Until now the solutions of the Dirac equation with negative energy have been a puzzle. Attempts similar to those we performed with the solutions of the Klein-Gordon equation, where the energy turned out to be positive (by the Lagrange formalism) for solutions with positive and negative time evolution factors, proved unsuccessful (cf. Exercise 2.3). Solutions with negative energy appear almost everywhere when we are concerned with processes of high energy or with strongly localized wave packets (see Exercises 8.4, 8.5). At this point we have to confront this dilemma and find a proper solution!

The existence of solutions with negative energy in the previous interpretation, as single-particle states of the electron, obviously leads to trouble and physical nonsense. Let us consider the electrons in an atom, the spectrum of which is once more given qualitatively in Fig. 12.1.

The bound states directly below the positive energy continuum, with $E < m_0c^2$, are in general in very good agreement with experiments. It is beyond any doubt that these are the bound states of the (one-electron) atom.

An electron in the lowest atomic state $(1s)$ could lose more energy by continuous radiative transitions. Thus an atom would be unstable and, because of the continuous emission of light, a radiation catastrophe would occur. However, such effects have never been observed! If this decay could happen, our world could not exist. Hence we have both a principle to uphold as well as a practical problem to solve to avoid electrons falling off into the states of negative energy. Neglecting the radiation field, the bound-state electrons would be stationary. By switching on the field (of course it is always “switched on”), and the use of radiation theory and of the wave functions found in Exercise 9.6, an infinite transition probability is obtained particularly if one takes into account the infinitely large number of final states in the lower continuum (see Exercise 12.1). But this is of course sheer nonsense! We must find a new physical idea to remove this dilemma, and, in its original form, this was provided by Dirac.¹ He assumed all states of negative energy to be occupied with electrons (see Fig. 12.2).

The vacuum state is defined by the absence of real electrons (electrons in states of positive energy), and all the states of negative energy are filled with electrons. The vacuum state is the energetically deepest stable state, which can be realized under certain conditions (constraints such as, e.g., external fields). In the absence of the field the vacuum represents the lower (negative) continuum (it is also called the “Dirac sea”), whose states are completely occupied with electrons.

12. The Hole Theory

Fig. 12.3. A photon of energy $\hbar \omega > 2m_0c^2$ creates an electron-electron-hole state. The hole is interpreted as a positron; hence the process is just $e^- - e^+$ pair creation electrons. This physical assumption of the negative energy continuum filled with electrons has very important consequences. We perceive at once that the radiation catastrophe mentioned above is now avoided because of the Pauli principle, which forbids transitions of real electrons into occupied lower states. On the other hand an electron of negative energy can absorb radiation. If the energy $\hbar \omega$ of the absorbed photon is greater than the energy gap ($\hbar \omega > 2m_0c^2$), an electron of negative energy can be excited into a state of positive energy (see Fig. 12.3).

In that case we get a real electron and a hole. The hole behaves like a particle with charge $+|e|$, because it can be annihilated by an electron ($e^-$) with charge $-|e|$; thus the hole is the antiparticle of the electron and is named positron ($e^+$). Obviously the creation of an electron and electron hole by photons is to be identified with electron–positron pair creation, with a threshold energy of

$$\hbar \omega = 2m_0c^2.$$  \hfill (12.1)

Alternatively, we call the process where an electron drops into a hole, thereby emitting an appropriate photon, pair annihilation (or matter–antimatter annihilation, or $e^+e^-$ annihilation). The energy balance of the pair creation is

$$\hbar \omega = E_{\text{electron with pos. energy}} - E_{\text{electron with neg. energy}}$$

$$= \left( +c \sqrt{p^2 + m_0c^2c^2} \right) - \left( -c \sqrt{p'^2 + m_0c^2c^2} \right)$$

$$E_{\text{electron}} + E_{\text{positron}}, \quad (12.2)$$

and we can associate the electron with the positive energy

$$E_{\text{electron}} = +c \sqrt{p^2 + m_0c^2c^2}.$$  \hfill (12.3)

So far this is nothing new. What is new, however, is that according to (12.2) we have to give the positron (electron of negative energy) a positive energy, namely

$$E_{\text{positron}} = +c \sqrt{p'^2 + m_0c^2c^2}.$$  \hfill (12.4)

In the special case of vanishing positron momentum $p'$, it follows that the positron has the rest mass

$$(E_{\text{positron}})_{\text{at rest}} = m_0c^2.$$  \hfill (12.5)

Therefore positrons (electron holes) have the same rest mass as electrons but opposite charge (as we have shown above). Similarly we obtain the following momentum balance: the photon has momentum $\hbar k$, which is distributed to electrons and positrons. We conclude that initial total momentum = final total momentum, i.e.

$$\hbar k + (p')_{\text{electron with neg. energy}} = (p)_{\text{electron with pos. energy}}$$  \hfill (12.6)

$$\hbar k = (p)_{\text{electron with pos. energy}} - (p')_{\text{electron with neg. energy}}$$  \hfill (12.7)

and write

$$\hbar k = (p)_{\text{electron}} + (p')_{\text{positron}}.$$  \hfill (12.8)