator). In the original, non-explicitly covariant, form it reads

$$i\hbar \frac{\partial}{\partial t}\psi(t, \mathbf{x}) = \left(-i\hbar c\ \alpha^i \frac{\partial}{\partial x^i} + \beta mc^2\right)\psi(t, \mathbf{x})\ , \tag{14.8}$$

where we have defined the matrices

$$\alpha^i \equiv \gamma^0 \gamma^i, \qquad \beta \equiv \gamma^0\ . \tag{14.9}$$

(their explicit forms in two different representations are given in the previous section). The RHS of (14.8) can be interpreted as the free Hamiltonian

$$H_0^{\mathrm{Dir}} = -i\hbar c\ \alpha^i \frac{\partial}{\partial x^i} + \beta mc^2 \tag{14.10}$$

acting on a four-component wave function, which describes the state of a single spin 1/2 particle. Moreover, the wave function $\psi(t, \mathbf{x})$ can be given probabilistic interpretation as it is possible to define the probability density

$$\rho(t, \mathbf{x}) = \sum_\alpha \psi_\alpha^* \psi_\alpha \equiv \psi^\dagger \psi = \bar{\psi}\gamma^0 \psi \tag{14.11}$$

and the probability current

$$\mathbf{j}(t, \mathbf{x}) = \sum_{\gamma\beta} \psi_\gamma^* c \boldsymbol{\alpha}_{\gamma\beta} \psi_\beta = c\bar{\psi}\boldsymbol{\gamma}\psi \tag{14.12}$$

which satisfy the continuity equation

$$\frac{\partial}{\partial t}\rho(t, \mathbf{x}) + \boldsymbol{\nabla} \cdot \mathbf{j}(t, \mathbf{x}) = 0\ . \tag{14.13}$$

Indeed, if we define the four-vector $j^\mu = (c\rho,\ \mathbf{j}) = c\bar{\psi}\gamma^\mu\psi$, the equation (14.13) can be rewritten as the four-current conservation condition $\partial_\mu j^\mu = 0$. Using then the equations (14.2) we easily verify that

$$\partial_\mu j^\mu \propto \partial_\mu(\bar{\psi}\gamma^\mu\psi) = (\bar{\psi}\ \overleftarrow{\not\partial})\psi + \bar{\psi}\ \not\partial\psi = -m\bar{\psi}\psi + m\bar{\psi}\psi = 0\ . \quad (14.14)$$

In our fully relativistic approach Lorentz covariance of the Dirac equation is obvious. We will however restate it here in the form which is encountered in standard textbooks. Recall that the field operator $\psi(x)$ has been constructed by requiring that

$$\psi'_\alpha(x) \equiv U_0(\Lambda, a)\psi_\alpha(x)U_0^{-1}(\Lambda, a) = D_{\alpha\beta}(\Lambda^{-1})\psi_\beta(\Lambda x - a)\ , \quad (14.15)$$

or, taking Λ^{-1} instead of Λ and $-\Lambda^{-1}a$ instead of a,

$$\psi'_\alpha(x) = D_{\alpha\beta}(\Lambda)\psi_\beta(\Lambda^{-1}(x + a)) \quad (14.16)$$

In other words,

$$\psi'_\alpha(x') = D_{\alpha\beta}(\Lambda)\psi_\beta(x) \quad (14.17)$$

where $x' = \Lambda x - a$. The same transformation rule is valid for the c-number wave function [87] satisfying the Dirac equation (14.2): if the wave function $\psi(x)$ satisfies $(i\ \not\partial - m)\psi(x) = 0$, then using (14.17) we can write

$$(i\gamma^\mu\partial_\mu - m)\, D^{-1}(\Lambda)\psi'(x') = 0\ . \quad (14.18)$$

Multiplying this from the left by D and using the basic property of $D(\Lambda)$ (see the preceding section)

$$D(\Lambda)\gamma^\mu D^{-1}(\Lambda) = \gamma^\nu \Lambda_\nu{}^\mu\ , \qquad \Lambda_\nu{}^\mu \frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial x'^\nu} \quad (14.19)$$

we find that the transformed wave function $\psi'(x')$ satisfies

$$\left(i\gamma^\mu\partial'_\mu - m\right)\psi'(x') = 0\ . \quad (14.20)$$

that is, the Dirac equation in the new Lorentz frame. Thus, the Dirac equation has the same form in all Lorentz frames. For $\bar{\psi}$ the transformation rule is

$$\bar{\psi}'(x') = \psi^\dagger(x)D^\dagger(\Lambda)\gamma^0 = \bar{\psi}(x)\gamma^0 D^\dagger(\Lambda)\gamma^0 = \bar{\psi}(x)D^{-1}(\Lambda)\ . \quad (14.21)$$

[87] In standard textbooks the matrix $D_{\alpha\beta}(\Lambda)$ is usually denoted $S_{\alpha\beta}(\Lambda)$. As it could lead to confusion with the S matrix introduced earlier, we retain the symbol D.

(we have used the pseudo-unitarity of the representation $D(\Lambda)$). It is therefore easy to check that the following rule holds

$$j'^{\mu}(x') = \bar{\psi}(x)D^{-1}(\Lambda)\gamma^{\mu}D(\Lambda)\psi(x) = (\Lambda^{-1})_{\nu}{}^{\mu}j^{\nu}(x) = \Lambda^{\mu}{}_{\nu}j^{\nu}(x) \ . \tag{14.22}$$

In the similar way, any object with n Lorentz indices constructed as a string of gamma matrices inserted between $\bar{\psi}$ and ψ will transform as a n-th rank Lorentz tensor.

We have also ensured that with respect to parity transformation the field operator $\psi(x)$ has the property

$$\psi'(x) \equiv \mathcal{P}\psi(x)\mathcal{P}^{-1} = \eta\beta\psi(Px) \ , \tag{14.23}$$

where η is the intrinsic parity of the particle annihilated by $\psi(x)$. Therefore, for the wave functions we also have

$$\psi'(Px) = \eta\beta\psi(x) \ . \tag{14.24}$$

We then find that

$$\bar{\psi}'(Px)\gamma^{\mu}\psi'(Px) = \bar{\psi}(x)\beta\gamma^{\mu}\beta\psi(x) = \begin{cases} \bar{\psi}(x)\gamma^{0}\psi(x) & \text{for } \mu = 0 \\ -\bar{\psi}(x)\gamma^{i}\psi(x) & \text{for } \mu = i \end{cases}$$

which shows that $\bar{\psi}\gamma^{\mu}\psi$ behaves under space reflection as a vector. Similarly, it is easy to see that $\bar{\psi}\gamma^{\mu}\gamma^{5}\psi$ behaves as a pseudovector (axial vector). Likewise, $\bar{\psi}\psi$ and $\bar{\psi}\gamma^{5}\psi$ are a scalar and a pseudoscalar, respectively. In general, insertion of an odd number of γ^{5}'s makes a quantity a Lorentz pseudo-tensor.

We have therefore a sort of a relativistic Hamiltonian quantum mechanics with the Hamiltonian (14.8) which can be rewritten in the form:

$$H_0^{\mathrm{Dir}} = c\ \boldsymbol{\alpha}\cdot\hat{\mathbf{p}} + \beta\ mc^2 \ , \tag{14.25}$$

(where the hat over $\hat{\mathbf{p}}$ reminds that this is an operator) and we can analyze it in the same way as we did it in the case of nonrelativistic quantum mechanics. Obviously H_0^{Dir} commutes with the momentum operator $\hat{\mathbf{p}}$ and they can be simultaneously diagonalized. Consider first the energy spectrum of states which have zero eigenvalue of the momentum operator $\hat{\mathbf{p}}$. It will be convenient to work in the representation, in which the matrix β is diagonal:

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \ . \tag{14.26}$$

For $\mathbf{p} = 0$ the eigenvectors and eigenvalues of $H_0^{\rm Dir}$ are determined by the equation

$$mc^2 \beta \psi^r = E_r \psi^r \ , \tag{14.27}$$

where r labels different solutions. The equation (14.27) has 4 independent solutions (for the lack of space $2mc^2$ is written as $2m$)

$$\psi^1 = \sqrt{2m}\begin{pmatrix}1\\0\\0\\0\end{pmatrix} , \ \ \psi^2 = \sqrt{2m}\begin{pmatrix}0\\1\\0\\0\end{pmatrix} , \ \ \psi^3 = \sqrt{2m}\begin{pmatrix}0\\0\\1\\0\end{pmatrix} , \ \ \psi^4 = \sqrt{2m}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} ,$$

(we have adopted the relativistic normalization of the eigenvectors) with the corresponding $H_0^{\rm Dir}$ eigenvalues

$$E_1 = E_2 = mc^2 \ , \qquad\qquad E_3 = E_4 = -mc^2 \ . \tag{14.28}$$

As we see, the Hamiltonian has negative eigenvalues which apparently correspond to some states of a free particle. (Remember, we are now investigating to which extent the Dirac equation can be interpreted as defining the relativistic Hamiltonian 1-particle quantum mechanics; in our original theory all free particle and antiparticle states had positive energies!).

It is easy to see that for nonzero eigenvalues $\mathbf{p}$ of the momentum operator the Hamiltonian (14.25) gives the relativistic relation between the energy and momentum. This follows from the fact that $(H_0^{\rm Dir})^2 = (c^2\ \hat{\mathbf{p}}^2 + m^2c^4)I_{4\times 4}$ which can be easily checked by using the relations

$$\left\{\alpha^i,\ \alpha^j\right\} = 2\delta^{ij} \ , \quad \left\{\alpha^i,\ \beta\right\} = 0 \ , \tag{14.29}$$

following from the anticommutation relations (12.34) and the definitions (14.9). Since for $\mathbf{p} = 0$ there are Hamiltonian eigenstates with negative energies, one expects that also for any $\mathbf{p} \neq 0$ there are solutions with energies equal $\pm\sqrt{c^2\hat{\mathbf{p}}^2 + m^2c^4}$.

To find the explicit form of all Hamiltonian eigenstates it will be convenient to introduce

$$\varepsilon_r = \begin{cases} +1 & \text{for} \quad r = 1,2 \\ -1 & \text{for} \quad r = 3,4 \end{cases} \tag{14.30}$$

and to write the $\hat{\mathbf{p}}$ eigenvectors as

$$\psi^r(t,\mathbf{x}) = w^r(\mathbf{p})\ e^{-i\varepsilon_r E_{\mathbf{p}} t/\hbar} e^{+i\varepsilon_r \mathbf{p}\cdot\mathbf{x}}\ . \tag{14.31}$$

Note that for $r = 1, 2$ these wave functions are $\hat{\mathbf{p}}$ eigenvectors with the eigenvalues $+\mathbf{p}$ but for $r = 3, 4$ they are $\hat{\mathbf{p}}$ eigenvectors with the eigenvalues $-\mathbf{p}$! Inserting these functions in the equation

$$\left(c\ \boldsymbol{\alpha}\cdot\hat{\mathbf{p}} + \beta mc^2\right)\psi^r(t,\mathbf{x}) = \varepsilon_r E_{\mathbf{p}}\psi^r(t,\mathbf{x}) \tag{14.32}$$

(where now by definition $E_{\mathbf{p}} > 0$) we find that the spinors $w^r(\mathbf{p})$ written in the form

$$w^r(\mathbf{p}) = \begin{pmatrix} \varphi^r \\ \chi^r \end{pmatrix} \tag{14.33}$$

must satisfy the equation

$$\begin{pmatrix} \varepsilon_r mc^2 - E_{\mathbf{p}} & c\ \boldsymbol{\sigma}\cdot\mathbf{p} \\ c\ \boldsymbol{\sigma}\cdot\mathbf{p} & -\varepsilon_r mc^2 - E_{\mathbf{p}} \end{pmatrix}\begin{pmatrix} \varphi^r \\ \chi^r \end{pmatrix} = 0\ . \tag{14.34}$$

The solutions can be now easily found: for $r = 1, 2$ we write

$$\chi^r = \frac{c\ \boldsymbol{\sigma}\cdot\mathbf{p}}{E_{\mathbf{p}} + mc^2}\varphi^r\ , \tag{14.35}$$

and for $r = 3, 4$

$$\varphi^r = \frac{c\ \boldsymbol{\sigma}\cdot\mathbf{p}}{E_{\mathbf{p}} + mc^2}\chi^r\ . \tag{14.36}$$

We have then

$$w^{1,2}(\mathbf{p}) = \begin{pmatrix} \varphi^{1,2} \\ \\ \frac{c\boldsymbol{\sigma}\cdot\mathbf{p}}{E_{\mathbf{p}}+mc^2}\varphi^r \end{pmatrix}\ , \qquad w^{3,4}(\mathbf{p}) = \begin{pmatrix} \frac{c\boldsymbol{\sigma}\cdot\mathbf{p}}{E_{\mathbf{p}}+mc^2}\chi^r \\ \\ \chi^r \end{pmatrix}\ , \tag{14.37}$$

where $\varphi_{1,2}$ and $\chi_{3,4}$ are two arbitrary pairs of linearly independent two-component spinors. Indeed, it is easy to check that for $r = 1, 2$ ($r = 3, 4$) the equation (14.34) is satisfied for arbitrary $\varphi_{1,2}$ ($\chi_{3,4}$) provided $\chi_{1,2}$ ($\varphi_{3,4}$) is given by the relation (14.35) (by the relation (14.36)). Of course we could construct four H_0^{Dir} eigenstates which have the same eigenvalue $+\mathbf{p}$ of the momentum operator $\hat{\mathbf{p}}$ (it suffices to change $\mathbf{p} \to -\mathbf{p}$

in solutions with $r = 3, 4$) but as we will see there is a deeper reason for collecting together solutions with H_0^{Dir} and $\hat{\mathbf{p}}$ eigenvalues $(E_{\mathbf{p}}, +\mathbf{p})$ with solutions having eigenvalues $(-E_{\mathbf{p}}, -\mathbf{p})$.

For given eigenvalues of the H_0^{Dir} and $\hat{\mathbf{p}}$ operators there is still a twofold degeneracy related to the spin variable and we need an operator commuting with H_0^{Dir} and $\hat{\mathbf{p}}$ to complete the set of observables. Let us construct the spin operator to see if it can play this role. To this end let us specify the transformation rule (14.17) to the spatial rotations $\Lambda = R$. We have then

$$\psi'(x) = D(R)\psi(R^{-1}x) = e^{\frac{i}{\hbar}\omega_{\mu\nu}\mathcal{J}^{\mu\nu}}\psi(R^{-1}x) = e^{\frac{i}{\hbar}\phi^i\mathcal{J}^i}\psi(R^{-1}x) \quad (14.38)$$

where $\mathcal{J}^{\mu\nu}$ is defined in eq. (12.35), and proceeding as in Section 4.3 (eqs. (4.61)-(4.64)) we identify the total angular momentum operator with

$$\hat{\mathbf{J}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} + \frac{\hbar}{2}\mathbf{\Sigma} \ , \quad (14.39)$$

where the spin operator $(\hbar/2)\Sigma^i = (\hbar/2)\epsilon^{ijk}\mathcal{J}^{jk}$ is

$$\frac{\hbar}{2}\Sigma^i = \frac{\hbar}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} \ . \quad (14.40)$$

It is easy to see that the spin operator $\mathbf{\Sigma}$ alone does not commute with the free Dirac Hamiltonian (14.25) and cannot be used in general to label different states having the same eigenvalues of H_0^{Dir} and $\hat{\mathbf{p}}$. However, the spin operator $\mathbf{\Sigma}$ does commute with H_0^{Dir} on the $\mathbf{p} = 0$ subspace and therefore, can be used to classify the states of a particle at rest. (Note that this is in agreement with the considerations of section 9 where we have shown that for massive particles the spin variable distinguishes different states when the four-momentum vector takes the standard form, that is in the frame, in which the massive particle is at rest.) It is also easy to verify that the total angular momentum operator (14.39) does commute with H_0^{Dir} but it does not commute with the momentum operator $\hat{\mathbf{p}}$. For $\mathbf{p} \neq 0$ there is however another operator, called the helicity operator[88]

$$\hat{h} = \frac{\hat{\mathbf{J}}\cdot\hat{\mathbf{p}}}{|\hat{\mathbf{p}}|} = \frac{\hbar}{2}\frac{\mathbf{\Sigma}\cdot\hat{\mathbf{p}}}{|\hat{\mathbf{p}}|} \tag{14.41}$$

which commutes both with $H_0^{\rm Dir}$ and $\hat{\mathbf{p}}$. One can therefore construct common eigenstates of $H_0^{\rm Dir}$, $\hat{\mathbf{p}}$ and $\hat{h}$. They are simple only for $\mathbf{p} = (0, 0, p_z)$:

$$w^1 \propto \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{m+E} \\ 0 \end{pmatrix}, \quad w^2 \propto \begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{p_z}{m+E} \end{pmatrix}, \quad w^3 \propto \begin{pmatrix} \frac{p_z}{m+E} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad w^4 \propto \begin{pmatrix} 0 \\ \frac{p_z}{m+E} \\ 0 \\ 1 \end{pmatrix}.$$

The functions (14.31) are then eigenfunctions of the helicity operator with eigenvalues

$$h_1 = +\frac{1}{2}, \qquad h_2 = -\frac{1}{2}, \qquad h_3 = -\frac{1}{2}, \qquad h_4 = +\frac{1}{2}, \qquad .$$

For arbitrary momentum $\mathbf{p}$ the helicity spinors can be also constructed but have a rather complicated form.

Alternatively one takes in (14.37)

$$\phi^1 = \chi^3 = \mathcal{N}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \phi^2 = \chi^4 = \mathcal{N}\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{14.42}$$

where $|\mathcal{N}| = \sqrt{E + mc^2}$ so that (as is easy to check) $w^{r\dagger}w^r = 2E_{\mathbf{p}}$. It is then possible to show, that

$$w^r_\alpha(\mathbf{p}) = D_{\alpha\beta}\left(-\frac{\mathbf{p}}{E_{\mathbf{p}}}\right) w^r_\beta(\mathbf{0}) \ , \tag{14.43}$$

where $D_{\alpha\beta}(\mathbf{v})$ is the matrix representation of the boost to the frame moving with the velocity $\mathbf{v}$. Therefore, the spinors $w^{1,3}$ ($w^{2,4}$) are the wave functions of the states which in the particle rest frame would correspond to eigenstates of the third component of the spin operator (14.40) with

[88]Obviously, for fixed $\hat{\mathbf{p}}$ the helicity operator $\hat{h}$ is given by $\hat{h} = (1/p^0)W^0$ where the four-vector W^μ defined in (10.25). Since W^μ commutes with P^ν, eigenvalues of $\hat{h}$ could in principle be also used to label the one-particle states; however $\hat{h}$ is not defined for particles at rest and for this reason we did not use it in Section 10.2.

eigenvalue $+\hbar/2$ $(-\hbar/2)$. Notice that this is precisely the construction, we have used in the abstract way in section 9 to give a meaning to the spin labels σ in an arbitrary frame; the boost $D_{\alpha\beta}(\mathbf{v})$ plays here the role of the standard Lorentz transformation L_p.

Because $w^{1,2}(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}$ and $w^{3,4}(-\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}$ are the four linearly independent eigenvectors of the momentum operator $\hat{\mathbf{p}}$ with the same eigenvalue $\mathbf{p}$, they satisfy the completeness and orthogonality relations:

$$\begin{aligned}
&\sum_{r=1}^{4} w^r_\alpha(\varepsilon_r\mathbf{p})w^{r\dagger}_\beta(\varepsilon_r\mathbf{p}) = 2E_\mathbf{p}\delta_{\alpha\beta} \ , \\
&\sum_{\alpha=1}^{4} w^{r\dagger}_\alpha(\varepsilon_r\mathbf{p})w^{r'}_\alpha(\varepsilon_{r'}\mathbf{p}) = 2E_\mathbf{p}\delta_{rr'} \ .
\end{aligned} \tag{14.44}$$

Another two important relations are

$$\begin{aligned}
&\sum_{r=1}^{4} \varepsilon_r w^r_\alpha(\mathbf{p})\bar{w}^r_\beta(\mathbf{p}) = 2mc^2\delta_{\alpha\beta} \ , \\
&\sum_{\alpha=1}^{4} \bar{w}^r_\alpha(\mathbf{p})w^{r'}_\alpha(\mathbf{p}) = 2mc^2\delta_{rr'}\varepsilon_r \ .
\end{aligned} \tag{14.45}$$

They can be derived by noticing first that they are satisfied in the rest frame and then using the boost (14.43) and pseudounitarity of the representation $D_{\alpha\beta}$.

The spinors ϕ^r and χ^r can be also chosen in such a way, that in the rest frame the corresponding spinors w^r are eigenvectors of the spin projection onto the direction $\mathbf{s}$ (where $\mathbf{s}^2 = 1$), that is the eigenvectors with the eigenvalue $+\hbar/2$ of the operator

$$\frac{\hbar}{2}\mathbf{s}\cdot\mathbf{\Sigma} \ . \tag{14.46}$$

(Spinors $w^{1,3}$ and $w^{2,4}$ are eigenvectors of the projection of the spin operator on the direction $\mathbf{u}_z = (0,0,1)$ and $-\mathbf{u}_z = (0,0,-1)$, respectively.)

The existence of states which can have arbitrarily large negative energy (the value of $\mathbf{p}^2$ can be arbitrarily large) is the main problem of the Dirac equation. In quantum mechanics one cannot simply reject such solutions because the completeness of the Hamiltonian eigenstates is essential. Moreover, if the particle described by the Dirac Hamiltonian has a nonzero electric charge, its interaction with the (quantized) electromagnetic field would lead to rapid spontaneous transitions to states having lower and lower energy with emission of arbitrarily energetic photons. (The probabilities of such transitions can be estimated using the semi-classical radiation theory described in section 3.) To save the situation Dirac proposed that in a normal situation all negative energy states are occupied by electrons which form the so-called "Dirac sea". The infinite negative energy and electric charge of the sea are simply not measurable; only the deviations from the normal state can be observed. Because spin 1/2 particles are fermions, a particle occupying initially a positive energy state cannot make transition to an occupied negative energy state and the problem of matter stability is "solved". (With this interpretation the Dirac equation inadvertently loses its 1-particle character.)

This idea led Dirac to the crucial prediction of the existence of antiparticles. His heuristic reasoning was as follows. A photon of energy $E > 2mc^2$ can excite an electron occupying initially a negative energy state so that the electron makes a transition to one of the unoccupied positive energy states (the total energy is of course conserved). As a result of this process there appears an electron in a positive energy state and, at the same time, an electron is lacking in the Dirac sea. The hole in the Dirac sea, that is lack of an electron of electric charge $-e$ in the state characterized by the $H_0^{\rm Dir}$, $\hat{\mathbf{p}}$ and $(\hbar/2)\mathbf{s}\cdot\mathbf{\Sigma}$ eigenvalues[89] $-E_{\mathbf{p}}$, $-\mathbf{p}$ and $-\hbar/2$ is perceived by an external observer as the presence of a particle of electric charge $+e$, energy $E_{\mathbf{p}}$, momentum $\mathbf{p}$ and the spin projection on the direction $\mathbf{s}$ equal $+\hbar/2$ and, hence, can be interpreted as the positron. Thus, in the Dirac interpretation the negative energy states are related to the states of antiparticles. As we will show later, there is a transformation, the so-called charge conjugation transformation (which however should not be confused with the true charge conjugation operation realized in the multiparticle Hilbert space by a quantum operator $\mathcal{C}$ - unfortunately I had no time to discuss this in more detail) that transforms negative energy solutions of the Dirac equation for a charge $-e$ particle (interacting with an external electromagnetic field A_μ - see the next subsection) into the positive energy solutions of the Dirac equation for a charge $+e$ particle.

[89] The eigenvalue of the spin operator is ascribed to the state via the appropriate boost as explained above.

To comply fully with the Dirac interpretation we define new spinors $u(\mathbf{p}, s)$ and $v(\mathbf{p}, s)$ as follows:

- $u(\mathbf{p}, s)e^{-ip\cdot x}$ is the eigenvector of H_0^{Dir}, $\hat{\mathbf{p}}$ with eigenvalues $E_{\mathbf{p}}$ and $+\mathbf{p}$ which, when transformed to the rest frame by using the standard boost, is the eigenvector of the operator $(\hbar/2)\mathbf{s}\cdot\mathbf{\Sigma}$ with the eigenvalue $+\hbar/2$.

- $v(\mathbf{p}, s)e^{+ip\cdot x}$ is the eigenvector of H_0^{Dir}, $\hat{\mathbf{p}}$ with eigenvalues $-E_{\mathbf{p}}$ and $-\mathbf{p}$ which, when transformed to the rest frame by using the standard boost, is the eigenvector of the operator $(\hbar/2)\mathbf{s}\cdot\mathbf{\Sigma}$ with the eigenvalue $-\hbar/2$.

For $s = u_z$ we have therefore

$$\begin{aligned} u(\mathbf{p}, u_z) &= w^1(\mathbf{p}), \qquad u(\mathbf{p}, -u_z) = w^2(\mathbf{p}), \\ v(\mathbf{p}, u_z) &= w^4(\mathbf{p}), \qquad v(\mathbf{p}, -u_z) = w^3(\mathbf{p}). \end{aligned} \tag{14.47}$$

Spinors $u(\mathbf{p}, s)$ and $v(\mathbf{p}, s)$ are just the ones which appear in the field operators (14.1). (This means that the spin projection onto the direction $\mathbf{s}$ can be taken as the label σ.) To show it more rigorously we take for simplicity $\mathbf{s} = (0, 0, 1)$ and note that in the rest frame of the particle the functions $u(\mathbf{0}, \sigma = +\frac{1}{2})$, $u(\mathbf{0}, \sigma = -\frac{1}{2})$ and $v(\mathbf{0}, \sigma = +\frac{1}{2})$, $v(\mathbf{0}, \sigma = -\frac{1}{2})$ coincided (in the Dirac representation for the γ^μ matrices) with the spinors $w^1(\mathbf{0})$, $w^2(\mathbf{0})$ and $w^4(\mathbf{0})$, $w^3(\mathbf{0})$, respectively.

We gave the Dirac interpretation here mainly for historical reasons and also because it is sometimes convenient for heuristic considerations. One should remember however that it cannot be taken very seriously. Although surprisingly successful for fermions, it cannot be used for integer spin particles because these, being bosons, would make even faster transitions to the occupied negative energy states. Fortunately, the Dirac interpretation is in fact not necessary because as we have seen, in the correct relativistic theory both free particle and free antiparticle states have always positive energies. In the field operators (14.1) the functions $u(\mathbf{p}, s)$ and $\bar{v}(\mathbf{p}, s)$ multiply the operators $b(\mathbf{p}, \sigma)$ and $d(\mathbf{p}, \sigma)$, respectively which annihilate the positive energy states of particle and antiparticle with momentum $\mathbf{p}$ and spin projection $+\hbar/2$ onto the direction $\mathbf{s}$. In this sense $u(\mathbf{p}, s)$ and $\bar{v}(\mathbf{p}, s)$ correspond to wave functions of a particle and antiparticle, respectively.[90]

To give the negative energy solutions of the Dirac equation the proper interpretations we compute the "second quantized" form of the free Dirac Hamiltonian

$$H = \int d^3\mathbf{x}\ \psi^\dagger(\mathbf{x}) \left(-i\hbar c\ \boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta mc^2\right) \psi(\mathbf{x})\ . \tag{14.48}$$

The field operators are taken here in the Schrödinger picture and are constructed following the general rule (5.37) $\psi_\alpha(\mathbf{x}) = \sum_\kappa u^{(\kappa)}_\alpha(\mathbf{x}) a_\kappa$ where the sum over κ runs over all Dirac Hamiltonian eigenstates, including the ones corresponding to negative energies. The operator $\psi_\alpha(\mathbf{x})$ can be then written in the form (14.1) taken at $t = 0$ with the interpretation that $b(\mathbf{p},\sigma)$ are the annihilation operators associated with positive energy Dirac Hamiltonian eigenstates whereas $d^\dagger(\mathbf{p},\sigma)$ are the *annihilation* operators associated with negative energy Dirac Hamiltonian eigenstates we find

$$\begin{aligned} H = \int d^3\mathbf{x}\ \int d\Gamma_\mathbf{q} \sum_{\sigma'} &\left[v^\dagger(\mathbf{q},\sigma')e^{i\mathbf{q}\cdot\mathbf{x}} d(\mathbf{q},\sigma') + u^\dagger(\mathbf{q},\sigma')e^{-i\mathbf{q}\cdot\mathbf{x}} b^\dagger(\mathbf{q},\sigma')\right] \\ \times \int d\Gamma_\mathbf{p} \sum_\sigma E_\mathbf{p} &\left[u(\mathbf{p},\sigma)e^{i\mathbf{p}\cdot\mathbf{x}} b(\mathbf{p},\sigma) - v(\mathbf{p},\sigma)e^{-i\mathbf{p}\cdot\mathbf{x}} d^\dagger(\mathbf{p},\sigma)\right]\ , \end{aligned}$$

where we have used the fact the $u(\mathbf{p},\sigma)e^{i\mathbf{p}\cdot\mathbf{x}}$ and $v(\mathbf{p},\sigma)e^{-i\mathbf{p}\cdot\mathbf{x}}$ are eigenvectors of $H_0^{\rm Dir}$ with eigenvalues $+E_\mathbf{p}$ and $-E_\mathbf{p}$, respectively. Recalling the identifications (14.47) and the second rule (14.44) we see that

$$\begin{aligned} u^\dagger(\mathbf{p},\sigma')u(\mathbf{p},\sigma) = 2E_\mathbf{p}\delta_{\sigma'\sigma}, &\qquad u^\dagger(-\mathbf{p},\sigma')v(\mathbf{p},\sigma) = 0\ , \\ v^\dagger(\mathbf{p},\sigma')v(\mathbf{p},\sigma) = 2E_\mathbf{p}\delta_{\sigma'\sigma}, &\qquad v^\dagger(-\mathbf{p},\sigma')u(\mathbf{p},\sigma) = 0\ , \end{aligned}$$

[90] The wave function is the probability amplitude of finding the particle at $\mathbf{x}$. The states representing the particle and antiparticle at $\mathbf{x}$ are

$$\psi^\dagger_\alpha(\mathbf{x})|\Omega\rangle \qquad \psi_\alpha(\mathbf{x})|\Omega\rangle\ .$$

Hence, the particle and antiparticle wave functions are

$$\begin{aligned} \langle\Omega|\psi_\alpha(\mathbf{x})b^\dagger(\mathbf{p},s)|\Omega\rangle &= u_\alpha(\mathbf{p},s)e^{i\mathbf{p}\cdot\mathbf{x}}\ , \\ \langle\Omega|\psi^\dagger_\alpha(\mathbf{x})d^\dagger(\mathbf{p},s)|\Omega\rangle &= v^\dagger_\alpha(\mathbf{p},s)e^{i\mathbf{p}\cdot\mathbf{x}}\ . \end{aligned}$$

but interpretation is obscured by the index α which cannot be interpreted here as directly referring to the basis states of the spin projection.

which leads to

$$H = \int d\Gamma_{\mathbf{p}} \sum_{\sigma} E_{\mathbf{p}} \left[b^{\dagger}(\mathbf{p},\sigma) b(\mathbf{p},\sigma) - d(\mathbf{p},\sigma) d^{\dagger}(\mathbf{p},\sigma) \right] \ .$$

To get rid of negative energy states the *annihilation* operators $d^{\dagger}$ are now reinterpreted as the *creation* operators (this is possible because the *anticommutation* rule satisfied by d and $d^{\dagger}$ is symmetric with respect to such a redefinition). Writing then the Hamiltonian in the form in which the creation operators stand to the left of the annihilation operators we recover, up to an infinite constant, the Hamiltonian of the free spin 1/2 particles and antiparticles

$$H_0 = \int d\Gamma_{\mathbf{p}} \sum_{\sigma} E_{\mathbf{p}} \left[b^{\dagger}(\mathbf{p},\sigma) b(\mathbf{p},\sigma) + d^{\dagger}(\mathbf{p},\sigma) d(\mathbf{p},\sigma) \right] \ , \qquad (14.49)$$

which was the starting point of our approach. (The dropped constant is just the energy of the "filled Dirac sea".) From this short calculation it becomes clear that in fact getting rid of the whole concept of the Dirac sea amounts simply to subtract an infinite constant from the Hamiltonian of spin 1/2 particles and adopting the second quantized version of the theory with both particles and antiparticles present.

14.2 Dirac equation with external potential

Now that we have understood the role of the free Dirac equation we want to explain what role in the theory is played by the Dirac equation with external electromagnetic potential. Its correct form can be easily written down by respecting Lorentz covariance of the equation and by invoking the so-called principle of minimal coupling. The latter is known also from classical mechanics and consists of introducing the vector potential with the help of the substitution $\hat{\mathbf{p}} \to \hat{\mathbf{p}} - \frac{qe}{c}\mathcal{A}^{\rm ext}$. In this way one obtains the Dirac equation in the form

$$i\hbar \frac{\partial}{\partial t}\psi(t,\mathbf{x}) = \left[-i\hbar c \alpha^i \left(\frac{\partial}{\partial x^i} + i\frac{qe}{\hbar c}\mathcal{A}_i^{\rm ext}(t,\mathbf{x}) \right) + \beta mc^2 + qe\phi(t,\mathbf{x}) \right] \psi(t,\mathbf{x}) \ ,$$

(remember: $\mathcal{A}_i = -\mathcal{A}^i$) or, in the covariant form (setting $\hbar = c = 1$),

$$\left(i \not{\partial} - qe \not{\mathcal{A}} - m \right) \psi(x) = 0 \qquad (14.50)$$

For future use, note also that even if in a special frame the external potential has the simple static form $\mathcal{A}^\mu(t, \mathbf{x}) = (\phi(\mathbf{x}), \mathbf{0})$, it will take the more general form with time dependent potential ϕ and nonzero $\mathcal{A}(t, \mathbf{x})$ in any other frame moving with respect to the original one.

Consider now a process in which many particles interact electromagnetically. We are now able to write down the complete amplitude for any such process using the covariant Feynman rules. Suppose however, that in the the process considered one particle is much heavier than the other ones and has a mass much bigger than energies (divided by c^2) of all other participating particles. The line representing this heavy particle goes through any Feynman diagram contributing to the amplitude of the process as shown in figure 20. Let us denote the electric charge (in units e) of the heavy particle by Q. The analytic expression for the contribution of the diagram shown in in figure 20 is easily written down:

$$-i\mathcal{M} = \int \frac{d^4p_1}{(2\pi)^4} \dots \int \frac{d^4p_{N-1}}{(2\pi)^4} \int \frac{d^4q_1}{(2\pi)^4} \dots \int \frac{d^4q_N}{(2\pi)^4} \left(\frac{-i}{q_1^2}\right) \dots \left(\frac{-i}{q_N^2}\right)$$
$$\times(-ieQ)^N (2\pi)^4 \delta^{(4)}(p_1 + q_1 - p_0) \dots (2\pi)^4 \delta^{(4)}(p_N + q_N - p_{N-1})$$
$$\times \bar{u}(\mathbf{p}', \sigma')\gamma^{\mu_N} \frac{i}{\not{p}_{N-1} - M} \gamma^{\mu_{N-1}} \dots \gamma^{\mu_2} \frac{i}{\not{p}_1 - M} \gamma^{\mu_1} u(\mathbf{p}, \sigma) \mathcal{M}^{\text{rest}}_{\mu_1, \dots \mu_N},$$

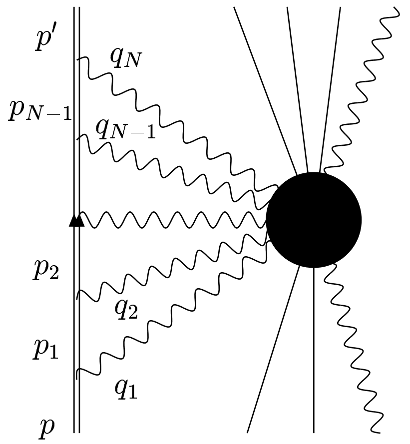
where $\mathcal{M}^{\text{rest}}_{\mu_1, \dots \mu_N}$ denotes the analytic expression corresponding to the black blob.

The integrations over the momenta of virtual states of the heavy particle can be done with the help of the delta functions. As a result one gets

$$-i\mathcal{M} = (-Qe)^N (i)^{N-1} \int \prod_{i=1}^{N} \left(\frac{d^4q_i}{(2\pi)^4} \frac{1}{q_i^2} \right) \ (2\pi)^4 \delta^{(4)}(p' - p + \sum_{i=1}^{N} q_i)$$
$$\times \bar{u}(\mathbf{p}', \sigma')\gamma^{\mu_N} \frac{1}{\not{p}_{N-1} - M} \gamma^{\mu_{N-1}} \dots \gamma^{\mu_2} \frac{1}{\not{p}_1 - M} \gamma^{\mu_1} u(\mathbf{p}, \sigma) \mathcal{M}^{\text{rest}}_{\mu_1, \dots \mu_N}, \quad (14.51)$$

where now the momenta $p_1 \dots p_{N-1}$ are expressed in terms of p and $q_1 \dots q_N$ as follows from four-momentum conservation in each vertex.

We will now consider the limit in which the four-momentum of the heavy particle changes very little compared to the scale set by its mass M. It is then possible to approximate the momenta $p_1 \dots p_{N-1}$ in the



+ diagrams with all possible permutations of the points in which photons are attached to the heavy particle line

Figure 20: Feynman diagrams describing a process with one very heavy nonrelativistic particle. The four-momenta q_i are all directed towards the black blob.

numerators of the heavy particle propagators simply by $p_i \approx p \approx p'$. The string of the gamma matrices in the numerator can be then rewritten as:

$$\bar{u}(\mathbf{p}', \sigma')\gamma^{\mu_N}(\not{p}_{N-1} + M)\gamma^{\mu_{N-1}}(\not{p}_{N-2} + M)\gamma^{\mu_{N-2}} \ldots \gamma^{\mu_2}(\not{p}_1 + M)\gamma^{\mu_1}u(\mathbf{p}, \sigma)$$

$$\approx \bar{u}(\mathbf{p}, \sigma')\gamma^{\mu_N}(\not{p} + M)\gamma^{\mu_{N-1}}(\not{p} + M)\gamma^{\mu_{N-2}} \ldots \gamma^{\mu_2}(\not{p} + M)\gamma^{\mu_1}u(\mathbf{p}, \sigma)$$

$$= \sum_{\sigma_1, \ldots \sigma_{N-1}} [\bar{u}(\mathbf{p}, \sigma')\gamma^{\mu_N}u(\mathbf{p}, \sigma_{N-1})][\bar{u}(\mathbf{p}, \sigma_{N-1})\gamma^{\mu_{N-1}}u(\mathbf{p}, \sigma_{N-2})][\bar{u}(\mathbf{p}, \sigma_{N-2})$$
$$\gamma^{\mu_{N-2}}u(\mathbf{p}, \sigma_{N-3})] \ldots [\bar{u}(\mathbf{p}, \sigma_2)\gamma^{\mu_2}u(\mathbf{p}, \sigma_1)][\bar{u}(\mathbf{p}, \sigma_1)\gamma^{\mu_1}u(\mathbf{p}, \sigma)]$$

where we have repeatedly used the relation

$$\not{p} + M = \sum_\sigma u(\mathbf{p}, \sigma)\bar{u}(\mathbf{p}, \sigma)$$

This leads to a great simplification because as it is easy to check

$$\bar{u}(\mathbf{p}, \sigma')\gamma^\mu u(\mathbf{p}, \sigma) = 2p^\mu \delta_{\sigma'\sigma} \tag{14.52}$$

(this can be checked e.g. by setting $\sigma' = \sigma$ and summing over σ which gives $4p^\mu$ on the RHS and amounts on the LHS to $tr[(\not{p} + M)\gamma^\mu]$ which is the same). Therefore in this limit the whole numerator in the last line of (14.51) becomes

$$(2p^{\mu_N})(2p^{\mu_{N-1}})\dots(2p^{\mu_2})(2p^{\mu_1})\ \delta_{\sigma'\sigma}\ .$$

In turn, using the mass shell conditions $p^2 = p'^2 = M^2$ the denominator of the last line of (14.51) can be rewritten as

$$\frac{(-1)}{[2p\cdot(q_1+\dots+q_{N-1})-i0]}\frac{(-1)}{[2p\cdot(q_1+\dots+q_{N-2})-i0]}\cdots\frac{(-1)}{[2p\cdot q_1-i0]}\ .$$

As a result, for the approximation to the amplitude corresponding to the diagram shown in figure 20 we get

$$-i\mathcal{M} = (-Qe)^N(-i)^{N-1}\int\prod_{i=1}^{N}\left(\frac{d^4q_i}{(2\pi)^4}\frac{1}{q_i^2}\right)\times(2\pi)^4\delta(p'-p+\sum_{i=1}^{N}q_i)$$
$$\frac{(2p^{\mu_N})(2p^{\mu_{N-1}})\dots(2p^{\mu_2})(2p^{\mu_1})\ \delta_{\sigma'\sigma}\ \mathcal{M}^{\text{rest}}_{\mu_1,\dots\mu_N}}{[2p\cdot(q_1+\dots+q_{N-1})-i0][2p\cdot(q_1+\dots+q_{N-2})-i0]\dots[2p\cdot q_1-i0]}.$$

In the considered limit the energy delta function $\delta(E_{\mathbf{p}'}-E_{\mathbf{p}}+\sum_{i=1}^{N}\omega_i)$ can be approximated as follows[91]

$$\delta\left(\sqrt{M^2+(\mathbf{p}-\mathbf{q}_1-\dots-\mathbf{q}_N)^2}+\sum_{i=1}^{N}\omega_i-\sqrt{M^2+\mathbf{p}^2}\right)$$
$$\approx\delta\left(-\frac{\mathbf{p}}{E_{\mathbf{p}}}\cdot(\mathbf{q}_1+\dots+\mathbf{q}_N)+\sum_{i=1}^{N}\omega_i\right) = E_{\mathbf{p}}\delta\left(p\cdot q\right)\ ,$$

where $q\equiv\sum_{i=1}^{N}q_i$ and ω_i are the virtual photon frequencies.

As the last step we have to add expressions corresponding to diagrams similar to the one shown in figure 20 but differing from it by permutations of the points at which the photons carrying momenta $q_1\dots q_N$ are

[91] At this point the our considerations become slightly formal, because the four-momenta q_i are integrated from $-\infty$ to $+\infty$. If the diagram contains no closed loops, then of course this is not a problem, as the momentum delta functions enforce that q_i are all expressed in terms of the small (by our assumption) four-momenta of the initial and/or final light particles. However if some of q_i are the four-momenta circulating in loops the argument is more tricky and relies on the renormalization theory: each diagram must be then considered together with a set of the so-called counterterm diagrams. The integrals over the loop momenta q_i are are then effectively cut-off at the scale of the momenta of the light particles and our treating q_i as small is justified.

attached to the heavy particle line. For the sum of all such expressions we get then

$$-i\mathcal{M} = (-Qe)^N(-i)^{N-1}\int\prod_{i=1}^{N}\left(\frac{d^4q_i}{(2\pi)^4}\frac{1}{q_i^2}\right)(2\pi)^4\delta^3(\mathbf{p}'-\mathbf{p}-\mathbf{q})2E_{\mathbf{p}}\delta(p\cdot q)$$

$$\delta_{\sigma'\sigma}\left(\sum_{\text{perm}}\frac{1}{[p\cdot q_1-i0]}\cdots\frac{1}{[p\cdot(q_1+\ldots+q_{N-1})-i0]}\right)p^{\mu_N}\ldots p^{\mu_1}\mathcal{M}^{\text{rest}}_{\mu_1,\ldots\mu_N}.$$

The sum in the bracket in the second line, when considered together with the delta $\delta(p\cdot q)$, turns out to be rather simple:

$$\delta(p\cdot q)\left(\sum_{\text{perm}}\cdots\right) = (2\pi i)^{N-1}\delta(p\cdot q_1)\delta(p\cdot q_2)\ldots\delta(p\cdot q_N).$$

For example for $N=2$, using the standard relation

$$\frac{1}{x\pm i0} = P\frac{1}{x}\mp i\pi\delta(x) \tag{14.53}$$

we have

$$\begin{aligned}
&\delta[p\cdot(q_1+q_2)]\left(\frac{1}{p\cdot q_1-i0}+\frac{1}{p\cdot q_2-i0}\right)\\
&=\delta[p\cdot(q_1+q_2)]\left(\frac{1}{p\cdot q_1-i0}-\frac{1}{p\cdot q_1+i0}\right)\\
&=2\pi i\delta[p\cdot(q_1+q_2)]\delta(p\cdot q_1) = 2\pi i\delta(p\cdot q_1)\delta(p\cdot q_2).
\end{aligned}$$

and this can be continued to higher N. The final expression for the the sum of diagrams of figure 20 then takes the form

$$\begin{aligned}
-i\mathcal{M} = (-Qe)^N\int\prod_{i=1}^{N}\left(\frac{d^4q_i}{(2\pi)^4}\frac{1}{q_i^2}\right)(2\pi)^3 2E_{\mathbf{p}}\delta^3(\mathbf{p}'-\mathbf{p}+\mathbf{q})\delta_{\sigma'\sigma}\\
\times(2\pi)^N\delta(p\cdot q_1)\delta(p\cdot q_2)\ldots\delta(p\cdot q_N)p^{\mu_N}\ldots p^{\mu_1}\mathcal{M}^{\text{rest}}_{\mu_1,\ldots,\mu_N}\ .
\end{aligned}$$

We now assume that the heavy particle is nonrelativistic, $p^0 = E_{\mathbf{p}} \gg |\mathbf{p}|$, so that $p^{\mu_i}\approx p^0 n^{\mu_i}$, where $n^\mu = (1,0,0,0)$ is the unit time-like four-vector. Therefore,

$$\delta(p\cdot q_i)\approx\frac{1}{p^0}\delta(q_i^0) \tag{14.54}$$

and the integrals over dq_i^0 can be taken explicitly. The amplitude reduces then to

$$-i\mathcal{M} = \int \prod_{i=1}^{N} \left(\frac{d^3\mathbf{q}_i}{(2\pi)^3} \frac{Qe}{\mathbf{q}_i^2} \right) (2\pi)^3 2E_{\mathbf{p}} \delta^3(\mathbf{p}' - \mathbf{p} + \mathbf{q}) \delta_{\sigma'\sigma} \mathcal{M}^{\text{rest}}_{0,\dots,0} \ .$$

Instead of taking for the initial and final states of the heavy particle definite momentum states, we can consider them to be the same state characterized by some momentum wave function $\chi_\sigma(\mathbf{p})$. The amplitude of such a process is obtained by integrating $-i\mathcal{M}$ with $\chi_\sigma(\mathbf{p})$:

$$-i\mathcal{M}(\chi \leftarrow \chi) = -i \int d\Gamma_{\mathbf{p}'} \int d\Gamma_{\mathbf{p}} \sum_{\sigma'\sigma} \chi^*_{\sigma'}(\mathbf{p}') \chi_\sigma(\mathbf{p}) \mathcal{M}(\mathbf{p}'\sigma' \leftarrow \mathbf{p}\sigma)$$

where $\chi_\sigma(\mathbf{p})$ is normalized in such a way that

$$\int d\Gamma_{\mathbf{p}} \sum_{\sigma} |\chi_\sigma(\mathbf{p})|^2 = 1 \tag{14.55}$$

The factor $2E_{\mathbf{p}}$ can be written as $\approx \sqrt{2E_{\mathbf{p}}}\sqrt{2E_{\mathbf{p}'}}$ and we get

$$-i\mathcal{M}(\chi \leftarrow \chi) = \int d^3\mathbf{r} \sum_{\sigma} |\tilde{\chi}_\sigma(\mathbf{r})|^2 \int \prod_{i=1}^{N} \left(\frac{d^3\mathbf{q}_i}{(2\pi)^3} \frac{Qe}{\mathbf{q}_i^2} e^{-i\mathbf{q}_i\cdot\mathbf{r}} \right) \mathcal{M}^{\text{rest}}_{0,\dots,0}, \tag{14.56}$$

where the position space wave function

$$\tilde{\chi}_\sigma(\mathbf{r}) \equiv \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{\chi_\sigma(\mathbf{p})}{\sqrt{2E_{\mathbf{p}}}}\, e^{i\mathbf{p}\cdot\mathbf{r}} \approx \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{\chi_\sigma(\mathbf{p})}{\sqrt{2M}}\, e^{i\mathbf{p}\cdot\mathbf{r}} \tag{14.57}$$

is normalized to unity[92]

$$\int d^3\mathbf{r} \sum_{\sigma} |\tilde{\chi}_\sigma(\mathbf{r})|^2 = 1 \ . \tag{14.59}$$

---

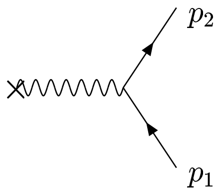
[92] Because

$$\int \frac{d^3\mathbf{p}}{(2\pi)^3} \sum_{\sigma} \left| \frac{\chi_\sigma(\mathbf{p})}{\sqrt{2E_{\mathbf{p}}}} \right|^2 = 1 \tag{14.58}$$

this follows from the so-called Parseval's identity.

In this way the amplitude (14.56) is averaged over the position of the heavy particle. If the wave function $\tilde{\chi}_\sigma(\mathbf{r})$ is peaked around some point $\mathbf{X}$ we get

$$-i\mathcal{M}(\chi \leftarrow \chi) = \int \prod_{i=1}^{N} \left(\frac{d^3\mathbf{q}_i}{(2\pi)^3} \frac{Qe}{\mathbf{q}_i^2} e^{-i\mathbf{q}_i\cdot\mathbf{X}} \right) \mathcal{M}^{\text{rest}}_{0,\ldots 0} \ . \tag{14.60}$$



$$-ieq \textstyle\int \frac{d^4q}{(2\pi)^4}\, \mathcal{A}^{\text{ext}}(q)(2\pi)^4\delta^4(p_1 - q - p_2)$$

Figure 21: Feynman rule for charged particle interactions with an external four-potential $\mathcal{A}_\mu(x)$.

Formally, the same amplitude is obtained by adding to the interaction Hamiltonian a term describing charged particle interaction with an external Coulomb potential and removing from it all terms containing the field operators of the heavy particle. The effective interaction Hamiltonian therefore reads

$$\mathcal{H}_I^{\text{eff}} = eq\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu(x) + eq\bar{\psi}(x)\gamma^\mu\psi(x)\mathcal{A}^{\text{ext}}_\mu(x) \ , \tag{14.61}$$

where $A_\mu(x)$ is the photon quantum field operator and

$$\mathcal{A}^{\text{ext}}_\mu(t,\mathbf{r}) = \begin{cases} \mathcal{A}^{\text{ext}}_i = 0 \\ \mathcal{A}^{\text{ext}}_0(t,\mathbf{r}) = \frac{Qe}{4\pi|\mathbf{r}-\mathbf{X}|} \end{cases} \tag{14.62}$$

or, in other words,

$$\mathcal{A}^{\text{ext}}_0(t,\mathbf{r}) = Qe \int \frac{d^3\mathbf{q}}{(2\pi)^3} \frac{e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{X})}}{|\mathbf{q}|^2} = Qe \int \frac{d^4q}{(2\pi)^4} 2\pi\delta(q^0) \frac{e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{X})}}{|\mathbf{q}|^2} \tag{14.63}$$

is the c-number external field. The interaction with the external field $\mathcal{A}^{\text{ext}}_\mu(x)$ in the effective Hamiltonian (14.61) leads to an additional Feynman rule shown in figure 21 in which $\mathcal{A}^{\text{ext}}_\mu(q)$ is defined by

$$\mathcal{A}^{\text{ext}}_\mu(x) = \int \frac{d^4q}{(2\pi)^4} \mathcal{A}^{\text{ext}}_\mu(q) e^{iq\cdot x} \ . \tag{14.64}$$

The amplitude (14.60) is then obtained by computing the Feynman diagram shown in figure 22 with

$$\mathcal{A}^{\text{ext}}_{\mu}(q) = Qe(2\pi)\delta(q^0)\frac{e^{-i\mathbf{q}\cdot\mathbf{X}}}{|\mathbf{q}|^2}\delta_{\mu 0} \tag{14.65}$$

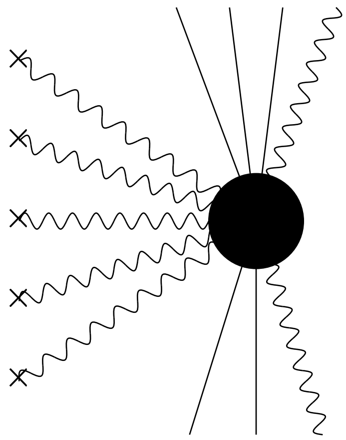


Figure 22: Feynman diagrams describing N interactions with an external field.

Indeed, from the Feynman rule shown in fig. 21 we get for the diagram of fig. 22 the following contribution to the amplitude the expression

$$\prod_{i=1}^{N}\left(\int \frac{d^4q_i}{(2\pi)^4}\mathcal{A}^{\text{ext}}_{\mu_i}(q_i)\right)\mathcal{M}^{\mu_1\ldots\mu_N}_{\text{rest}} \tag{14.66}$$

(the factors $-ieq(2\pi)^4\delta^{(4)}(k-q_i-pk')\gamma_{\mu_i}$ form the part of the amplitude $\mathcal{M}^{\mu_1,\ldots,\mu_N}_{\text{rest}}$) which for $\mathcal{A}^{\text{ext}}_{\mu_i}(q_i)$ of the form (14.65) is equal to (14.60).

Thus, contributions to the amplitude of the scattering of a light particle (say, electron) off a heavy one (say, proton) of diagrams shown in the upper part of figure 23 can be, in the limit, in which the heavy particle (denoted in figure 23 by the double lines) is nonrelativistic, approximated by diagrams shown in the lower part of figure 23, which describe light particle scattering in the external Coulomb potential. (If the light particle is nonrelativistic too, the dominant contribution can be shown to arise from the so-called ladder diagrams shown in figure 23, that is the ones in which photon lines do not cross each other. The sum of such diagrams is resummed by using the Bethe-Salpeter equation.)

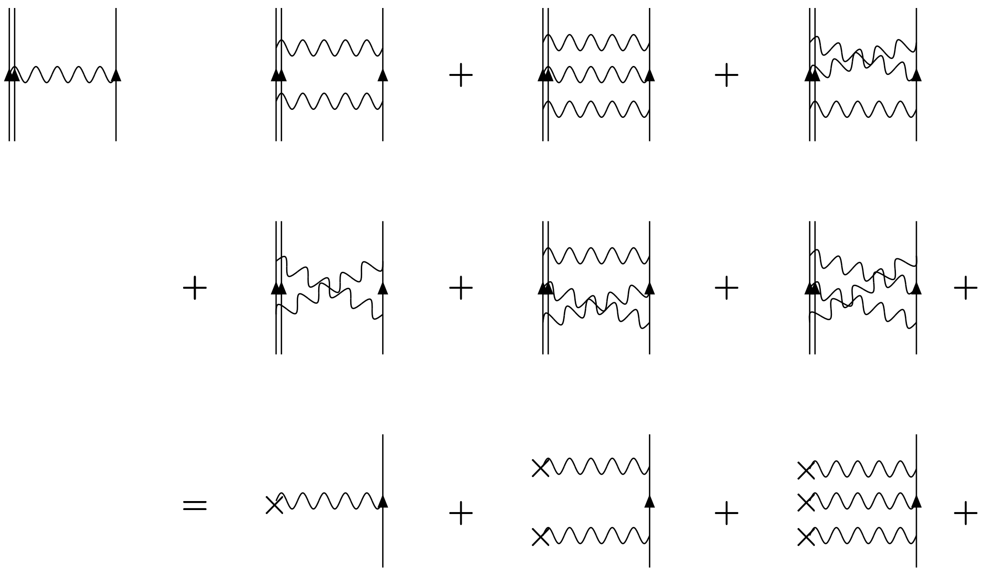


Figure 23: Two upper sets of Feynman diagrams contribe to the heavy-light particle scattering in full electrodynamics (there are also other diagrams). For nonrelativistic heavy particle (double line) these contributions are taken into account by using Coulomb external field (lower set of diagrams).

The total effective Hamiltonian (in the Schrödinger picture) describing the scattering in this approximation is

$$\begin{aligned} H = &\int d\Gamma_{\mathbf{p}} \sum_{\sigma} E_{\mathbf{p}} \left[b^{\dagger}(\mathbf{p},\sigma) b(\mathbf{p},\sigma) + d^{\dagger}(\mathbf{p},\sigma) d(\mathbf{p},\sigma) \right] \\ &+ eq \int d^3\mathbf{x}\ \bar{\psi}(0,\mathbf{x}) \gamma^{\mu} \psi(0,\mathbf{x}) \mathcal{A}^{\rm ext}_{\mu}(x)\ , \\ &+ eq \int d^3\mathbf{x}\ \bar{\psi}(0,\mathbf{x}) \gamma^{\mu} \psi(0,\mathbf{x}) A_{\mu}(0,\mathbf{x}) \\ &+ \int d\Gamma_{\mathbf{k}} \sum_{\lambda} \omega_{\mathbf{k}} a^{\dagger}(\mathbf{k},\lambda) a(\mathbf{k},\lambda)\ . \end{aligned} \tag{14.67}$$

As we have already shown the first line, up to an infinite constant, takes the form given in eq. (14.48). Recall now, that in the case of nonrelativistic quantum mechanics of many particles which apart from interacting with each other through two particle forces $V(\mathbf{x}_i - \mathbf{x}_j)$, interacted in addition with an external force represented in the second-quantized Hamiltonian by

$$V = \int d^3\mathbf{x}\ a^\dagger(\mathbf{x})V(\mathbf{x})a(\mathbf{x})\ ,$$

it was possible to get rid of this term by solving the 1-particle Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V(\mathbf{x}) - E_n\right] u_n(\mathbf{x}) = 0\ , \tag{14.68}$$

where the index n can take discrete as well as continuous values, and expanding the field operators $a(\mathbf{x})$ and $a^\dagger(\mathbf{x})$ not in operators creating eigenstates of the momentum operators but in the creation and annihilation operators creating particles in states $|n\rangle$:

$$a(\mathbf{x}) = \sum_n u_n(\mathbf{x})a_n, \qquad a^\dagger(\mathbf{x}) = \sum_n u_n^*(\mathbf{x})a_n^\dagger\ . \tag{14.69}$$

As a result the total Hamiltonian of the system took the form

$$H = \sum_n E_n a_n^\dagger a_n + \int d^3\mathbf{x} \int d^3\mathbf{x}' a^\dagger(\mathbf{x})a^\dagger(\mathbf{x}')V(\mathbf{x}-\mathbf{x}')a(\mathbf{x}')a(\mathbf{x}), \tag{14.70}$$

where in the last term one has to use (14.69). The advantage is then that the interaction with the external field is taken into account exactly.

Similarly, by solving the Dirac equation in an external Coulomb field

$$\left[-i\hbar c\ \alpha^i\frac{\partial}{\partial x^i} + \beta mc^2 + qe\mathcal{A}_0(\mathbf{x})\right]\psi(t,\mathbf{x}) = E\ \psi(t,\mathbf{x})\ , \tag{14.71}$$

one can take into account interactions of light particles with a nonrelativistic heavy particle exactly. More specifically, solving the equation (14.71) we find positive $u_N(\mathbf{x})$ and negative $v_M(\mathbf{x})$ energy solutions. The indices N and M can now have different ranges; for example if the light particles (annihilated by the operator ψ) are electrons and the Coulomb field is produced by a proton, the index N (corresponding to positive energy solutions) runs over discrete (bound states) and continuous values whereas the index M (corresponding to negative energy solutions) runs only over continuous values (positrons are not bound by protons). The field operators (taken in the interaction picture) can be then expanded as

$$\begin{aligned}\psi(t,\mathbf{x}) &= \sum_N u_N(\mathbf{x})e^{-iE_Nt}b_N + \sum_M v_M(\mathbf{x})e^{+i|E_M|t}d^\dagger_M \\ \bar{\psi}(t,\mathbf{x}) &= \sum_M \bar{v}_M(\mathbf{x})e^{-i|E_M|t}d_M + \sum_N \bar{u}_N(\mathbf{x})e^{+iE_Nt}b^\dagger_N\end{aligned} \tag{14.72}$$

and the first two lines of the Hamiltonian (14.67) take the form

$$\begin{aligned}H_{\text{“free”}} &= \sum_N E_N b^\dagger_N b_N + \sum_M E_M d_M d^\dagger_M \\ &= \sum_N E_N b^\dagger_N b_N + \sum_M |E_M| d^\dagger_M d_M + \infty\ .\end{aligned} \tag{14.73}$$

The infinite constant can be subtracted and we see that the operators $b^\dagger_N$ and $d^\dagger_M$ all create states having always positive energies. These energies take into account all possible interactions of positively and negatively charged particles with the heavy particle but do not take yet into account the virtual corrections. These can be now computed perturbatively. This is the way to calculate for example the Lamb shift of the $2S^{1/2}$ and $2P^{1/2}$ states in the hydrogen atom.

Amplitudes of processes like the one shown in figure 20 are now computed from the standard formula for the S matrix. The influence of the interaction with the heavy particle is now reflected in the modified form of the electron propagator $iS(x,y)$ which now depends on x and y separately and in the position representation is the solution of the equation

$$\left[\gamma^\mu\left(i\partial_\mu - qe\mathcal{A}^{\rm ext}_\mu(0,\mathbf{x})\right) - m\right]S(x,y) = \delta^{(4)}(x-y)\ , \tag{14.74}$$

where $\mathcal{A}^{\rm ext}_\mu$ is given by (14.62) and with appropriate boundary conditions. (This can be deduced in the same way we derived the standard propagator $iS^{\rm F}(x-y)$.)

More generally, one can consider also external fields other that the Coulomb one, because in principle light particles can interact with many nonrelativistic heavy particles which together create a complicated external electromagnetic field (after all any electromagnetic field has to be somehow generated by some charged particles). Moreover, since in a given reference frame some particles can be moving it makes sense to consider also time varying external fields. Amplitudes of processes in which only light particles participate are then computed with Feynman diagrams with propagators given by the equation (14.74) with some $\mathcal{A}^{\rm ext}_\mu(x)$, not necessarily of the Coulomb form.

14.3 Dirac equation with external fields

In the previous section we have explained what role in the complete theory is played by the solutions of the Dirac equation with external fields. In particular we have shown, that the solutions of the Dirac equation in the external Coulomb field should give a good first relativistic approximation to the electron energy in the Hydrogen atom. It is now time to investigate the equation

$$i\hbar\frac{\partial}{\partial t}\psi(t,\mathbf{x}) = \left[-i\hbar c\alpha^i\left(\frac{\partial}{\partial x^i}+i\frac{qe}{\hbar c}A_i(t,\mathbf{x})\right)+\beta mc^2+qe\phi(t,\mathbf{x})\right]\psi(t,\mathbf{x})\ ,$$

more closely (recall: $e>0$ and $q=-1$ for the electron).

First, let us introduce the so-called "large" and "small" spinors φ and χ, respectively by writing

$$\psi(t,\mathbf{x}) = e^{-imc^2t/\hbar}\begin{pmatrix}\varphi(t,\mathbf{x})\\ \chi(t,\mathbf{x})\end{pmatrix}\ . \tag{14.75}$$

Inserting this form of ψ in the Dirac equation and using the Dirac representation of the Dirac matrices (in which γ^0 is diagonal) we get a system of two coupled equations for ϕ and χ:

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\varphi\\ \chi\end{pmatrix} = c\begin{pmatrix}0 & \boldsymbol{\sigma}\cdot\boldsymbol{\pi}\\ \boldsymbol{\sigma}\cdot\boldsymbol{\pi} & 0\end{pmatrix}\begin{pmatrix}\varphi\\ \chi\end{pmatrix}+qe\phi\begin{pmatrix}\varphi\\ \chi\end{pmatrix}-2mc^2\begin{pmatrix}0\\ \chi\end{pmatrix}, \tag{14.76}$$

where

$$\boldsymbol{\pi} = \hat{\mathbf{p}}-\frac{eq}{c}\mathbf{A}\ . \tag{14.77}$$

Consider now stationary positive energy solutions. If we assume that the particle is nonrelativistic, that is its kinetic energy as well as the interaction energies (like $qe\phi$) are small compared with mc^2, the dominant part of the e^{-iEt} factor of ψ is already factorized out by the substitution (14.75). Hence, in the second equation (14.76) we can (in the first approximation) neglect the time derivative of χ as well as the $eq\phi(x)\chi$ term with respect to the $-2mc^2\chi$ term. This allows to express χ in terms of φ

$$\chi = \frac{\boldsymbol{\sigma} \cdot \boldsymbol{\pi}}{2mc}\varphi \ . \tag{14.78}$$

Inserting this in the first equation (14.76) we get the closed equation for the spinor φ:

$$i\hbar\frac{\partial\varphi}{\partial t} = \left[\frac{1}{2m}\left(\boldsymbol{\sigma} \cdot \boldsymbol{\pi}\right)^2 + qe\phi\right]\varphi \ . \tag{14.79}$$

It is then an easy exercise to work out the term $(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2$ and to cast (14.79) in the form of the Pauli's equation

$$i\hbar\frac{\partial\varphi}{\partial t} = \left[\frac{1}{2m}\left(\hat{\mathbf{p}} - \frac{eq}{c}\mathbf{A}\right)^2 - \frac{eq\hbar}{2mc}\boldsymbol{\sigma} \cdot \mathbf{B} + eq\phi\right]\varphi \ . \tag{14.80}$$

It is remarkable, that the nonrelativistic reduction of the Dirac equation automatically yields the correct giromagnetic ratio $g = 2$:

$$\boldsymbol{\mu} = g \ \mu_B \ \mathbf{s} \tag{14.81}$$

where $\boldsymbol{\mu}$ is the magnetic moment, $\mu_B = qe\hbar/2mc$ is the so-called Bohr magneton and $\mathbf{s} = (\hbar/2)\boldsymbol{\sigma}$ is the spin operator. That this is not what was expected on classical ground can be seen by working out the first term in the RHS of (14.80) in the case of constant magnetic field $\mathbf{B}$. The Hamiltonian then is

$$H = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{eq}{2mc}\left(\mathbf{L} + 2\mathbf{s}\right) \cdot \mathbf{B} + \ldots \tag{14.82}$$

This shows that the magnetic moment related to the orbital angular momentum is just $\mu_B\mathbf{L}$ whereas for the spin part $g = 2$ appears as a consequence of special relativity and - as we will explain later - of the minimal coupling of electrons with the electromagnetic field.

A more systematic way of unraveling the physical content of the Dirac equation in the case of nonrelativistic particle is provided by the Foldy-Wouthuysen transformation. Since this is treated in detail in the monograph of Bjorken & Drell, we will summarize below only its main points.

Let us write the Dirac Hamiltonian

$$H^{\text{Dir}} = \left[-i\hbar c \alpha^i \left(\frac{\partial}{\partial x^i} + i \frac{qe}{\hbar c} A_i(t, \mathbf{x}) \right) + \beta mc^2 + qe\phi(t, \mathbf{x}) \right] ,$$

in the form (we set now $\hbar = 1$ and $c = 1$)

$$H^{\text{Dir}} = \beta m + \boldsymbol{\alpha} \cdot (\hat{\mathbf{p}} - qe\mathbf{A}) + qe\phi \equiv \beta m + \mathcal{O} + \mathcal{E} \ , \tag{14.83}$$

in which only the operator $\mathcal{O} = \boldsymbol{\alpha} \cdot (\hat{\mathbf{p}} - qe\mathbf{A})$ mixes "large" and "small" components of the spinor ψ. In the standard terminology $\mathcal{O}$ is called "odd" operator whereas $\mathcal{E}$ and βm are called "even" operators. Note that if the kinetic and the interaction energies of the particle are nonrelativistic, the odd operator $\mathcal{O}$ is subleading (i.e. of order $(m)^0$ with respect to the even one (which is of order $(m)^1$). Moreover, treating formally the interaction energies ($qe\phi$ and $qe\mathbf{A}$) as much smaller than the kinetic energy ($\boldsymbol{\alpha} \cdot \hat{\mathbf{p}}$), i.e. assuming the following hierarchy of the operators

$$\beta m \gg \mathcal{O} \gg \mathcal{E} \gg \frac{\dot{\mathcal{O}}}{m} \sim \frac{\dot{\mathcal{E}}}{m} \ , \tag{14.84}$$

(we assume also that the extermal fields are slowly varying with time) will allow us to organize the expansion in a systematic way. The main idea is to perform a similarity transformation

$$\psi' = e^{iS} \psi \ , \tag{14.85}$$

in such a way that the odd operator becomes more subleading, that is suppressed by more powers of m. After this is achieved (to the desired order), one can neglect it and consider only the equation for the "large" spinor φ. Using (14.85) we write the Dirac equation as

$$i \frac{\partial}{\partial t} e^{-iS} \psi' = H^{\text{Dir}} e^{-iS} \psi' \ . \tag{14.86}$$

Multiplying this from the left by e^{iS}, expanding the exponents (assuming that S is "small" - see below) and rearranging terms we get

$$i \frac{\partial}{\partial t} \psi' = H^{\text{Dir}\prime} \psi' \tag{14.87}$$

with

$$
\begin{aligned}
H^{\mathrm{Dir}\prime} = {} & H^{\mathrm{Dir}} + i\left[S,\ H^{\mathrm{Dir}}\right] + \frac{i^2}{2!}\left[S,\ \left[S,\ H^{\mathrm{Dir}}\right]\right] \\
& + \frac{i^3}{3!}\left[S,\ \left[S,\ \left[S,\ H^{\mathrm{Dir}}\right]\right]\right] + \frac{i^4}{4!}\left[S,\ \left[S,\ \left[S,\ \left[S,\ H^{\mathrm{Dir}}\right]\right]\right]\right] \\
& - \dot{S} - \frac{i}{2}\left[S,\ \dot{S}\right] + \frac{1}{6}\left[S,\ \left[S,\ \dot{S}\right]\right] \ ,
\end{aligned} \tag{14.88}
$$

where the commutators of S with H^{Dir} arise from the expansion of $e^{iS}H^{\mathrm{Dir}}e^{-iS}$ and the terms with $\dot{S}$ arise from the left-hand side of (14.86).

Suppose we chose S in such a way that the odd operator of order $(m)^0$ (that is $\mathcal{O}$) disappears. This is possible with $S \propto 1/m$ because the commutator of βm with S is of order $(m)^0$. It turns out that the right choice is

$$
S = \frac{-i}{2m}\beta\mathcal{O} \ . \tag{14.89}
$$

Indeed, since $\beta\mathcal{O} = -\mathcal{O}\beta$ (the matrices β and α^i anticommute) and $\beta^2 = 1$

$$
\frac{-i}{2m}\left[\beta\mathcal{O},\ \beta m\right] = \frac{-i}{2}\beta\left[\mathcal{O},\ \beta\right] = i\mathcal{O} \ .
$$

Computing the remaining commutators in (14.88) we get

$$
\begin{aligned}
& i\left[S,\ H^{\mathrm{Dir}}\right] = -\mathcal{O} + \frac{\beta}{m}\mathcal{O}^2 + \frac{\beta}{2m}\left[\mathcal{O},\ \mathcal{E}\right] \\
& \frac{i^2}{2!}\left[S,\ \left[S,\ H^{\mathrm{Dir}}\right]\right] = -\frac{\beta}{2m}\mathcal{O}^2 - \frac{1}{2m^2}\mathcal{O}^3 - \frac{1}{8m^2}\left[\mathcal{O},\ \left[\mathcal{O},\ \mathcal{E}\right]\right] \\
& \frac{i^3}{3!}\left[S,\ \left[S,\ \left[S,\ H^{\mathrm{Dir}}\right]\right]\right] = \frac{1}{6m^2}\mathcal{O}^3 - \frac{\beta}{6m^3}\mathcal{O}^4 + \ldots \\
& \frac{i^4}{4!}\left[S,\ \left[S,\ \left[S,\ \left[S,\ H^{\mathrm{Dir}}\right]\right]\right]\right] = \frac{\beta}{24m^3}\mathcal{O}^4 + \ldots \\
& -\dot{S} = \frac{i\beta}{2m}\dot{\mathcal{O}} \\
& -\frac{i}{2}\left[S,\ \dot{S}\right] = -\frac{i}{8m^2}\left[\mathcal{O},\ \dot{\mathcal{O}}\right] \ .
\end{aligned}
$$

In each line we have written the operators following the hierarchy (14.84) and the dots indicate the terms which are smaller than

$$
\mathcal{O}\left(\frac{\mathcal{O}}{m}\right)^3 \sim E_{\mathrm{kin}}\left(\frac{E_{\mathrm{kin}}}{m}\right)^3 , \qquad\qquad \text{or}
$$

$$\mathcal{O}\left(\frac{\mathcal{O}}{m}\right)\left(\frac{\mathcal{E}}{m}\right) \sim E_{\rm kin}\left(\frac{E_{\rm kin}}{m}\right)\left(\frac{E_{\rm int}}{m}\right) . \tag{14.90}$$

The last term of (14.88) does not contribute in this order (by virtue of the assumption that the external fields are slowly varying). The new Hamiltonian then is

$$\begin{aligned} H^{\rm Dir\prime} &= \beta\left(m + \frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^3}\right) + \mathcal{E} - \frac{1}{8m^2}\left[\mathcal{O},\ \left[\mathcal{O},\ \mathcal{E}\right]\right] - \frac{i}{8m^2}\left[\mathcal{O},\ \dot{\mathcal{O}}\right] \\ &+ \frac{\beta}{2m}\left[\mathcal{O},\ \mathcal{E}\right] - \frac{1}{3m^2}\mathcal{O}^3 + \frac{i\beta}{2m}\dot{\mathcal{O}} \\ &\equiv \beta m + \mathcal{E}' + \mathcal{O}' . \end{aligned} \tag{14.91}$$

In $H^{\rm Dir\prime}$ the new odd operator (the second line), denoted by $\mathcal{O}'$, is already of order $1/m$. To make it of order $1/m^2$ we perform another similarity transformation $\psi'' = e^{iS'}\psi'$ so that $H^{\rm Dir\prime\prime}$ is given by the formula (14.88) with $H^{\rm Dir}$ replaced by $H^{\rm Dir\prime}$ and S replaced by

$$S' = \frac{-i}{2m}\beta\mathcal{O}' . \tag{14.92}$$

We get in this way

$$\begin{aligned} H^{\rm Dir\prime\prime} &= \beta m + \mathcal{E}' + \frac{\beta}{2m}\left[\mathcal{O}',\ \mathcal{E}'\right] + \frac{i}{2m}\beta\dot{\mathcal{O}}' \\ &\equiv \beta m + \mathcal{E}'' + \mathcal{O}'' \end{aligned} \tag{14.93}$$

where we have retained only terms not smaller than (14.90). To the order $1/m$ the Hamiltonian $H^{\rm Dir\prime\prime}$ does not contain operators mixing "large" and "small" components of ψ''. If we are interested only in the order $E_{\rm kin}(E_{\rm int}/m)$ corrections to the nonrelativistic treatment of the electron motion we can stop here: when restricted to the "large" component of ψ'' the Hamiltonian $H^{\rm Dir\prime\prime}$

$$H^{\rm Dir\prime\prime}\Big|_{\rm large} = m + \frac{\mathcal{O}^2}{2m} + \mathcal{E} ,$$

coincides (up to the additive constant m) with the Hamiltonian in the Pauli's equation (14.80). To go one step further, that is to take into account the corrections of order (14.90) we have to perform the third similarity transformation with $S'' = -i\beta\mathcal{O}''/2m$ ater which up to the orders (14.90) the new Hamiltonian is free from operators mixing "large" and "small" components. The relevant part of $H^{\mathrm{Dir}'''}$ reads:

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

which is just the first line of (14.91). We have now to work out its terms to the order (14.90):

$$\begin{aligned} &\frac{1}{2m}\mathcal{O}^2 = \frac{1}{2m}\left(\hat{\mathbf{p}} - qe\mathbf{A}\right)^2 - \frac{qe}{2m}\boldsymbol{\Sigma}\cdot\mathbf{B} \ , \\ &-\frac{1}{8m^3}\mathcal{O}^4 = \frac{\hat{\mathbf{p}}^4}{8m^3} + \ldots, \\ &\frac{1}{8m^2}\left([\mathcal{O},\ \mathcal{E}] + i\dot{\mathcal{O}}\right) = i\frac{qe}{8m^2}\boldsymbol{\alpha}\cdot\mathbf{E} \ , \\ &-i\frac{qe}{8m^2}\left[\mathcal{O},\ \boldsymbol{\alpha}\cdot\mathbf{E}\right] = -\frac{qe}{8m^2}\mathrm{div}\mathbf{E} - i\frac{qe}{8m^2}\boldsymbol{\Sigma}\cdot\mathrm{rot}\mathbf{E} - \frac{qe}{4m^2}\boldsymbol{\Sigma}\cdot(\mathbf{E}\times\hat{\mathbf{p}}) \ , \end{aligned}$$

so that the part of the final Hamiltonian acting on the "big" spinor φ (restoring all $\hbar$ and c) reads

$$\begin{aligned} H^{\mathrm{Dir}'''} &= mc^2 + \frac{1}{2m}\left(\hat{\mathbf{p}} - \frac{qe}{c}\mathbf{A}\right)^2 + qe\phi - \frac{qe}{mc}\frac{\hbar}{2}\boldsymbol{\sigma}\cdot\mathbf{B} - \frac{\hat{\mathbf{p}}^4}{8m^3c^2} \\ &\quad - \frac{qe\hbar^2}{8m^2c^2}\mathrm{div}\mathbf{E} - i\frac{qe\hbar}{8m^2c}\boldsymbol{\sigma}\cdot\mathrm{rot}\mathbf{E} - \frac{qe\hbar}{4m^2c}\boldsymbol{\sigma}\cdot(\mathbf{E}\times\hat{\mathbf{p}}) \end{aligned}$$

Compared to the Pauli's Hamiltonian in (14.80) we get here three different kinds of corrections. The first one (given by the last term in the first line) is the obvious relativistic correction to the energy-momentum relation. The second one (the first term in the second line), called the Darwin term, is related to the so-called *Zitterbewegung*. The nonrelativistic motion of a particle is an average over the very rapid oscillations in the region characterized by the length $|\delta\mathbf{r}| \sim \hbar/mc$ with frequency greater than $2mc^2/\hbar$. This becomes obvious when one considers e.g. the time evolution of a Gaussian wave packet (see Bjorken & Drell). Because of these oscillations the particle feels the potential ϕ not exactly at $\mathbf{r}$ but rather averaged over the region $\sim |\delta\mathbf{r}|$. This brings in a correction to the potential energy $V = qe\phi$ of order

$$\begin{aligned} \delta V &= \langle qe\phi(\mathbf{r} + \delta\mathbf{r})\rangle - \langle qe\phi(\mathbf{r})\rangle \\ &\approx qe\langle\delta r^i\partial_i\phi(\mathbf{r}) + \frac{1}{2}\delta r^i\delta r^j\partial_i\partial_j\phi(\mathbf{r})\rangle = \frac{1}{6}\ qe\ |\delta\mathbf{r}|^2\Delta\phi(\mathbf{r}) \ . \end{aligned}$$

(the last equality holds after averaging over directions of the vector $\delta\mathbf{r}$). For $\mathbf{E} = -\boldsymbol{\nabla}\phi$ (in the radiation gauge $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$) and setting $|\delta\mathbf{r}| = \hbar/mc$ this can be rewritten as

$$\delta V = -\frac{qe\hbar^2}{6m^2c^2}\mathrm{div}\mathbf{E} \ ,$$

which is of the same order of magnitude and sign as the Darwin term. Finally the two last terms (which only taken together give the Hermitian operator) can be most easily investigated in the case of a spherically symmetric static electric potential ϕ. For $\mathbf{E} = -\boldsymbol{\nabla}\phi$ the first of these terms vanishes and the second one gives the *spin-orbit* interaction

$$H_{\mathrm{spin-orbit}} = \frac{qe\hbar}{4m^2c}\frac{1}{r}\frac{\partial\phi}{\partial r}\boldsymbol{\sigma}\cdot\hat{\mathbf{L}} = \frac{qe}{2m^2c}\frac{1}{r}\frac{\partial\phi}{\partial r}\mathbf{s}\cdot\hat{\mathbf{L}} \ . \tag{14.94}$$

This interaction can be understood as follows. Consider the Hydrogen atom. In the instantaneous reference system, in which the electron is at rest, the nucleus is moving around it generating in addition to the electric field $\mathbf{E}' \approx \mathbf{E} = -\boldsymbol{\nabla}\phi$ also the magnetic field, which (using the Lorentz transformation of the tensor $F^{\mu\nu}$) is given by

$$\mathbf{B}' \approx -\mathbf{v}\times\mathbf{E} \ ,$$

where $\mathbf{v} = \mathbf{p}/m$ is the instantaneous velocity of the electron in the nucleus rest reference system. The interaction of the electron magnetic moment with the magnetic field $\mathbf{B}'$ gives then the correction

$$-\frac{qe}{mc}\frac{\hbar}{2}\boldsymbol{\sigma}\cdot\mathbf{B} = \frac{qe}{m^2c}\mathbf{s}\cdot(\hat{\mathbf{p}}\times\mathbf{E}) = \frac{qe}{m^2c}\frac{1}{r}\frac{\partial\phi}{\partial r}(\mathbf{r}\times\hat{\mathbf{p}})\cdot\mathbf{s} = \frac{qe}{m^2c}\frac{1}{r}\frac{\partial\phi}{\partial r}\mathbf{s}\cdot\hat{\mathbf{L}} \ .$$

The fact that this reasoning gives the correction larger by a factor 2 than the actual interaction $H_{\mathrm{spin-orbit}}$ is the manifestation of the so-called Thomas precession. Roughly speaking, the Lorentz transformation we have performed was too naive because boosts do not commute with the ordinary space rotations.

We can now consider relativistic corrections to the Hydrogen-like atom spectrum. As they are small (the only one large is a trivial shift of all energy levels by mc^2), we will use the ordinary first-order perturbation calculus. In the absence of external a magnetic field the perturbation consists of three terms:

$$
\begin{aligned}
H^{(1)} &= -\frac{\mathbf{p}^4}{8m^3c^2} \\
H^{(2)} &= -\frac{Zqe^2\hbar^2}{8m^2c^2}4\pi\delta^{(3)}(\mathbf{r}) \\
H^{(3)} &= -\frac{Zqe^2}{2mc^2}\frac{1}{r^3}\mathbf{s}\cdot\mathbf{L}
\end{aligned}
\tag{14.95}
$$

where we have used $\phi = Ze/r$. We will also need the following formulae

$$
\langle\frac{1}{r}\rangle = \frac{Z}{a_0}\frac{1}{n^2}, \qquad \langle\frac{1}{r^2}\rangle = \frac{Z^2}{a_0^2}\frac{1}{n^3}\frac{1}{l+\frac{1}{2}}, \qquad \langle\frac{1}{r^3}\rangle = \frac{Z^3}{a_0^3}\frac{1}{n^3}\frac{1}{l(l+1)(l+\frac{1}{2})}
$$

in which $\langle\ldots\rangle$ denotes the expectation value in the unperturbed nonrelativistic Hydrogen atom state $|n,l,m\rangle$ and $a_0 = \hbar^2/me^2 = \hbar c/mc^2\alpha$ is the Bohr radius. These relations are derived in the Appendix.

Since the operators $H^{(1)}$ and $H^{(2)}$ are rotationally invariant it is clear that $\langle n,l',m',s'_z|H^{(1,2)}|n,l,m,s_z\rangle \propto \delta_{l'l}\delta_{m'm}$ (labels s_z, s'_z stand for the z-axis spin projection). Hence, these corrections will not lift the accidental degeneracy of the states with the same n and different l (discussed in section 4.5).

The diagonal matrix elements of $H^{(1)}$ are most easily computed by using the trick.[93]

$$
\begin{aligned}
\langle -\frac{\mathbf{p}^4}{8m^3c^2}\rangle &= -\frac{1}{2mc^2}\langle (H_0 - qe\phi)^2\rangle \\
&= -\frac{1}{2mc^2}\left[E_n^2 - 2E_nZqe^2\langle\frac{1}{r}\rangle + Z^2q^2e^4\langle\frac{1}{r^2}\rangle\right] \\
&= \frac{1}{2}mc^2(Z^2\alpha^2)\ \frac{Z^2\alpha^2}{n^3}\left(\frac{3}{4n} - \frac{1}{l+\frac{1}{2}}\right)
\end{aligned}
\tag{14.96}
$$

where we have exploited the fact that $H_0 = \mathbf{p}^2/2m + qe^2Z/r$. The correction introduced by the Darwin term $H^{(2)}$ is

$$
\begin{aligned}
-\frac{Zqe^2\hbar^2}{8m^2c^2}\langle 4\pi\delta^{(3)}(\mathbf{r})\rangle &= -4\pi\frac{Zqe^2\hbar^2}{8m^2c^2}|\psi_{nlm}(0)|^2 \\
&= -q\frac{1}{2}mc^2(Z^2\alpha^2)\ \frac{Z^2\alpha^2}{n^3}\delta_{l,0}
\end{aligned}
\tag{14.97}
$$

[93]Direct computation of the $\mathbf{p}^4$ matrix element can lead to a nonsensical result $\langle\mathbf{p}^4\rangle < 0$ because $\mathbf{p}^4|n,l,m\rangle$ may not be in the Hilbert space, i.e. may be a non-normalizable state.

where we have used $|\psi_{nlm}(0)|^2 = (Z^3/\pi a_0^3 n^3)\delta_{l,0}$.

The only perturbation that can lift the accidental degeneracy of the Hydrogen spectrum is $H^{(3)}$. As folows from the example discussed in section 4.4, it becomes diagonal in the basis $|n, j, j_z, l, s\rangle$ states (obviously the perturbations $H^{(1)}$ and $H^{(2)}$ remain diagonal also in the $|n, j, j_z, l, s\rangle$ basis; their diagonal matrix elements are then still given by (14.96) and (14.97)). For $l \neq 0$ one gets:

$$
\begin{aligned}
&-\frac{Zqe^2}{2mc^2}\langle\frac{1}{r^3}\mathbf{s}\cdot\mathbf{L}\rangle \\
&\quad = -q\frac{1}{2}mc^2(Z^2\alpha^2)\ \frac{Z^2\alpha^2}{n^3}\frac{j(j+1)-l(l+1))-s(s+1)}{2l(l+1)(l+\frac{1}{2})}
\end{aligned}
\tag{14.98}
$$

where for $s = 1/2$ one has either $j = l + \frac{1}{2}$ or $j = l - \frac{1}{2}$. Expectation values of $H^{(3)}$ are of course zero in the $|n, l = 0, s_z\rangle$ states.

It is easy to check, that

$$
\frac{j(j+1)-l(l+1))-s(s+1)}{2l(l+1)(l+\frac{1}{2})} = \begin{cases} \frac{1}{l+\frac{1}{2}} - \frac{1}{l+1} & \text{for} \quad j = l + \frac{1}{2} \\ \frac{1}{l+\frac{1}{2}} - \frac{1}{l} & \text{for} \quad j = l - \frac{1}{2} \end{cases}
$$

Therefore, the total correction to the energy of the $|n, j, j_z, l, s\rangle$ Hydrogen atom state is

$$
\Delta E = \frac{1}{2}mc^2(Z^2\alpha^2)\ \frac{Z^2\alpha^2}{n^3}\left(\frac{3}{4n} - \frac{1}{l+\frac{1}{2}} - q\delta_{l,0} - q\left\{\begin{matrix} \frac{1}{l+\frac{1}{2}} - \frac{1}{l+1} \\ \frac{1}{l+\frac{1}{2}} - \frac{1}{l} \end{matrix}\right\}\right)
$$

where the upper line of the last term is for $j = l + \frac{1}{2}$ and the lower one is for $j = l - \frac{1}{2}$. For $l = 0$ the expressions in curly brackets should be set to zero. For $l \neq 0$ the terms with $1/(l + \frac{1}{2})$ cancel out (because for electron $q = -1$) and the terms $1/(l + 1)$ for $j = l + \frac{1}{2}$ and $1/l$ for $j = l - \frac{1}{2}$ both can be written simply as $1/(j + \frac{1}{2})$. Moreover, the resulting term $-q/(j + \frac{1}{2})$ absent for $l = 0$ is in this case (when it would be formally equal $-q$ because then $j + \frac{1}{2} = 1$) provided by the contribution of the Darwin term $-q\delta_{l,0}$. As a result, the correction can be compactly written as

$$
\Delta E = \frac{1}{2}mc^2(Z^2\alpha^2)\ \frac{Z^2\alpha^2}{n^3}\left(\frac{3}{4n} - \frac{1}{j+\frac{1}{2}}\right) \tag{14.99}
$$

Surprisingly, the dependence on l has dissappeared from the final formula. Hence, even with the relativistic corrections there remains some degeneracy of the spectrum: the energy is the same for the pairs of states whose j arises from $l = j - \frac{1}{2}$ and $l = j + \frac{1}{2}$. In other words, degenerate remain the following multilets of states: $2S_{1/2}$ and $2P_{1/2}$, $3S_{1/2}$ and $3P_{1/2}$, $3P_{3/2}$ and $3D_{3/2}$ and so on. Actually, these degeneracies are the property of the full solution to the Dirac equation for electron moving in the Coulomb external potential $\phi = Ze/r$. Exact formula for the electron energy levels in the Hydrogen atom obtained by solving the Dirac equation is

$$E = mc^2 \left[1 + \left(\frac{Z\alpha}{n - (j + \frac{1}{2}) + \sqrt{(j + \frac{1}{2})^2 - Z^2\alpha^2}} \right)^2 \right]^{-1/2} \tag{14.100}$$

(Its expansion in powers of $Z\alpha$ gives (14.99) as a first correction to the nonrelativistic result plus mc^2 term). From the formula (14.100), which was first obtained by Sommerfeld in the framework of the old quantum theory , it is clear that the degeneracy is a property of the full solution to the Dirac equation.

The degeneracy of the states with the same j but different l is however lifted by order $mc^2(Z^2\alpha^2)$ $Z^2\alpha^2\alpha^2$ corrections which can be computed only in the framework of full electrodynamics. The resulting so-called hyperfine splitting of $2S_{1/2}$ and $2P_{1/2}$ states is referred to as the Lamb shift.

14.4 Spin projections

In many applications, for example in studies of polarization in scattering process, we are interested in transition probabilities between initial and/or final states having definite spin projections. In such cases one does not sum transition probabilities over spins of initial and/or final particles. Instead, spinors describing well defined spin states should be used. This is not very convenient because their explicit form can be rather complicated. There is however a simple technique allowing to avoid constructing them explicitly.

Consider first a free massive spin 1/2 particle in its rest frame. It is then easy to write down the operator projecting on the state in which particle has spin projection $\hbar/2$ in the direction specified by the unit vector $\mathbf{u}$:

$$P(u) \equiv \frac{1 + \mathbf{u} \cdot \mathbf{\Sigma}}{2} \ . \tag{14.101}$$

The polarization vector $\mathbf{u}$ can be of course promoted to a four-vector $u^\mu_{\rm rest} = (0, \mathbf{u})$ which can be then transformed to any other inertial frame:

$$u^\mu = \Lambda^\mu{}_\nu u^\nu_{\rm rest} \ . \tag{14.102}$$

Let us recall that the spin label of a moving particle is ascribed to it just by looking at the spin in the particle rest frame. Therefore for a given rest frame spin vector $\mathbf{u}$ we would like to find a projection operator whose eigenstates when transformed to the particle rest frame are just the eigenstates of the operator (14.101). To this end we rewrite (14.101) in a more "covariant way"

$$P(u) \equiv \frac{1 + \gamma^5 \gamma_i u^i_{\rm rest} \gamma^0}{2} = \frac{1 + \gamma^5 \gamma_\mu u^\mu_{\rm rest} \gamma^0}{2} \ . \tag{14.103}$$

Because of γ^0 this is not a truly covariant expression. Redefining it by removing γ^0 we gain however something more: namely we obtain the operator which projects positive energy states on the spin direction $+\mathbf{u}$ and negative energy states on the spin direction $-\mathbf{u}$. Thus,

$$\hat{\Sigma}(s) = \frac{1 + \gamma^5 \gamma_\mu s^\mu}{2} \ . \tag{14.104}$$

where $s^\mu = \Lambda^\mu{}_\nu s^\nu_{\rm rest}$ have the following properties

$$\begin{aligned} \hat{\Sigma}^2(s) &= \hat{\Sigma}(s) \ , \\ \hat{\Sigma}(s)\hat{\Sigma}(-s) &= 0 \ , \\ \hat{\Sigma}(s) + \hat{\Sigma}(-s) &= 1 \ , \end{aligned} \tag{14.105}$$

and

$$\begin{aligned} \hat{\Sigma}(s)u(\mathbf{p}, s) = u(\mathbf{p}, s), &\quad \hat{\Sigma}(-s)u(\mathbf{p}, s) = 0, \\ \hat{\Sigma}(s)v(\mathbf{p}, s) = v(\mathbf{p}, s), &\quad \hat{\Sigma}(-s)v(\mathbf{p}, s) = 0, \end{aligned} \tag{14.106}$$

where $-s \equiv \Lambda\,(-s_{\rm rest})$; $-s^{\mu}_{\rm rest} \equiv (0 - \mathbf{s})$. To prove this it is sufficient to notice that due to the Lorentz transformation properties of the gamma matrices

$$\hat{\Sigma}(s)D(\Lambda) = D(\Lambda)\hat{\Sigma}(s_{\rm rest}) \ . \tag{14.107}$$

The spin projection operator $\hat{\Sigma}(s)$ simplifies the calculations of rates of processes with polarized particles. Consider for example a scattering of a particle having initially spin $\hbar/2$ in the direction of s_1. Insted of using explicit form of $u(\mathbf{p}, s_1)$ in the calculation we can insert $\hat{\Sigma}(s_1)u(\mathbf{p}, s)$ in the calculation of the matrix element and perform the summation over the spin s using the standard formula $\sum_s u(\mathbf{p}, s)\bar{u}(\mathbf{p}, s) = \not{p} + m$. The presence of the spin projector $\hat{\Sigma}(s_1)$ under the trace over the spinor indices will ensure that only the spin state s_1 is included in this way. Because computing traces of the gamma matrices is easier than handling explicit spinors $u(\mathbf{p}, s_1)$ this technique is of great practical value.

14.5 Particle-antiparticle symmetry

As we have discussed, the Dirac interpretation of negative energy solution of his equation as "holes" establishes a one to one correspondence between these solutions and wave functions of antiparticles. That such a correspondence should exist is clear from our construction of the field operators $\psi_\alpha(x)$, $\bar{\psi}_\alpha(x)$, in which we arbitrarily took electron for the "particle" annihilated by the operator $b(\mathbf{p}, s)$ and positron for the "antiparticle" annihilated by the operator $d(\mathbf{p}, s)$. Equally well we could have taken positron for the "particle" and electron for the "antiparticle". Obviously, the physical quantities like the scattering amplitudes should not depend on this choice. Yet, this is not immediately evident from the Feynman rules in which arrows on fermion lines represent the direction of the particle momenta and different spinors are used for final state wave functions of particles and antiparticles. In this subsection demonstrate, that nevertheless the amplitudes computed in the formalism in which electrons are "particles" are the same as in the formalism in which they are treated as "antiparticles".

If the negative energy solution of the Dirac equation for the electron are related to positron states, there should be a transformation allowing to construct a *positive* energy solution $\psi^{\rm pos}_{-q}$ of the equation

$$\left[\gamma_\mu \left(i\partial^\mu - (-qe)A^\mu\right) - m\right] \psi^{\rm pos}_{-q}(t, \mathbf{x}) = 0 \ , \tag{14.108}$$

out of the *negative* energy solution $\psi^{\rm neg}_{+q}$ of the equation

$$[\gamma_\mu\,(i\partial^\mu - qeA^\mu) - m]\,\psi^{\rm neg}_{+q}(t,\mathbf{x}) = 0\ . \tag{14.109}$$

To find this transformation we note that in a static field A_μ, when the time dependence of the solutions $\psi^{\rm pos}_{-q}$ and $\psi^{\rm neg}_{+q}$ corresponding to states of definite energy factorizes:

$$\psi^{\rm neg}_{+q}(t,\mathbf{x}) \propto e^{-i(-|E|)t}\ , \qquad \psi^{\rm pos}_{-q}(t,\mathbf{x}) \propto e^{-i|E|t}\ , \tag{14.110}$$

the transformation has to change the sign of the exponent. It must therefore involve the complex conjugation operation. Performing the complex conjugation of the equation (14.109)

$$\left[-\gamma^*_\mu\,(i\partial^\mu + qeA^\mu) - m\right]\psi^{\rm neg *}_{+q}(t,\mathbf{x}) = 0\ . \tag{14.111}$$

we see that it also changes in the right way the relative sign between A_μ and $i\partial^\mu$. To recover the equation (14.108) it is sufficient to find a matrix C such that

$$(C\gamma^0)\gamma^*_\mu(C\gamma^0)^{-1} = -\gamma_\mu\ , \quad \text{or, equivalently,} \quad C\gamma^{\rm T}_\mu C^{-1} = -\gamma_\mu, \tag{14.112}$$

for then the function

$$\psi^{\rm pos}_{-q} \equiv C\gamma^0\psi^{\rm neg *}_{+q} = C\bar\psi^{\rm T}\ , \tag{14.113}$$

will satisfy the equation (14.108). It is easy to see, that in the two representations of the gamma matrices we have been using

$$C = i\gamma^2\gamma^0 = -C^{-1} = -C^\dagger = -C^{\rm T}\ . \tag{14.114}$$

The transformation (14.113) is therefore the required one.[94]

[94]In textbooks it is called the charge conjugation operation. We refrain from using this term which should be reserved for the operation represented in the multiparticle Hilbert space by the unitary operator $\mathcal{C}$ with the properties

$$\begin{aligned} \mathcal{C}b(\mathbf{p},s)\mathcal{C}^{-1} &= \eta d(\mathbf{p},s)\ , \qquad \mathcal{C}b^\dagger(\mathbf{p},s)\mathcal{C}^{-1} = \eta^* d^\dagger(\mathbf{p},s)\ , \\ \mathcal{C}d(\mathbf{p},s)\mathcal{C}^{-1} &= \eta^* b(\mathbf{p},s)\ , \qquad \mathcal{C}d^\dagger(\mathbf{p},s)\mathcal{C}^{-1} = \eta b^\dagger(\mathbf{p},s)\ . \end{aligned} \tag{14.115}$$

For the plane wave solution

$$\psi^{\text{neg}} = v(\mathbf{p}, s)e^{ip\cdot x} \tag{14.116}$$

of the free Dirac equation it follows from (14.113) that

$$C\gamma^0 v^*(\mathbf{p}, s)e^{-ip\cdot x} \tag{14.117}$$

is the positive energy solution of the same (because in the free equation there is no charge distinguishing particles and antiparticles) equation. Hence the relation

$$C\gamma^0 v^*(\mathbf{p}, s) = e^{i\phi}u(\mathbf{p}, s) \ , \tag{14.118}$$

(where the phase ϕ can depend on the momentum and spin label) should hold. That the spin label on both sides of this relation should be the same can be seen by considering transformation of the negative energy solution having zero eigenvalue of $\hat{\mathbf{p}}$:

$$C\gamma^0 v^*(\mathbf{0}, s = \frac{1}{2}) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = u(\mathbf{0}, s = \frac{1}{2}) \ .$$

Conversly, the operation of complex conjugation followed by multiplication by $C\gamma^0$ transforms the positive energy solution $u(\mathbf{p}, s)e^{-ip\cdot x}$ into the negative energy solution, whence

$$C\gamma^0 u^*(\mathbf{p}, s) = e^{i\phi}v(\mathbf{p}, s) \ . \tag{14.119}$$

Consider now the scattering of positrons in an external electromagnetic field $\mathcal{A}_\mu^{\text{ext}}$. Suppose electrons are taken for particles. In the lowest order of perturbation calculus we are therefore interested in the matrix element

$$\begin{aligned} -i\mathcal{M} &\propto \langle\Omega|d(\mathbf{p}_f, s_f)\left(-i\int d^4x\ \mathcal{H}_I\right) d^\dagger(\mathbf{p}_i, s_i)|\Omega\rangle \\ &\propto -iqe\int d^4x\mathcal{A}_\mu^{\text{ext}}(x)\ \langle\Omega|d(\mathbf{p}_f, s_f) : \bar{\psi}^{(+)}(x)\gamma^\mu\psi^{(-)}(x) : d^\dagger(\mathbf{p}_i, s_i)|\Omega\rangle \\ &\propto +iqe\bar{v}(\mathbf{p}_i, s_i)\gamma^\mu v(\mathbf{p}_f, s_f) \ . \end{aligned} \tag{14.120}$$

The change of sign comes from the normal ordering of the operators (which we use here to get rid of an infinite contribution to the forward scattering amplitude)

$$: \bar{\psi}^{(+)}_{\alpha}(x)\gamma^{\mu}_{\alpha\beta}\psi^{(-)}_{\beta}(x) := -\psi^{(-)}_{\beta}(x)\gamma^{\mu}_{\alpha\beta}\bar{\psi}^{(+)}_{\alpha}(x) \ . \tag{14.121}$$

Up to a sign[95] this is what we obtain from the Feynman rules applied to the diagram shown in figure 24a.

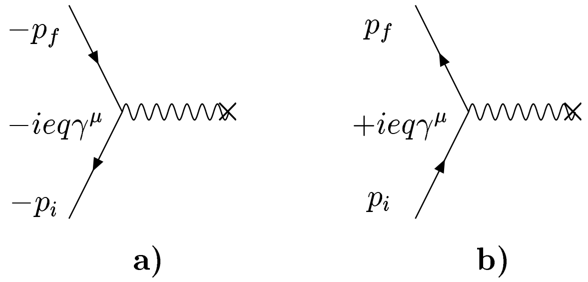


Figure 24: Lowest order contribution to the amplitude of positron scatering in an external field.

Inserting now in the expression (14.120) the relations

$$\begin{aligned} v(\mathbf{p}_f, s_f) &= e^{-i\phi}C\bar{u}^{\mathrm{T}}(\mathbf{p}_f, s_f) \ , \\ \bar{v}(\mathbf{p}_f, s_f) &= -e^{+i\phi}u^{\mathrm{T}}(\mathbf{p}_i, s_i)C^{-1} \ , \end{aligned} \tag{14.122}$$

which follow from the relations (14.118) and (14.119), we get

$$\begin{aligned} -i\mathcal{M} &\propto +iqe\bar{v}(\mathbf{p}_i, s_i)\gamma_{\mu}v(\mathbf{p}_f, s_f) \\ &= iqe\left(-e^{i\phi}\right)u^{\mathrm{T}}(\mathbf{p}_i, s_i)C^{-1}\gamma_{\mu}e^{-i\phi}C\bar{u}^{\mathrm{T}}(\mathbf{p}_f, s_f) \\ &= iqeu^{\mathrm{T}}(\mathbf{p}_i, s_i)\gamma^{\mathrm{T}}_{\mu}\bar{u}^{\mathrm{T}}(\mathbf{p}_f, s_f) = -i(-qe)\bar{u}(\mathbf{p}_f, s_f)\gamma_{\mu}u(\mathbf{p}_i, s_i) \end{aligned} \tag{14.123}$$

which is exactly what we would obtain by taking positrons for particles and computing the amplitude using the Feynman diagram shown in

[95] The overall sign of the amplitude is irrelevant and for this reason one forgets about it using the diagramatic Feynman rules; for pedagogical reason we have derived the amplitude directly from the formula for the S matrix in order to get this sign right as it is necessary for our argument here.

figure24b with the $-ie(-q)\gamma^{\mu} = +ieq\gamma^{\mu}$ rule for the vertex because the "particle" charge would be then $-q$.

It is also instructive to repeat the same for the Compton scattering on positrons. If electrons are taken for the "particles", the lowest order contribution to the scattering amplitude is described by the Feynman diagram shown in figure 25a (there is also the second diagram with photon lines interchanged) in which photon with momentum k_1 is incoming and the one with momentum k_2 is outgoing.

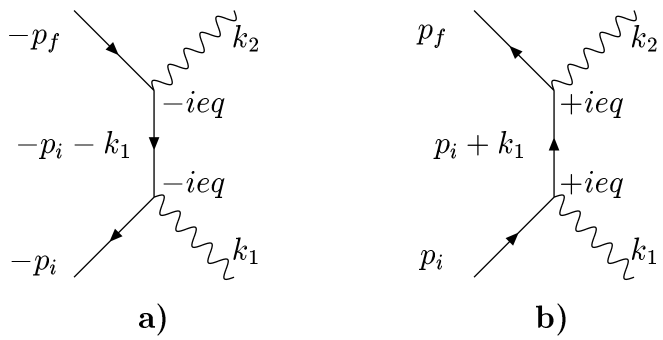


Figure 25: Two different views on the lowest order contribution to the amplitude of Compton scattering on positrons. a) Positron treated as an antiparticle. b) Positron treated as a particle

The corresponding amplitude (neglecting the denominators of the propagators and the "i" factors which are inessential for the argument) reads

$$\begin{aligned}-i\mathcal{M} &\propto \bar{v}(\mathbf{p}_i, s_i)(-ieq)\ \not{\epsilon}(k_1)(-\not{p}_i - \not{k}_1 + m)(-ieq)\ \not{\epsilon}^*(k_2)v(\mathbf{p}_f, s_f)\\ &= -u^{\mathrm{T}}(\mathbf{p}_i, s_i)C^{-1}(-ieq)\ \not{\epsilon}(k_1)(-\not{p}_i - \not{k}_1 + m)(-ieq)\ \not{\epsilon}^*(k_2)C\bar{u}^{\mathrm{T}}(\mathbf{p}_f, s_f)\ .\end{aligned}$$

where in the second line we have used the relations (14.122). Inserting now the CC^{-1} factors judiciously we arrive at

$$\begin{aligned}&u^{\mathrm{T}}(\mathbf{p}_i, s_i)(+ieq)\ \not{\epsilon}^{\mathrm{T}}(k_1)(\not{p}_i + \not{k}_1 + m)^{\mathrm{T}}(+ieq)\ \not{\epsilon}^{*\mathrm{T}}(k_2)\bar{u}^{\mathrm{T}}(\mathbf{p}_f, s_f)\\ &= u(\mathbf{p}_f, s_f)(+ieq)\ \not{\epsilon}^*(k_2)(\not{p}_i + \not{k}_1 + m)(+ieq)\ \not{\epsilon}(k_1)u(\mathbf{p}_i, s_i)\end{aligned}$$

that is, at the result which would be obtained treating positrons as the "particles" and computing the amplitude from the diagram 25b. It is clear that this reasoning extends to arbitrary processes.