8

Quantum Mechanics

Summary

In this chapter we will talk about quantum mechanics. The foundation for everything here are the identifications we discussed in Chapter 5. As a first result, we derive the **relativistic energy-momentum relation**.

After a discussion of how the quantum formalism works, we take the non-relativistic limit of the Klein-Gordon equation, because this is the equation of motion for the simplest type of particles: scalars. This results in the famous **Schrödinger equation**. The solution of this equation is interpreted as a probability amplitude and two simple examples are analysed using this **wave-mechanic** approach.

Afterwards, the **Dirac notation** is introduced, which is useful to understand the general structure of quantum mechanics. In this notation, the initial state of a system is denoted by an abstract state vector $|i\rangle$, called **ket**. The probability amplitude for measuring this initial state in a specific final state can then be computed formally by multiplying it with a **bra**, denoted $\langle f|$. The combination of a bra with a ket results in a complex number that is interpreted as probability amplitude A for the process $i \to f$. The probability for this process is $|A|^2$. Then we talk about **projection operators**. We will see how they can be used, together with the **completeness relation**, to expand an arbitrary state in the eigenstate basis of an arbitrary operator. The previously used wave-mechanics can then be seen as a special case, where we expand the states in the location basis. In the Dirac notation the Schrödinger equation is used to compute the time-evolution of states. To make the connection explicit, we use the Dirac notation to reanalyze an example which we already solved using wave-mechanics.

J. Schwichtenberg, *Physics from Symmetry*, Undergraduate
Lecture Notes in Physics, https://doi.org/10.1007/978-3-319-66631-0_8

8.1 Particle Theory Identifications

[1] The Klein-Gordon, Dirac, Proka and Maxwell equations

The equations we derived so far[1] can be used in particle- and field-theories. In this chapter we want to investigate how they can be used in a particle-theory. Our dynamic variables are then the location, the angular momentum, the energy and the momentum of the particle in question. As explained in Chapter 5, we identify these with the generators of the corresponding symmetry[2]:

[2] See Eq. 3.248, Eq. 3.252 and Chapter 5

- momentum $\hat{p}_i = -i\partial_i$,
- location $\hat{x}_i = x_i$,
- energy $\hat{E} = i\partial_0$,
- angular momentum $\hat{L}_i = i\frac{1}{2}\epsilon_{ijk}(x^j\partial^k - x^k\partial^j)$.

Before we discuss how these operators are used in quantum mechanics, we use them to derive one of the most important equations of modern physics.

8.2 Relativistic Energy-Momentum Relation

In Section 6.2 we derived the equation of motion for a free spin 0 particle, the Klein-Gordon equation:

$$(\partial_\mu\partial^\mu + m^2)\Phi = 0.$$

[3] Using $p_\mu = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} p_0 \\ \vec{p} \end{pmatrix} = \begin{pmatrix} E \\ \vec{p} \end{pmatrix}$

With the identifications reiterated above this equation reads[3]

$$\begin{aligned}(\partial_\mu\partial^\mu + m^2)\Phi &= (\partial_0\partial_0 - \partial_i\partial_i + m^2)\Phi \\ &= \left(\left(\frac{1}{i}E\right)\left(\frac{1}{i}E\right) - \left(-\frac{1}{i}p_i\right)\left(-\frac{1}{i}p_i\right) + m^2\right)\Phi \\ &= (-E^2 + \vec{p}^2 + m^2)\Phi = 0\,. \end{aligned} \tag{8.1}$$

Therefore

$$\rightarrow E^2 = \vec{p}^2 + m^2 \quad \text{or using four-vectors} \quad p_\mu p^\mu = m^2\,, \tag{8.2}$$

which is the famous **energy-momentum relation** of special-relativity. For a particle at rest ($\vec{p} = 0$) this gives us Einstein's famous equation

$$E^2 = m^2 \leftrightarrow E = mc^2$$

where we restored c^2 for clarity. We can now understand why we gave the scalar value of the first Casimir operator of the Poincaré group $p_\mu p^\mu$ the suggestive name m^2 (Eq. 3.266). The combination $p_\mu p^\mu$ is indeed the squared mass of the particle in question, which can be measured in experiments, for example, by measuring the energy and momentum of the particle: $m = \sqrt{E^2 - \vec{p}^2}$. For the same reason, we understand now why the constant in the Lagrangian we derived in Section 6.2 is also called m^2.

8.3 The Quantum Formalism

Now that our physical quantities are given by operators, we need something they can act on. First take note that we have for each operator a set of eigenfunctions, which is completely analogous to the eigenvectors of matrices. Matrices are finite-dimensional and therefore we get finite dimensional eigenvectors. Here our operators act on an infinite-dimensional vector space and we therefore have eigenfunctions. For example, the equation we must solve to get the eigenfunctions of the momentum operator is

$$\underbrace{-i\partial_i}_{\text{operator}}\Psi = \underbrace{p_i}_{\text{eigenvalue}}\overbrace{\Psi}^{\text{eigenfunction}}, \tag{8.3}$$

with some number p_i. A solution is

$$\begin{aligned}\Psi &= \underbrace{C}_{=const} e^{ip_i x_i} \quad \text{because} \\ \rightarrow -i\partial_i C e^{ip_i x_i} &= p_i C e^{ip_i x_i} \quad \checkmark\end{aligned} \tag{8.4}$$

but take note that this a solution for arbitrary p_i. Therefore we have found an infinite number of eigenfunctions for the momentum operator $\hat{p}_i = -i\partial_i$. Equivalently, we can search for energy eigenfunctions

$$i\partial_0\Phi = E\Phi \tag{8.5}$$

or angular momentum eigenfunctions[4]. Analogous to eigenvectors for matrices, these eigenfunctions are bases for the corresponding abstract vector space[5]. This means we can expand an arbitrary function Ψ in terms of eigenfunctions. For example, we can expand Ψ in terms of momentum eigenfunctions[6] (for brevity in one dimension)

$$\Psi = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dp \Psi_p e^{-ipx}, \tag{8.6}$$

where Ψ_p are the coefficients in this expansion analogous to v_1, v_2, v_3 in $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$.

For some systems we have boundary conditions such that we have a discrete instead of a continuous basis. Then we can expand an arbitrary state, for example, in terms of discrete set of energy eigenfunctions Φ_{E_n}:

$$\Psi = \sum_n c_n \Phi_{E_n}. \tag{8.7}$$

An important observation is that, in general, a set of eigenfunctions for one operator is not a set of eigenfunctions for another operator.

[4] The discussion for angular momentum eigenfunctions is a bit more complicated, because the operator is more complicated than the others. We can't find a set of eigenfunctions for all three components at the same time, because $[\hat{L}_i, \hat{L}_j] \neq 0$. This will be discussed in a moment. The final result of a lengthy discussion is that the corresponding eigenfunctions of the third component of the angular momentum operator $\hat{L}_3$ (and of the squared angular momentum operator $\hat{L}^2$, which commutes with all components $[\hat{L}^2, \hat{L}_j] = 0$) are the famous **spherical harmonics**. These form an orthonormal basis.

[5] Recall that for matrices the eigenvectors are a basis for the corresponding vector space.

[6] Take note that this is exactly the Fourier transform of a function Ψ, which is introduced in Appendix D.1 and the factor $\frac{1}{\sqrt{2\pi}}$ is a matter of convention.

Only for operators that commute $[A, B] = AB - BA = 0$, we can find a simultaneous set of eigenfunctions for both operators. We don't prove this here, but to see the connection between a common set of eigenfunctions and the commutator property, let's assume we have two commutating operators $[A, B] = 0 \rightarrow AB = BA$. For the set of eigenfunctions Ψ_i that belong to A, we have $A\Psi_i = a_i\Psi_i$, where a_i denotes the corresponding eigenvalues. Now, we can calculate[7]

[7] The brackets () are used here solely for illustrative purposes.

$$A(B\Psi_i) \underbrace{=}_{AB=BA} BA\Psi_i \underbrace{=}_{A\Psi_i = a_i\Psi_i} Ba_i\Psi_i \overbrace{=}^{a_i \text{ is just a number}} a_i(B\Psi_i)\,. \tag{8.8}$$

This shows that $(B\Psi_i)$ is also an eigenfunction of A for all Ψ_i, with exactly the same eigenvalues a_i. Therefore the only difference between $(B\Psi_i)$ and Ψ_i can be a constant: $B\Psi_i = b_i\Psi_i$. We conclude that the A eigenfunctions Ψ_i are also eigenfunctions of B if $AB = BA$ holds.

[8] The usage of Ψ is conventional in quantum mechanics and we use it here, although we used it so far exclusively for spinors, to describe spin 0 particles, too.

In general our operators act on something we call[8] Ψ, which denotes the **state** of the physical system in question. We get this Ψ, by solving the corresponding equation of motion.

[9] This means all other coefficients in the expansion $\Psi = \sum_n c_n\Phi_{E_n}$ are zero.

Usually, such a solution will have more than one term if we expand it in some basis. For example, consider a state that can be written in terms of two energy eigenstates[9] $\Psi = c_1\Phi_{E_1} + c_2\Phi_{E_2}$. Acting with the energy operator on this state yields

$$\hat{E}\Psi = \hat{E}(c_1\Phi_{E_1} + c_2\Phi_{E_2}) = c_1E_1\Phi_{E_1} + c_2E_2\Phi_{E_2} \neq E(c_1\Phi_{E_1} + c_2\Phi_{E_2}). \tag{8.9}$$

A **superposition** of states with different energy is therefore, in general, no eigenstate of the energy operator, because for an eigenstate we have by definition $\hat{E}\Psi = E\Psi$ for some number E. But what is then the energy of the system described by Ψ? What does it mean that a state is a superposition of two energy eigenstates? How can we interpret all this in physical terms?

A first hint towards an interpretation is the $U(1)$ symmetry of our Lagrangians, which shows us that the solution of an equation of motion Ψ cannot be directly physically relevant[10]. Secondly, we observe that a solution to any equation of motion we derived so far is a function of[11] $\vec{x}$ and t, i.e. $\Psi = \Psi(\vec{x}, t)$.

[10] If we assume that Ψ describes our particle directly in some way, what would the $U(1)$ transformed solution $\Psi' = e^{i\alpha}\Psi$, which is equally allowed, describe?

[11] This will be made explicit in the next sections.

The standard interpretation is that the absolute value squared $|\Psi(\vec{x}, t)|^2$ of the **wave function** $\Psi(\vec{x}, t)$ gives the **probability density** of its location. Observe that the $U(1)$ symmetry has no influence on this quantity $|\Psi|^2 = \Psi^\dagger\Psi \rightarrow (\Psi')^\dagger(\Psi)' = \Psi^\dagger e^{-i\alpha}e^{i\alpha}\Psi = \Psi^\dagger\Psi$. In

other words: $\Psi(x,t)$ is the probability amplitude that a measurement of the location gives a value in the interval $[x, x+dx]$. Consequently we have, if we integrate over all space

$$\int dx \Psi^\star(x,t)\Psi(x,t) \stackrel{!}{=} 1, \tag{8.10}$$

which is called the normalization prescription, because the probability for finding the particle anywhere in space must be $100\% = 1$.

If we want to make predictions about any other physical quantity, we must expand the wave-function in terms of the corresponding basis. For example, we can expand it in terms of energy eigenfunctions: $\Psi = c_1\Phi_{E_1} + c_2\Phi_{E_2} + \ldots$. Then, the standard interpretation of quantum mechanics is: the probability for measuring a given energy value E_1 for the system described by Ψ is given by the absolute value squared of the overlap between Ψ and Φ_{E_1}

$$P(E_1) = \left| \int dx \Phi^\star_{E_1}(x,t)\Psi(x,t) \right|^2 .$$

In the example above this means

$$\begin{aligned} P(E_1) &= \left| \int dx \Phi^\star_{E_1}(x,t)\Psi(x,t) \right|^2 = \left| \int dx \Phi^\star_{E_1}(x,t) \left(c_1\Phi_{E_1} + c_2\Phi_{E_2}\right) \right|^2 \\ &= \Big| c_1 \underbrace{\int dx \Phi^\star_{E_1}(x,t)\Phi_{E_1}}_{=1 \text{ as explained above}} + c_2 \underbrace{\int dx \Phi^\star_{E_1}(x,t)\Phi_{E_2}}_{=0 \text{ because eigenstates are orthogonal}} \Big|^2 \\ &= |c_1|^2 . \end{aligned} \tag{8.11}$$

Analogously, when we expand some other $\Psi(\vec{x},t)$ in terms of momentum eigenfunctions

$$\Psi(\vec{x},t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dp \tilde{\Psi}(\vec{p},t) \mathrm{e}^{-i\vec{p}\vec{x}},$$

we have $\tilde{\Psi}(\vec{p},t)$ as the probability amplitude for finding the system with momentum in the interval $[p, p+dp]$.

This interpretation can be used to make probabilistic predictions about the system, for example using the statistical expectation value, which is the topic of the next section. Afterwards we will derive the equation of motion for non-relativistic quantum mechanics and look at two examples.

8.3.1 Expectation Value

In statistics the expectation value is defined in analogy to the weighted average. For example, if tossing a dice ten times results in 2, 4, 1, 3, 3,

6, 3, 1, 4, 5, the average value is

$$\langle x\rangle = (2+4+1+3+3+6+3+1+4+5)\cdot\frac{1}{10} = 3.2.$$

An alternative way of computing this is to collect equal results and weighing them by their empirical probability:

$$\langle x\rangle = \frac{2}{10}\cdot 1+\frac{1}{10}\cdot 2+\frac{3}{10}\cdot 3+\frac{2}{10}\cdot 4+\frac{1}{10}\cdot 5+\frac{1}{10}\cdot 6 = 3.2.$$

We can write this in general as

$$\langle x\rangle = \sum_i \rho_i x_i \tag{8.12}$$

where ρ_i denotes the probability. Equally for a continuous distribution we have

$$\langle x\rangle = \int dx\rho(x)x \tag{8.13}$$

In quantum mechanics the **expectation value** for a physical quantity $\hat{O}$ is defined analogously

$$\langle\hat{O}\rangle = \int d^3x\Psi^\star\hat{O}\Psi. \tag{8.14}$$

In general, we must expand Ψ in terms of eigenfunctions of $\hat{O}$, for example momentum eigenfunctions. Then, acting with the operator $\hat{O}$ on these states yields the corresponding eigenvalues and we get a weighted sum.

For example, the expectation value for the location of some particle is

$$\langle\hat{x}\rangle = \int d^3x\Psi^\star\hat{x}\Psi = \int d^3x\Psi^\star x\Psi = \int d^3x x\underbrace{\Psi^\star\Psi}_{\text{probability density of its location}}. \tag{8.15}$$

In the next section, we consider the non-relativistic limit of the Klein-Gordon equation and will derive this way the equation of motion of non-relativistic quantum mechanics[12].

[12] It is also possible to consider the non-relativistic limit of the other equations that we have derived so far. For example, the Dirac equation becomes in the non-relativistic limit the Pauli equation. However, these additional equations are less important for the goals of this book.

8.4 The Schrödinger Equation

The Klein-Gordon equation is solved by plane waves

$$\Phi = \mathrm{e}^{\pm ip_\mu x^\mu} \equiv \mathrm{e}^{\pm ip\cdot x},$$

where $p_\mu = (E,\vec{p})^T$ is the conserved four-momentum of the particle. We check

$$\begin{aligned} 0 &= (\partial_\mu\partial^\mu + m^2)\Phi \\ &= (\partial_\mu\partial^\mu + m^2)\mathrm{e}^{\pm ip_\mu x^\mu} \\ &= (i^2 p_\mu p^\mu + m^2)\mathrm{e}^{\pm ip_\mu x^\mu} = 0 \\ &= (-m^2+m^2)\mathrm{e}^{\pm ip_\mu x^\mu} = 0 \qquad \checkmark \end{aligned} \tag{8.16}$$

We now write the solution with a minus sign a little differently

$$\Phi = e^{-ip_\mu x^\mu} = \Phi = e^{i(-Et+\vec{p}\cdot\vec{x})}$$

and then we see that the time dependence is given by e^{-iEt}, i.e. $\Phi \propto e^{-iEt}$. From Eq. 8.2 we know the energy is

$$E = \sqrt{\vec{p}^2 + m^2} = \sqrt{m^2(\frac{\vec{p}^2}{m^2} + 1)} = m\sqrt{\frac{\vec{p}^2}{m^2} + 1}.$$

In the non-relativistic limit $|\vec{p}| \ll m$, i.e. for an object that moves much slower than the speed of light, which implies that its momentum is much smaller than its mass, we can approximate the energy using the Taylor-series as

$$\begin{aligned} E &= m\sqrt{\frac{\vec{p}^2}{m^2} + 1} \\ &= m(1 + \frac{1}{2}\frac{\vec{p}^2}{m^2} + \ldots) \\ &\approx \underbrace{m}_{\text{rest-mass}} + \underbrace{\frac{\vec{p}^2}{2m}}_{\text{kinetic energy}}. \end{aligned} \tag{8.17}$$

We can therefore write

$$\Phi = e^{i(-Et+\vec{p}\cdot\vec{x})} \approx e^{-imt}\underbrace{e^{i\vec{p}\cdot\vec{x}-i(\vec{p}^2/2m)t}}_{\equiv\phi(\vec{x},t)} = e^{-imt}\phi(\vec{x},t). \tag{8.18}$$

From $|\vec{p}| \ll m$ it follows that the rest-mass is much bigger than the kinetic energy and therefore the remaining time dependence in $\phi(\vec{x},t)$ oscillates more slowly than e^{-imt}. If we put our ansatz into the Klein-Gordon equation we get

$$(\partial_\mu\partial^\mu + m^2)e^{-imt}\phi(\vec{x},t) = (\partial_t\partial^t - \partial_i\partial^i + m^2)e^{-imt}\phi(\vec{x},t) = 0.$$

Then we use $\partial_t e^{-imt}(\ldots) = e^{-imt}(-im + \partial_t)(\ldots)$, which is just the product rule, twice. This yields

$$e^{-imt}((-im + \partial_t)^2 - \partial_i\partial^i + m^2)\phi(\vec{x},t) = 0,$$

which we can divide by e^{-imt}, because this never becomes zero. Therefore

$$\begin{aligned} &\rightarrow ((-im + \partial_t)^2 - \nabla^2 + m^2)\phi(\vec{x},t) = 0 \\ &\rightarrow (-m^2 - 2im\partial_t + (\partial_t)^2 - \nabla^2 + m^2)\phi(\vec{x},t) = 0\,. \end{aligned}$$

Comparing now the third term

$$\begin{aligned} (\partial_t)^2\phi(\vec{x},t) &= (\partial_t)^2\exp\left[i\vec{p}\cdot\vec{x} - i(\vec{p}^2/2m)t\right] \\ &= (\vec{p}^2/2m)^2\exp\left[i\vec{p}\cdot\vec{x} - i(\vec{p}^2/2m)t\right] \propto \frac{p^4}{m^2} \end{aligned} \tag{8.19}$$

with the second term

$$\begin{aligned} im\partial_t\phi(\vec{x},t) &= im\partial_t \exp\left[i\vec{p}\cdot\vec{x} - i(\vec{p}^2/2m)t\right] \\ &= m(\vec{p}^2/2m)\exp\left[i\vec{p}\cdot\vec{x} - i(\vec{p}^2/2m)t\right] \propto p^2 \end{aligned} \tag{8.20}$$

shows us that, because of $|\vec{p}| \ll m$, we can neglect the third term in this limit. Therefore

$$(-2im\partial_t - \nabla^2)\phi(\vec{x},t) = 0$$

$$\underbrace{\rightarrow}_{\text{dividing by (-2m)}} (i\partial_t + \frac{1}{2m}\nabla^2)\phi(\vec{x},t) = 0$$

$$\rightarrow (i\partial_t + \frac{\nabla^2}{2m})\phi(\vec{x},t) = 0, \tag{8.21}$$

which is the famous **Schrödinger equation**. When we now make the identifications we recited at the beginning of this chapter, the equation reads

$$\rightarrow (E - \underbrace{\frac{\vec{p}^2}{2m}}_{=\text{kinetic energy}})\phi(\vec{x},t) = 0$$

$$\rightarrow E = \frac{\vec{p}^2}{2m}, \tag{8.22}$$

which is the usual **non-relativistic energy-momentum relation**. From this point-of-view, it is easy to see how we can include an external potential, because movement in an external potential simply adds a term describing the potential energy to the energy equation:

$$\rightarrow E = \frac{\vec{p}^2}{2m} + V$$

A famous example is the potential of a harmonic oscillator $V = -kx^2$. It is conventional to rewrite the Schrödinger equation, using the Hamiltonian operator $\hat{H}$, which collects all contributing energy operators, for example, the operator for the kinetic energy $\frac{\nabla^2}{2m}$ and the operator for the potential energy $\hat{V}$. Then we have

$$i\partial_t\phi(\vec{x},t) = \underbrace{\left(\frac{\nabla^2}{2m} + \hat{V}\right)}_{\equiv\hat{H}}\phi(\vec{x},t) \rightarrow i\partial_t\phi(\vec{x},t) = \hat{H}\phi(\vec{x},t). \tag{8.23}$$

This is the standard way to write the Schrödinger equation.

8.4.1 Schrödinger Equation with an External Field

We can also follow the same procedure as in the last section and derive the non-relativistic limit of the **interacting** Klein-Gordon equation (Eq. 7.42), i.e. the equation that describes the interaction between a massive spin 0 field ϕ and a massless spin 1 field A_μ (the photon field). The resulting equation is

$$\left(i\partial_t - \frac{1}{2m}\left(\nabla - iq\vec{A}\right)^2 - qA_0\right)\phi(\vec{x},t) = 0\,. \tag{8.24}$$

8.5 From Wave Equations to Particle Motion

Now, let's look at two examples of how the standard interpretation of quantum mechanics works in practice.

8.5.1 Example: Free Particle

A solution for the free (without external potential) Schrödinger equation (Eq. 8.21) is given by

$$\Psi = \mathrm{e}^{-i(Et-\vec{p}\vec{x})} \tag{8.25}$$

because

$$\begin{aligned} i\partial_t \mathrm{e}^{-i(Et-\vec{p}\vec{x})} &\stackrel{!}{=} -\frac{\nabla^2}{2m}\mathrm{e}^{-i(Et-\vec{p}\vec{x})} \\ \rightarrow E\mathrm{e}^{-i(Et-\vec{p}\vec{x})} &\stackrel{!}{=} \frac{\vec{p}^2}{2m}\mathrm{e}^{-i(Et-\vec{p}\vec{x})} \quad \checkmark\,, \end{aligned} \tag{8.26}$$

where E is just the numerical value for the total energy of the free particle, which was derived in Eq. 8.22 as $E = \frac{\vec{p}^2}{2m}$. A more general solution is a linear combination

$$\Psi = A\mathrm{e}^{-i(Et-\vec{p}\vec{x})} + B\mathrm{e}^{-i(E't-\vec{p}'\vec{x})} + \ldots.$$

We interpret the wave-function as a probability amplitude and therefore it must be normalized, because the total probability for finding the particle must be 100% $= 1$. This is not possible for the wave-function above which spreads out over all of space. To describe an individual free particle we have to use a suited linear combination, called a wave-packet:

$$\Psi_{WP}(\vec{x},t) = \int dp^3 A(\vec{p}) e^{i(\vec{p}\vec{x}-Et)}, \tag{8.27}$$

where the complex numbers $A(\vec{p})$ have to be chosen in a way that makes the wave packet normalizable. One possibility is a Gaussian

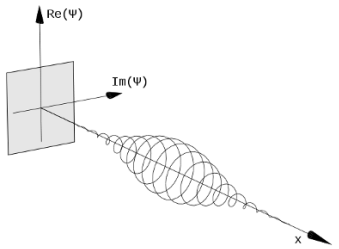

Fig. 8.1: Free wave-packet with Gaussian envelope. Figure by Inductiveload (Wikimedia Commons) released under a public domain licence. URL: `http://commons.wikimedia.org/wiki/File:Travelling_Particle_Wavepacket.svg`, Accessed: 4.5.2014

wave-packet, where $A(\vec{p})$ is a Gauss distribution.

$$\Psi_{GWP}(\vec{x},t) = \int dp^3 A(\vec{p}) e^{i(\vec{p}\vec{x}-Et)} = \int dp^3 \psi_0 e^{i(\vec{p}-\vec{\bar{p}})^2/4\sigma^2} \mathrm{e}^{i(\vec{p}\vec{x}-Et)} .$$

An example of such a Gaussian wave-packet is plotted in Fig. 8.1. For many computations clever tricks can be used in order to avoid working with complicated wave packets, allowing us to work with simple wave functions instead.

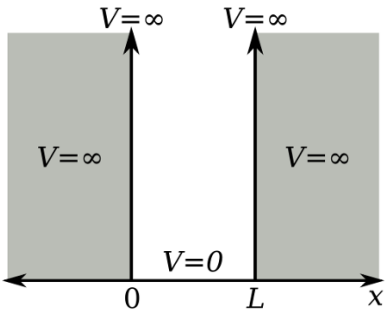

Fig. 8.2: Infinite potential well. Figure by Benjamin D. Esham (Wikimedia Commons) released under a public domain licence. URL: `http://commons.wikimedia.org/wiki/File:Infinite_potential_well.svg` , Accessed: 4.5.2014

8.5.2 Example: Particle in a Box

Now we look at one of the standard examples of quantum mechanics: a particle confined in a box, here 1-dimensional, with infinitely high potential walls. Inside the box the potential is zero, outside it's infinite (see Fig. 8.2).

The potential is defined piece-wise

$$V(x) = \begin{cases} 0, & 0 < x < L \\ \infty, & \text{otherwise} \end{cases} \tag{8.28}$$

and therefore, we have to solve the one-dimensional Schrödinger equation

$$i\partial_t \Psi(x,t) = -\frac{\partial_x^2}{2m}\Psi(x,t) + V(x)\Psi(x,t)$$

piece-wise.

- Inside the box, the solution is equal to the free particle solution, because $V(x) = 0$ for $0 < x < L$.
- Outside, because here $V(x) = \infty$, the only possible, physical solution is $\Psi(x,t) = 0$.

We can rewrite the general free particle solution as[13]

$$\begin{aligned} \Psi(x,t) &= A\mathrm{e}^{-i(Et-px)} + B\mathrm{e}^{-i(Et+px)} \\ &= \big(C\sin(px) + D\cos(px)\big)\mathrm{e}^{-iEt} , \end{aligned}$$

which we can rewrite again using the non-relativistic energy-momentum relation (Eq. 8.22)

$$E = \frac{p^2}{2m} \qquad \rightarrow \qquad p = \sqrt{2mE}$$

$$\Psi(x,t) = \big(C\sin(\sqrt{2mE}x) + D\cos(\sqrt{2mE}x)\big)\mathrm{e}^{-iEt} . \tag{8.29}$$

Next we use that the wave-function must be a continuous function. If there are any jumps in the wave-function, the momentum of

[13] Using $\sin(x) = \frac{1}{2i}\left(\mathrm{e}^{ix} - \mathrm{e}^{-ix}\right)$ and $\cos(x) = \frac{1}{2}\left(\mathrm{e}^{ix} + \mathrm{e}^{-ix}\right)$, which follows directly from the series expansion of $\cos(x), \sin(x)$ and e^{ix} as derived in Appendix B.4.1.

the particle $\hat{p}_x\Psi = -i\partial_x\Psi$ is infinite, because the derivative at the jump would be infinite. Therefore, we have the boundary conditions $\Psi(0) = \Psi(L) \stackrel{!}{=} 0$. Because $\cos(0) = 1$, we see that we have $D \stackrel{!}{=} 0$. Furthermore, we see that these conditions impose

$$\sqrt{2mE_n} \stackrel{!}{=} \frac{n\pi}{L}, \tag{8.30}$$

with arbitrary integer n. This follows because for[14]

$$\Phi_n(x,t) = C\sin(\frac{n\pi}{L}x)\mathrm{e}^{-iE_nt} \tag{8.31}$$

both boundary conditions are satisfied

$$\rightarrow \Phi_n(L,t) = C\sin(\frac{n\pi}{L}L)\mathrm{e}^{-iE_nt} = C\sin(n\pi)\mathrm{e}^{-iE_nt} = 0 \qquad \checkmark$$

$$\rightarrow \Phi_n(0,t) = C\sin(\frac{n\pi}{L}0)\mathrm{e}^{-iE_nt} = C\sin(0)\mathrm{e}^{-iE_nt} = 0 \qquad \checkmark$$

The normalization constant C, can be found to be $C = \sqrt{\frac{2}{L}}$, because the probability for finding the particle anywhere inside the box must be $100\% = 1$ and the probability outside is zero, because there we have $\Psi = 0$. Therefore

$$P = \int_0^L dx\Phi_n^\star(x,t)\Phi_n(x,t) \stackrel{!}{=} 1$$

$$P = \int_0^L dxC^2\sin(\frac{n\pi}{L}x)\mathrm{e}^{+iE_nt}\sin(\frac{n\pi}{L}x)\mathrm{e}^{-iE_nt}$$

$$= C^2\int_0^L dx\sin^2(\frac{n\pi}{L}x) = C^2\left[\frac{x}{2} - \frac{\sin(\frac{2n\pi}{L}x)}{4\frac{n\pi}{L}}\right]_0^L$$

$$= C^2\left(\frac{L}{2} - \frac{\sin(\frac{2n\pi}{L}L)}{4\frac{n\pi}{L}}\right) = C^2\frac{L}{2} \stackrel{!}{=} 1$$

$$\rightarrow C^2 \stackrel{!}{=} \frac{2}{L}.$$

We can now solve Eq. 8.30 for the energy E

$$E_n \stackrel{!}{=} \frac{n^2\pi^2}{L^2 2m}. \tag{8.32}$$

We see here that the **possible energies are quantized**, which means that only a discrete set of values is allowed and not arbitrary continuous values[15]. Hence the name quantum mechanics.

Take note that we have a solution for each n and linear combinations of the form

$$\Phi(x,t) = A\Phi_1(x,t) + B\Phi_2(x,t) + ...$$

[14] Take note that we put an index n on our wave-function, because we have a different solution for each n.

[15] Recall that n is necessarily an integer. The possible energy values are $\frac{\pi^2}{L^2 2m}$ times this integer squared.

are solutions, too. These solutions have to be normalised again because of the probabilistic interpretation[16].

[16] A probability of more than $1 = 100\%$ doesn't make sense.

Next we can ask[17], what is the probability for measuring the particle having Energy $E = E_2 = \frac{2^2\pi^2}{L^2 2m}$. Say our particle is in the normalized state given by

[17] Instead of what's the probability to find a particle at place x.

$$\Psi(x,t) = \sqrt{\frac{3}{5}}\Phi_2(x,t) + \sqrt{\frac{2}{5}}\Phi_3(x,t).$$

The answer in the conventional interpretation of quantum mechanics is: it is the absolute value squared of the overlap between Ψ and Φ_2

$$P\left(E = \frac{2^2\pi^2}{L^2 2m}\right) = \Big| \int dx \Phi_2^\star(x,t)\Psi(x,t)\Big|^2,$$

where the overlap can be seen as a scalar product[18] of Φ_2 and Ψ:

[18] In fact, this is the scalar product of the Hilbert space in which our state vectors Ψ, Φ_n live.

$$(\Phi_2, \Psi) = \int dx \Phi_2^\star \Psi = \underbrace{c}_{\text{complex number}}.$$

The computation is easy, because the solutions we found are orthogonal, i.e.

$$\int dx \Phi_n^\star(x,t)\Phi_{n'}(x,t) = \delta_{nn'}\,.$$

For example[19],

[19] You can check this easily using integration by parts or something like Wolframalpha.com

$$\begin{aligned}\int_0^L dx \Phi_2^\star \Phi_3(x,t) &= \int_0^L dx C\sin(\frac{2\pi}{L}x)\mathrm{e}^{+iEt} C\sin(\frac{3\pi}{L}x)\mathrm{e}^{-iEt} \\ &= C^2\int_0^L dx \sin(\frac{2\pi}{L}x)\sin(\frac{3\pi}{L}x) = 0.\end{aligned}$$

Therefore, we get the probability for finding the energy $E = \frac{2^2\pi^2}{L^2 2m}$

$$\begin{aligned}P\left(E = \frac{2^2\pi^2}{L^2 2m}\right) &= \Big| \int dx \Phi_2^\star(x,t)\Psi(x,t)\Big|^2 \\ &= \Big| \int dx \Phi_2^\star(x,t)\left(\sqrt{\frac{3}{5}}\Phi_2(x,t) + \sqrt{\frac{2}{5}}\Phi_3(x,t)\right)\Big|^2 \\ &= \Big| \int dx \left(\sqrt{\frac{3}{5}}\underbrace{\Phi_2^\star(x,t)\Phi_2(x,t)}_{=1 \text{ if integrated}} + \sqrt{\frac{2}{5}}\underbrace{\Phi_2^\star(x,t)\Phi_3(x,t)}_{=0 \text{ if integrated}}\right)\Big|^2 \\ &= \left(\sqrt{\frac{3}{5}}\right)^2. \end{aligned} \tag{8.33}$$

Take note that we are able to call the functions we just found in Eq. 8.31, eigenstates of the energy operator $i\partial_t$ or equivalently of the Hamiltonian[20] $\hat{H} \equiv -\frac{\partial_x^2}{2m}$, because

[20] This follows directly from the Schrödinger equation $i\partial_t\Phi = -\frac{\partial_x^2}{2m}\Phi \equiv \hat{H}\Phi$.

$$\hat{H}\Phi_n = E_n\Phi_n. \tag{8.34}$$

If we act with the energy operator on an eigenstate, we get the same state multiplied with a constant, which we call energy of the state. In contrast, an arbitrary state is changed when the energy operator or the Hamiltonian operator $\hat{H}$ act on it. For example, if we take a look at the linear combination

$$\Psi = \sqrt{\frac{3}{5}}\Phi_2 + \sqrt{\frac{2}{5}}\Phi_3$$

we see that

$$\hat{H}\Psi = \hat{H}\left(\sqrt{\frac{3}{5}}\Phi_2 + \sqrt{\frac{2}{5}}\Phi_3\right) \underbrace{=}_{\text{Eq. 8.34}} \sqrt{\frac{3}{5}}E_2\Phi_1 + \sqrt{\frac{2}{5}}E_3\Phi_3$$

which cannot be written as multiple of Ψ because $E_2 \neq E_3$. Therefore, Ψ is not an eigenstate of the energy operator. Nevertheless, every wave function can be expressed in terms of the eigenstates Φ_n, because they form a complete basis set.

Next we want to introduce a useful notation invented by Dirac, which extremely helpful to understand the structure of quantum mechanics.

8.5.3 Dirac Notation

In the Dirac notation the state of a physical system is denoted abstractly by

$$|\Psi\rangle\,, \tag{8.35}$$

which is called a **ket**[21]. For example, if we prepare a particle in a box in an energy eigenstate we have

[21] The pun will become clear in a second.

$$\hat{H}\,|\Phi_n\rangle = E_n\,|\Phi_n\rangle\ .$$

To each ket we can define a **bra**, denoted by $\langle\Psi|$ which is given by

$$\langle\Psi| = |\Psi\rangle^\dagger\,, \tag{8.36}$$

where the † symbol (called dagger) denotes the Hermitian conjugate, i.e. transposing plus complex conjugation. A bra is an object that acts on a ket. We can define an inner product as follows:

$$(|\Phi\rangle\,, |\Psi\rangle) \equiv \langle\Phi|\Psi\rangle\ .$$

If a ket is multiplied with a **bra** from the left-hand side, the result is a complex number

$$\langle\Phi|\Psi\rangle = c\,. \tag{8.37}$$

This complex number is the probability amplitude for a physical system in the state $|\Psi\rangle$ to be measured in the state $|\Phi\rangle$. Consequently, the probability is given by $|\langle\Phi|\Psi\rangle|^2$. For example, the probability amplitude for finding a particle in the state $|\Psi\rangle$ in the interval $[x, x+dx]$ is given by

$$\langle x|\Psi\rangle \equiv \Psi(x).$$

This is the wave function we used in the last chapters. Furthermore, we could ask: What's the probability amplitude for finding the same particle with momentum in the interval $[p, p+dp]$? The answer in the Dirac notation is

$$\langle p|\Psi\rangle \equiv \Psi(p).$$

Recall that our states must fulfil a normalization condition, because we are using a probabilistic interpretation[22]. For example, we have

$$\int dx|\Psi(x,t)|^2 = \int dx\Psi^\star(x,t)\Psi(x,t) \stackrel{!}{=} 1 \tag{8.38}$$

and equally

$$\int dp|\Phi(p,t)|^2 = \int dp\Phi^\star(p,t)\Phi(p,t) \stackrel{!}{=} 1 \tag{8.39}$$

or written in our new notation

$$\int dx|\langle x|\Psi\rangle|^2 = \int dx(\langle x|\Psi\rangle)^\dagger\langle x|\Psi\rangle = \int dx\langle\Psi|x\rangle\langle x|\Psi\rangle \stackrel{!}{=} 1 \tag{8.40}$$

and

$$\int dp|\langle p|\Phi\rangle|^2 = \int dp(\langle p|\Phi\rangle)^\dagger\langle p|\Phi\rangle = \int dp\langle\Phi|p\rangle\langle p|\Phi\rangle \stackrel{!}{=} 1\,. \tag{8.41}$$

We can see here a new kind of operator: $|p\rangle\langle p|$ and $|x\rangle\langle x|$, which are called **projection operators**[23]. They are operators because they transform a ket into another ket. For example,

$$|x\rangle \underbrace{\langle x|\Psi\rangle}_{=\text{some complex function we call }\Psi(x)} = |x\rangle\Psi(x),$$

which is again a ket, because the product of a complex function with a ket is again a ket. In general, an operator is any object that acts on a ket to generate another ket. Looking at Eq. 8.40 motivates us to introduce another operator

$$\int dx\langle\Psi|x\rangle\langle x|\Psi\rangle = \langle\Psi|\underbrace{\left(\int dx|x\rangle\langle x|\right)}_{\equiv\hat{I}}|\Psi\rangle = \langle\Psi|\hat{I}|\Psi\rangle \stackrel{!}{=} 1. \tag{8.42}$$

From this we can conclude

$$\hat{I}|\Psi\rangle \stackrel{!}{=} |\Psi\rangle \tag{8.43}$$

[22] The probability for finding the particle anywhere must be 100% = 1 or equally the probability for finding the particle with any momentum must be 100% = 1. In other words: the sum of probabilities for all possible outcomes must add up to 1.

[23] Exactly like the projection operators for left-chiral and right-chiral spinors we introduced earlier, these projection operators fulfil the defining condition $P^2 = P$.

for an arbitrary ket $|\Psi\rangle$, because $\langle\Psi|\Psi\rangle = 1$. This follows, because the probability for a system prepared in state the $|\Psi\rangle$ to be found in the state $|\Psi\rangle$, must be, of course, 100% $= 1$. For example, if we prepare a particle to be at some point x_0, the probability for finding it at point x_0 is 1. Therefore, $\langle x_0|x_0\rangle = 1$. Because of this, $\hat{I}$ is called the unit operator and plays the same role as the number 1 in the multiplication of numbers. The result

$$\int dx\, |x\rangle \langle x| = \hat{I} \tag{8.44}$$

or for a discrete basis

$$\sum_i |i\rangle \langle i| = \hat{I} \tag{8.45}$$

are called **completeness relations**. In general, we say the component of a ket $|a\rangle$ in the basis $|i\rangle$ is

$$|i\rangle^\dagger |a\rangle \equiv \langle i|a\rangle \equiv a_i,$$

which is a complex number. Using the completeness relation, i.e. $\sum_i |i\rangle \langle i| = \hat{I}$, we can write

$$|a\rangle = \sum_i |i\rangle \langle i|a\rangle = \sum_i |i\rangle\, a_i.$$

This can be seen as the series expansion[24] of the ket $|a\rangle$ in terms of the basis $|i\rangle$. Consequently, the complex numbers a_i can be seen as the expansion coefficients. Equally for a continuous complete basis, we have

$$|\Psi\rangle = \int dx\, |x\rangle \underbrace{\langle x|\Psi\rangle}_{\equiv \Psi(x) \text{ complex number}} = \int dx\, |x\rangle\, \Psi(x).$$

[24] Analogous to how we can write a vector in terms of a given basis: $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$, as explained in Appendix A.1.

The expectation value we introduced in Section 8.3.1 is in the Dirac notation given by

$$\langle\hat{O}\rangle = \langle\Psi|\, \hat{O}\, |\Psi\rangle\ .$$

Now let's return to the example of the particle in a box, and solve it using the Dirac notation.

8.5.4 Example: Particle in a Box, Again

The question we asked was: What is the probability for finding the particle having energy $E = E_2 = \frac{n^2\pi^2}{2mL^2}$? This question can be answered in the Dirac notation in a very natural way. The probability is given by

$$P(E_2) = \left|\, \langle E = \frac{n^2\pi^2}{2mL^2} |\Psi\rangle \,\right|^2 = \left|\, \langle E = \frac{n^2\pi^2}{2mL^2} | \underbrace{\left(\int dx\, |x\rangle \langle x| \right)}_{=\hat{I}} |\Psi\rangle \,\right|^2$$

$$= \Big| \int dx \, \langle E = \frac{n^2\pi^2}{2mL^2} |x\rangle \, \langle x|\Psi\rangle \Big|^2 = \Big| \int dx \Phi_2^\star(x)\Psi(x) \Big|^2 ,$$

which is exactly the result we derived using the standard wave mechanics in Section 8.5.2. All of this becomes much clearer as soon as you learn more about quantum mechanics and solve some problems on your own, using both notations.

8.5.5 Spin

Now it's time to return to the operator we derived in Section 5.1.1 for a new kind of angular momentum, we called spin. So far we have two loose ends. On the one hand, we have used spin as a label for the representations of the Lorentz group. For example, if we want to describe an elementary particle with spin $\frac{1}{2}$, we have to use an object transforming according to the spin $\frac{1}{2}$ representation. On the other hand, we derived from rotational symmetry something that we called spin, too[25]. In the quantum framework, when we act with the operator that we derived in Section 5.1.1 on a given state, the result is the spin of the particle. For the scalar representation, this operator is given by $\hat{S} = 0$, which of course always yields 0 when with it on a state Ψ.

For the $(\frac{1}{2}, 0)$ representation we have to use the two-dimensional representation of the rotation generator, which was derived in Section 3.7.5

$$\hat{S}_i = \frac{\sigma_i}{2} \tag{8.46}$$

and where σ_i denotes the Pauli matrices. If we want to know the spin of a particle described by Ψ, we have to act with the spin-operator on Ψ. For example, for $\hat{S}_3$ this will give us the spin of the corresponding particle in the 3-, or more familiar called z-direction. Analogously for $\hat{S}_2$ in the y- and $\hat{S}_1$ in the x-direction.

The explicit form of the operator $\hat{S}_3$ is

$$\hat{S_3} = \frac{\sigma_3}{2} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \tag{8.47}$$

The corresponding eigenstates are

$$v_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad v_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{8.48}$$

with eigenvalues $\frac{1}{2}$ and $-\frac{1}{2}$, respectively.

[25] The conserved quantity that follows from invariance under rotations has two parts. One part follows from the invariance under rotations of the spacetime coordinates and is called orbital angular momentum. The second parts follows from the invariance under the mixing of the field components. See Section 4.5.4.

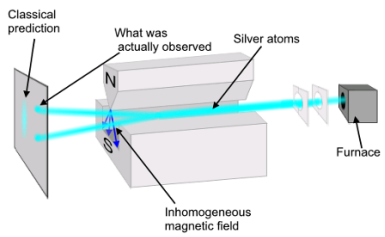

Fig. 8.3: Illustration of the Stern-Gerlach experiment. The original experiment was performed with silver atoms, whose spin behavior is dominated by the one electron in the outermost atomic orbital. The experimental result is the same as with electrons. A beam of particles is affected by an inhomogeneous magnetic field. For a classical type of angular momentum the deflection of the particles through the magnetic field should be a continuous distribution. Measured are just two deflection types, i.e. the beam splits in two parts, one corresponding to spin $\frac{1}{2}$ and one to $-\frac{1}{2}$. Figure by Theresa Knott (Wikimedia Commons) distributed under a CC BY-SA 3.0 license: `http://creativecommons.org/licenses/by-sa/3.0/deed.en`. URL: `http://commons.wikimedia.org/wiki/File:Stern-Gerlach_experiment.PNG`, Accessed: 24.5.2014.

This means a particle described by a spinor[26] has spin $\frac{1}{2}$, which can be aligned or anti-aligned to some arbitrary measurement axis. This is why we call this representation spin $\frac{1}{2}$ representation[27]. In the quantum framework this is interpreted by saying that a measurement of spin can only result in $\frac{1}{2}$ and $-\frac{1}{2}$. In Section 4.5.4 we learned that spin is something similar to orbital angular momentum, because both notions arise from rotational invariance. Here we can see that this kind of angular momentum gives quite surprising results, when measured. The most famous experiment proving this curious fact of nature is the Stern-Gerlach experiment (see Fig. 8.3).

[26] Recall that a spinor is an object transforming according to the $(\frac{1}{2},0)$, the $(0,\frac{1}{2})$ or $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ representation.

[27] Analogous statements can be made for the spin 1 or other higher representations.

The same is true for a measurement of spin in any direction. A measurement of the spin in the x-,y- or z-direction can only result in $\frac{1}{2}$ and $-\frac{1}{2}$.

Let's look at one concrete example of how a spin-measurement works in the quantum formalism. As mentioned above, the explicit form of the spin operator (Eq. 5.4), say for a measurement along the z-axis, is

$$\hat{S}_z = \frac{1}{2}\sigma_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}. \tag{8.49}$$

The eigenstates are $|\frac{1}{2}\rangle_z \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|-\frac{1}{2}\rangle_z \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, where the subscript z denotes that we are dealing with eigenstates of $\hat{S}_z$. A general spinor is not a spin-eigenstate, but a superposition $|X\rangle = a\,|\frac{1}{2}\rangle_z + b\,|-\frac{1}{2}\rangle_z$. The coefficients depend on how we prepare a given particle. If we did a measurement of the spin along the z-axis and filtered out all particles with spin $-\frac{1}{2}$, the coefficient b would be zero and a would be 1. Without any measurement and filtering the coefficients are $a = b = \frac{1}{\sqrt{2}}$, which means probability[28] $\frac{1}{2}$ for each possibility. Things get interesting if we make a measurement along the z-axis and afterwards a measurement, for example, along the x-axis. Even if we did filter out all $-\frac{1}{2}$ components along the z-axis, there will be particles with spin $-\frac{1}{2}$ along the x-axis.

[28] The coefficients are directly related to the probability amplitude which we need to square in order to get the probability. This will be shown in a moment.

Acting with the spin operator $\hat{S}_z$ on a state means measuring spin along the z-axis. For a general state $|X\rangle$ both outcomes $+\frac{1}{2}$ and $-\frac{1}{2}$ are possible and the probability is directly related to the factors a and b. If we want to know the probability for measuring $-\frac{1}{2}$, the quantum formalism tells us that the corresponding probability amplitude is

$$_z\langle -\frac{1}{2}|X\rangle = a\underbrace{{}_z\langle -\frac{1}{2}|\frac{1}{2}\rangle_z}_{=0} + b\underbrace{{}_z\langle \frac{1}{2}|-\frac{1}{2}\rangle_z}_{=1} = b \tag{8.50}$$

Therefore, the probability for measuring spin $-\frac{1}{2}$ along the z-axis is $P_{z=-\frac{1}{2}} = |b|^2$. If we want to measure the spin along another axis, say the x-axis, we need to expand our two states in terms of the eigenstates of $\hat{S}_x$, which reads in explicit matrix form (Eq. 5.4)

$$S_x = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \tag{8.51}$$

and the corresponding normalized eigenstates are $|\frac{1}{2}\rangle_x \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $|-\frac{1}{2}\rangle_x \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. If we want to know the probability for measuring spin $-\frac{1}{2}$ along the x-axis, we need to rewrite $|\frac{1}{2}\rangle_z$ and $|-\frac{1}{2}\rangle_z$ in terms of $|\frac{1}{2}\rangle_x$ and $|-\frac{1}{2}\rangle_x$:

$$\underbrace{|\frac{1}{2}\rangle_z}_{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} = \frac{1}{\sqrt{2}} \Big(\underbrace{|\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}} + \underbrace{|-\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}} \Big) \tag{8.52}$$

$$\underbrace{|-\frac{1}{2}\rangle_z}_{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} = \frac{1}{\sqrt{2}} \Big(\underbrace{|\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}} - \underbrace{|-\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}} \Big) . \tag{8.53}$$

And therefore

$$|X\rangle = a\,|\frac{1}{2}\rangle_z + b\,|-\frac{1}{2}\rangle_z = a\frac{1}{\sqrt{2}}\Big(|\frac{1}{2}\rangle_x + |-\frac{1}{2}\rangle_x \Big) + b\frac{1}{\sqrt{2}}\Big(|\frac{1}{2}\rangle_x - |-\frac{1}{2}\rangle_x \Big). \tag{8.54}$$

The probability amplitude for measuring $-\frac{1}{2}$ for the spin along the x-axis is then

$$\begin{aligned} {}_x\langle -\frac{1}{2}|X\rangle &= {}_x\langle -\frac{1}{2}| \left(a\frac{1}{\sqrt{2}}\Big(|\frac{1}{2}\rangle_x + |-\frac{1}{2}\rangle_x \Big) + b\frac{1}{\sqrt{2}}\Big(|\frac{1}{2}\rangle_x - |-\frac{1}{2}\rangle_x \Big) \right) \\ &= \frac{a}{\sqrt{2}} - \frac{b}{\sqrt{2}} \end{aligned} \tag{8.55}$$

and the probability is $P_{x=-\frac{1}{2}} = |\frac{a}{\sqrt{2}} - \frac{b}{\sqrt{2}}|^2$.

Now, let's come back to the example outlined at the beginning of this computation. If we filter out all particles with spin $-\frac{1}{2}$ along the z-axis the state $|X\rangle$ is

$$|X\rangle_{\text{after z-axis filtering}} = |\frac{1}{2}\rangle_z . \tag{8.56}$$

This means $a = 1$ and $b = 0$ and we get a non-zero probability for measuring spin $-\frac{1}{2}$ along the x-axis $P_{x=-\frac{1}{2}} = |\frac{1}{\sqrt{2}} - \frac{0}{\sqrt{2}}|^2 = \frac{1}{2}$. If we now filter out all particles with spin $-\frac{1}{2}$ along the x-axis and repeat our measurement of spin along the z-axis we notice something quite remarkable. After filtering the particles with spin $-\frac{1}{2}$ along the x-axis we have the state

$$|X\rangle_{\text{after x-axis filtering}} = |\frac{1}{2}\rangle_x \,. \tag{8.57}$$

If we want to know the probability for measuring spin $-\frac{1}{2}$ along the z-axis, we need to write $|\frac{1}{2}\rangle_x$ in terms of $|\frac{1}{2}\rangle_z$ and $|-\frac{1}{2}\rangle_z$:

$$|X\rangle_{\text{after x-axis filtering}} = \underbrace{|\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}} = \frac{1}{\sqrt{2}}\Big(\underbrace{|\frac{1}{2}\rangle_z}_{\begin{pmatrix}1\\0\end{pmatrix}} + \underbrace{|-\frac{1}{2}\rangle_z}_{\begin{pmatrix}0\\1\end{pmatrix}}\Big) \tag{8.58}$$

and we get the probability $P_{z=-\frac{1}{2}} = |\,_z\langle -\frac{1}{2}|X\rangle|^2 = \frac{1}{2}$.

To summarize, this means:

- We start with a spin measurement along the z-axis and filter out all particles with $-\frac{1}{2}$. This leaves us with a state

$$|X\rangle_{\text{after z-axis filtering}} = |\frac{1}{2}\rangle_z \,. \tag{8.59}$$

- If we now measure again the spin along the z-axis we get a very unsurprising result: The probability for measuring spin $-\frac{1}{2}$ is zero and for spin $+\frac{1}{2}$ the probability is 100%.

$$\,_z\langle \frac{1}{2}|X\rangle_{\text{after z-axis filtering}} = 1\,, \tag{8.60}$$

$$\,_z\langle -\frac{1}{2}|X\rangle_{\text{after z-axis filtering}} = 0\,. \tag{8.61}$$

- If we then measure the spin along the x-axis, for our z-filtered particle stream, we get a probability of $\frac{1}{2} = 50\%$ for a measurement of $+\frac{1}{2}$. Equivalently, we have a probability of $\frac{1}{2} = 50\%$ for a measurement of $-\frac{1}{2}$. If we then filter out all components with spin $-\frac{1}{2}$ along the x-axis we are in the state $|X\rangle_{\text{after x-axis filtering}} = |\frac{1}{2}\rangle_x$.
- Now measuring the spin along the z-axis again, gives us the surprising result that the probability for measuring spin $-\frac{1}{2}$ is $\frac{1}{2} = 50\%$. The measurement along the x-axis did change the state and therefore we are again getting components with spin $-\frac{1}{2}$ along the z-axis, even though we did filter these out in the first step!

A brilliant discussion of these matters, involving real measuring devices, can be found in the Feynman Lectures[29] Vol. 3.

[29] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics, Volume 3*. Addison Wesley, 1st edition, 1 1971. ISBN 9780201021189

8.6 Heisenberg's Uncertainty Principle

Now it's time to talk about one of the most curious features of quantum mechanics. We learned in the last section that a measurement of spin in the x-direction makes everything we knew previously about spin along the z-direction useless. This kind of thing happens for many observables in quantum mechanics. We can trace this behavior back to the fact that[30] $\hat{S}_x\hat{S}_z \neq \hat{S}_z\hat{S}_x$. This means that a measurement of spin along the z-axis followed by a measurement of spin along the x-axis is different from a measurement of spin along the x-axis followed by a measurement of spin along the z-axis. After measuring the spin along the z-axis the system is in an eigenstate of $\hat{S}_z$ and after a measurement of spin along the x-axis, in an eigenstate of $\hat{S}_x$. The eigenstates for $\hat{S}_z$ and $\hat{S}_x$ are all different and therefore this is no surprise.

[30] Recall that we identify the spin operators with the corresponding finite-dimensional representations for the rotation generators. These fulfil the commutator relation $[J_i, J_j] = J_iJ_j - J_jJ_i = i\epsilon_{ijk}J_k \neq 0 \rightarrow J_iJ_j \neq J_jJ_i$. For example, if we describe spin $\frac{1}{2}$ particles, we must use the two-dimensional representation $J_i = \frac{\sigma_i}{2}$.

We can look at this from a different perspective: **We aren't able to know the spin of a system along the z-axis and the x-axis at the same time!** Each time we measure spin along the z-axis the spin along the x-axis becomes undetermined and vice-versa. The same is true for spin along the z-axis/x-axis and spin along the y-axis. Spin may be something really strange, but we can observe the same behavior for measurements of position and momentum. Take a look again at Eq. 5.3, which we recite here for convenience:

$$[\hat{p}_i, \hat{x}_j] = \hat{p}_i\hat{x}_j - \hat{x}_j\hat{p}_i = i\delta_{ij}. \tag{8.62}$$

Following the line of thought as above tells us that a measurement of momentum in the x-direction changes what we can expect for a measurement of position on the x-axis. In other words this means that we can't know position and momentum in the same direction at the same time with arbitrary precision. Take note that the commutator is only non-zero for measurements along the same axis[31]. A measurement of momentum in the y-direction has no influence on what we can expect for the position on the x-axis.

[31] The Kronecker delta δ_{ij} is zero for $i \neq j$ and one for $i = j$ as defined in Appendix B.5.5.

Everytime we measure momentum the position becomes uncertain and vice versa. This is known as **Heisenberg's uncertainty principle**. Analogous observations can be made for angular momentum along different axes, because the commutator for the corresponding operator is non-zero, too. In general, we can check for any two physical quantities if they commute with each other. If they don't, we know that they can't be measured at the same time with arbitrary precision.

Maybe this shouldn't surprise us. Quantum mechanics uses the generators of the corresponding symmetries as measurement oper-

ators. For instance, this has the consequence that a measurement of momentum is equivalent to the action of the translation generator[32]. The translation generator moves our system a little bit and therefore the location is changed. What is more surprising is that nature actually works this way. Over the years there have been many experimental tests of the Heisenberg's uncertainty principle and all proved it to be correct.

[32] Recall that invariance under translations leads us to conservation of momentum.

8.7 Comments on Interpretations

The interpretation and notations described in this chapter are the standard ones. Nevertheless, there are other formalisms equally powerful. For example, the Feynman path integral formalism[33] is, in terms of results, completely equivalent to the wave mechanics, we described in this chapter. Yet computations in this formalism are completely different. If we want to compute the probability for a particle to get from point a to point b, we have to sum over all possible paths that a particle can go between a and b. As absurd as it sounds, this approach leads to the same result, which can be proved formally as well. Freeman Dyson once told the story[34]

> Dick Feynman told me about his "sum over histories" version of quantum mechanics. "The electron does anything it likes," he said. "It just goes in any direction at any speed, forward or backward in time, however it likes, and then you add up the amplitudes and it gives you the wavefunction." I said to him, "You're crazy." But he isn't.

[33] To learn more about this see for example, Richard P. Feynman and Albert R. Hibbs. *Quantum Mechanics and Path Integrals: Emended Edition*. Dover Publications, emended editon edition, 7 2010. ISBN 9780486477220

[34] Harry Woolf, editor. *Some Strangeness in the Proportion*. Addison-Wesley, 1st edition, 2 1981. ISBN 9780201099249

Another interpretation for the basic equations of quantum mechanics, further away from the mainstream, is Bohmian Mechanics. The starting point is to put the ansatz: Re^{St} into the Schrödinger equation. Separating the imaginary and real part results in two equations, one of which can be seen as completely analogous to the Hamilton-Jacobi equation of classical mechanics plus an additional term. This additional term can be interpreted as an extra potential, which is responsible for the strange quantum effects. Further computations are completely analogous to classical mechanics. A new force is computed from the extra potential using the gradient, which is then put into Newton's classical equation: $F = ma$. Therefore, in Bohmian mechanics one still has classical particle trajectories. The results from this approach are, as far as I know, equal to those computed by standard non-relativistic quantum mechanics. Nevertheless, this approach has fallen into disfavour because the extra potential undergoes non-local changes.

Further Reading Tips

- **Richard P. Feynman - The Feynman Lectures on Physics, Vol. 3**[35] is a great book to start learning about quantum mechanics. Most concepts of quantum mechanics are explained here more lucidly than anywhere else.
- **David J. Griffiths - Introduction to Quantum Mechanics**[36] is a very readable and enlightening book.
- **J. J. Sakurai - Modern Quantum Mechanics**[37] is a brilliant book, which often offers a unique perspective on the concepts of quantum mechanics.
- **Paul A. M. Dirac - Lectures on Quantum Mechanics**[38] is an old book by one of the fathers of quantum mechanics. Highly recommended, because reading about ideas by the person who discovered them is always a good idea.
- You can find a detailed discussion of the interpretation of the components of a Dirac spinor in Appendix 8.8.

[35] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics, Volume 3*. Addison Wesley, 1st edition, 1 1971. ISBN 9780201021189

[36] David J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Prentice Hall, 2nd edition, 4 2004. ISBN 9780131118928

[37] J. J. Sakurai. *Modern Quantum Mechanics*. Addison Wesley, 1st edition, 9 1993. ISBN 9780201539295

[38] Paul A. M. Dirac and Physics. *Lectures on Quantum Mechanics*. Dover Publications, 1st edition, 3 2001. ISBN 9780486417134

8.8 Appendix: Interpretation of the Dirac Spinor Components

In this and the corresponding appendices, u *and* v *denote two-component objects inside a Dirac spinor and* u *and* v *four-component objects. For example* u_1 *and* u_2 *describe two different four-component objects. For a general four-component object* u*, we denote the two two-component objects by* u_1 *and* u_2*, i.e.* $u = \begin{pmatrix} \mathrm{u}_1 \\ \mathrm{u}_2 \end{pmatrix}$.

Up to this point we have been relatively vague about the two Weyl spinors inside a Dirac spinor. What do they stand for? How can they be interpreted? In addition, each such Weyl spinor inside a Dirac spinor consists of two components. How to interpret these? Now, we are finally in the position to give answers to these questions.

In short:

- The two Weyl spinors χ_L, ξ_R inside a Dirac spinor $\psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}$ describe "different particles". Nevertheless, it's conventional to call them the *same* particle, for example an electron, with different chirality:
  - χ_L describes a left-chiral electron,
  - ξ_R describes a right-chiral electron,

but the crucial point is that these are really distinct particles/fields[39], because they aren't related by a parity transformation or charge conjugation. This is why we use different symbols. There is, of course, some sort of connection between them, which is why we write them in one object. This will be discussed in detail in a moment.

[39] They are labelled by different quantum numbers and therefore behave differently in experiments!

- The two components of each Weyl spinor describe different spin configurations[40] of the corresponding particle.
  - A Weyl spinor proportional to $\begin{pmatrix}1\\0\end{pmatrix}$ describes a particle with **spin up**
  - A Weyl spinor proportional to $\begin{pmatrix}0\\1\end{pmatrix}$ describes a particle with **spin down**
  - Any other Weyl spinor is simply a mixture of spin up and spin down.

[40] This was discussed in Section 8.5.5. Recall that spin can be measured like angular momentum, but for a spin $\frac{1}{2}$ particle the result of such a measurement can only be $+\frac{1}{2}$ or $-\frac{1}{2}$, no matter what axis we choose. These two measurement results are commonly called **spin up** and **spin down**.

Let's see how this comes about in detail:

The important thing we learn from weak interactions and parity violation is that left-chiral and right-chiral particles are really different particles. Left-chiral particles carry weak charge (isospin) and therefore interact via the weak force. Right chiral particles do not, which was explained in Section 7.7.1.

We have for every particle in nature a corresponding antiparticle and in general, we get the antiparticle description through charge conjugation. Charge conjugation flips all particle labels, which includes isospin. Let's see what particles we can expect that are related to, say electrons. We have

- A left-chiral *electron* χ_L, with isospin $-\frac{1}{2}$ and electric charge $-e$, which is part of a doublet.
- A right-chiral *anti-left-chiral-electron* $(\chi_L)^c = \chi_R$ with isospin $\frac{1}{2}$, electric charge $+e$, which is part of a doublet, too. This property does not simply vanish through charge conjugation. This may be confusing at the moment, but so far we have only talked about the coupling of the weak force to particles. It turns out that the weak force couples to right-chiral antiparticles as well.
- A right-chiral *electron* ξ_R with isospin 0 and electric charge $-e$
- A left-chiral *anti-right-chiral-electron* $(\xi_R)^c = \xi_L$ with isospin 0 and electric charge $+e$

Therefore, when talking about electrons, we have in fact **four** different "things" we need to consider. These are all really different particles and we need to give them different names! Usually one talks about just two particles related to an electron: The electron and the positron and we will see how this comes about in a moment.

We restrict the following discussion to the rest frame of the particles in question. In other frames the discussion works analogous but is more cumbersome.

The objects inside a Dirac spinor and its charge conjugate are directly related to these four particles. The **physical electron** and the **physical positron** are commonly identified as the solutions of the Dirac equation. The Dirac equation is an equation of motion, i.e. an equation that determines the dynamics of the particles in question. The explicit solutions are Dirac spinors that evolve in time. We need such Dirac spinors with a definite time-evolution in order to describe how our particles evolve in time. As we will see this requires that we always use two of the particles listed above at once.

An explicit derivation of the solutions of the Dirac equation can be found in the appendix Section 8.9. Here we just use the results. There are four independent solutions of the Dirac equation and two are of the form

$$\psi_i = \begin{pmatrix} u_i \\ u_i \end{pmatrix} \tag{8.63}$$

with for example[41] $u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-imt}$ and $u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-imt}$ and two solutions are of the form

$$\tilde{\psi}_i = \begin{pmatrix} -v_i \\ v_i \end{pmatrix}, \tag{8.64}$$

with for example $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{+imt}$ and $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{+imt}$

[41] This is a basis choice. The only requirement is that they are linearly independent.

The Dirac equation tells us that the spin configuration of the two particles described by the upper and lower Weyl spinors inside a Dirac spinor, are directly related. In addition, their time dependence must be the same. These solutions describe what is commonly known as a **physical electron** and a **physical positron**, with different spin configurations[42].

- ψ_1 is an electron with spin up
- ψ_2 is an electron with spin down

[42] Take note that the connection between these objects is charge conjugation. This can be seen by using the explicit form of the charge conjugation operator: $\psi_1^c = i\gamma_2\psi_1^\star$. Therefore: $(\psi_1)^c = \tilde{\psi}_2$ and $(\psi_2)^c = \tilde{\psi}_1$.

- $\tilde{\psi}_1$ is a positron with spin up
- $\tilde{\psi}_2$ is a positron with spin down

We can see nicely that a physical electron has a left-chiral (the upper two components) and a right-chiral part (the lower two components). For a physical electron the spin configuration of the left-chiral and the right-chiral part and the time-dependence must be the same for both parts. As discussed above, this left-chiral and right-chiral parts are really different, because they have different weak charge! Nevertheless, in order to describe a dynamical, physical electron we always need both parts.

Take note that these solutions do not mean that the object we use to describe a physical electron consists of exactly the same upper and lower two-component objects. Only their spin configuration and their time dependence must be equivalent. The upper object is still part of a doublet, whereas the lower object isn't. The upper object transforms under $SU(2)$ transformations and the lower doesn't. Using the notation from above we have

$$\text{physical electron} = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \propto \begin{pmatrix} \mathrm{u} \\ \mathrm{u} \end{pmatrix}. \tag{8.65}$$

This does not mean that $\chi_L = \xi_R$. The Weyl spinor χ_L is part of a doublet and describes a particle with isospin, whereas the particle described by ξ_R has isospin zero. In addition, we already know that the upper and lower Weyl spinor inside a Dirac spinor transform differently under Lorentz boosts. Therefore it is important that we use different symbols. In other words: The object describing a left-chiral electron χ_L carries an additional $SU(2)$ index, because χ_L transforms as part of a doublet under $SU(2)$ transformations. ξ_R has no such index and transforms as a singlet under $SU(2)$ transformations.

Equivalently we have

$$\text{physical positron} = \begin{pmatrix} -\xi_L \\ \chi_R \end{pmatrix} \propto \begin{pmatrix} -\mathrm{v} \\ \mathrm{v} \end{pmatrix} \tag{8.66}$$

which we can see through charge conjugation[43]

$$\begin{aligned} (\text{physical electron})^c &= i\gamma_2 \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}^\star = \begin{pmatrix} -\xi_L \\ \chi_R \end{pmatrix} \propto i\gamma_2 \begin{pmatrix} \mathrm{u} \\ \mathrm{u} \end{pmatrix}^\star = \begin{pmatrix} -\mathrm{u}^c \\ \mathrm{u}^c \end{pmatrix} \\ &= \text{physical positron} \end{aligned} \tag{8.67}$$

The message to take away is that the physical electron we observe in nature most of the time is a mixture of two different particles: The

[43] See Section 7.1.5 and use the explicit form of the matrix as defined in Eq. 6.13: $\gamma_2 = \begin{pmatrix} 0 & \bar{\sigma}_2 \\ \sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\sigma_2 \\ \sigma_2 & 0 \end{pmatrix}$.

left-chiral electron that carries isospin and the right-chiral electron with isospin zero! Equivalently the physical positron is a mixture of an anti-left-chiral electron, which carries isospin and an anti-right-chiral electron with isospin zero.

The solutions of the Dirac equation tell us how our particles evolve in time. Let's say we start with an electron with spin up that was created in a weak interaction and is therefore purely left-chiral. How does this particle evolve in time? A purely left-chiral electron with spin up

$$e_L^{\uparrow} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{8.68}$$

is not a solution of the Dirac equation and therefore, in order to determine its time evolution, we must rewrite this in terms of solutions of the Dirac equation.

$$e_L^{\uparrow} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}\right) = \Psi_1(t=0) - \tilde{\Psi}_1(t=0)\,. \tag{8.69}$$

We know how Ψ_1 and $\tilde{\Psi}_1$ evolve in time

$$\Psi_1(t) - \tilde{\Psi}_1(t) = \rightarrow \frac{1}{2}\left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{-imt} - \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{imt}\right) \tag{8.70}$$

For $t = 0$ this reduces to the left-chiral state as it should be, but as time evolves, say $t = \frac{\pi}{2m}$ we have

$$\begin{aligned} \rightarrow \frac{1}{2}\left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \underbrace{\mathrm{e}^{-i\frac{\pi}{2}}}_{=-\mathrm{i}} - \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \underbrace{\mathrm{e}^{i\frac{\pi}{2}}}_{=\mathrm{i}}\right) &= \frac{i}{2}\left(\begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}\right) \\ &= -i\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = -ie_R^{\uparrow} \end{aligned} \tag{8.71}$$

which describes a right-chiral electron with spin up! The lesson here is that as time evolves, a left-chiral particle changes into a right-chiral

particle and vice versa. To describe the time-evolution of a particle like an electron we need e_L and e_R, which is why we wrote them together in one object: the Dirac spinor. The same is true for the positron.

Recall that the two different particles e_L and e_R, carry different weak charge, i.e. isospin. Nevertheless as time evolves these two particles can transform into each other. Most of the time it will be a mixture of both and not a definite eigenstate. Isospin and chirality are therefore not conserved as time evolves, only in interactions.

We can now see that the notation with Dirac spinors is necessary, because we have a close, dynamical connection between each two particles of the four particles listed at the beginning of this section. Chirality and therefore isospin are not conserved during propagation. A propagating electron can sometimes be found as left-chiral and sometimes as right-chiral.

8.9 Appendix: Solving the Dirac Equation

As explained at the beginning of the last section, we use here the symbols u *and* v *for the two-component objects inside a Dirac spinor, and* u *and* v *for four-component objects. This means, for example* u_1 *and* u_2 *describe two different four-component objects. If we don't want to be specific and want to consider both four-component objects at the same time we simply write* u*. Then* u_1 *and* u_2 *are the two two-component objects inside such a four-component object* $u = \begin{pmatrix} \mathrm{u}_1 \\ \mathrm{u}_2 \end{pmatrix}$

In this appendix we will solve the Dirac equation in the *rest frame* in the *chiral basis*. The solution for an arbitrary frame can be computed by acting with a boost transformation on the solution derived in this section. In addition to the discussion in the last section, we will use these solutions in Chapter 9, when we talk about quantum field theory. The Dirac equation is

$$(i\partial_\mu \gamma^\mu - m)\psi = 0. \tag{8.72}$$

Anticipating plane wave solutions, we make the ansatz $\Psi = u\mathrm{e}^{-ipx}$, with some four-component object u, because the matrices γ_μ in the equation are 4×4. In the rest frame, which means momentum zero $\vec{p} = 0$, the exponent reduces to $-ipx = -i(p_0 x_0 - \vec{p}\vec{x}) = -ip_0 x_0$. Now using the relativistic energy-momentum relation $E = \sqrt{\vec{p} + m^2}$, which we derived at the beginning of this chapter, and using that

$p_0 = E$ and $x_0 = t$, we have $-ipx = -iEt = -i\sqrt{\underbrace{\vec{p}}_{=0} + m^2}t = -imt$. Putting this ansatz into the Dirac equation yields

$$\begin{aligned}
&(i\partial_\mu\gamma^\mu - m)u\mathrm{e}^{-imt} = 0 \\
&\rightarrow (i(\partial_0\gamma^0 + \partial_i\gamma^i) - m)u\mathrm{e}^{-imt} = 0 \\
&\rightarrow i((-im)\gamma^0 - m)u\mathrm{e}^{-imt} = 0 \\
&\rightarrow (m\gamma^0 - m)u = 0 \\
&\underbrace{\rightarrow}_{\text{dividing by m}} \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)c = 0 \\
&\rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 0 \\
&\rightarrow \begin{pmatrix} -u_1 + u_2 \\ u_1 - u_2 \end{pmatrix} = 0
\end{aligned} \tag{8.73}$$

The 1 inside the remaining matrix here is the 2×2 unit matrix and therefore u_1 and u_2 are two-component objects. We see that our ansatz solves the equation, if $u_1 = u_2$. Therefore, we have found two linearly independent solutions of the Dirac equation

$$\Psi_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}\mathrm{e}^{-imt} \quad \Psi_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}\mathrm{e}^{-imt} \tag{8.74}$$

We can find two other solutions by making the ansatz $\tilde{\Psi} = v\mathrm{e}^{ipx}$, which analogously reduces in the rest frame to $\tilde{\Psi} = v\mathrm{e}^{imt}$. This ansatz yields

$$\begin{aligned}
&(i\partial_\mu\gamma^\mu - m)v\mathrm{e}^{imt} = 0 \\
&\rightarrow (-m\gamma^0 - m)v = 0 \\
&\rightarrow \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0 \\
&\rightarrow \begin{pmatrix} -v_1 - v_2 \\ -v_1 - v_2 \end{pmatrix} = 0.
\end{aligned} \tag{8.75}$$

We therefore conclude that we have a solution with time dependence e^{imt}, if the upper and lower two-component objects in the Dirac spinor are related by $-v_1 = v_2$. Two linearly independent

solutions following from this ansatz are

$$\tilde{\Psi}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} e^{imt} \quad \tilde{\Psi}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} e^{imt}. \tag{8.76}$$

8.10 Appendix: Dirac Spinors in Different Bases

In the Lagrangian the Dirac spinors ψ appear always in combination with the matrices γ_μ. This can be used to simplify computations, by switching to a different basis. This works, because we can add terms of the form $1 = N^{-1}N$, with some arbitrary invertible matrix N, between ψ and γ_μ and then redefine both. For example

$$\partial_\mu \bar{\psi}\gamma_\mu\psi = \partial_\mu \bar{\psi} \underbrace{N^{-1}N}_{=1} \gamma_\mu \underbrace{N^{-1}N}_{=1} \psi = \partial_\mu \underbrace{\bar{\psi}N^{-1}}_{\equiv\bar{\psi}'} \underbrace{N\gamma_\mu N^{-1}}_{\equiv\gamma'_\mu} \underbrace{N\psi}_{\equiv\psi'} = \partial_\mu \bar{\psi}'\gamma'_\mu\psi'. \tag{8.77}$$

The basis we worked with in this text so far is called the **chiral basis** or **Weyl basis**. Conventionally the Dirac equation is solved in another basis, called **mass basis** or **Dirac basis**. In the chiral basis we worked with so far, the Dirac Lagrangian

$$\mathcal{L}_D = i\chi_L^\dagger \sigma^\mu \partial_\mu \chi_L + i\xi_R^\dagger \bar{\sigma}^\mu \partial_\mu \xi_R - m\chi_L^\dagger \xi_R - m\xi_R^\dagger \chi_L \tag{8.78}$$

has non-diagonal mass terms, i.e. mass terms that mix different states. We can use the freedom to choose a basis to pick a basis where the mass terms are **diagonal**, which is then called mass basis.

This means we want a mass term $\psi^\dagger m\psi$, with $m = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}$, which gives us mass terms of the form

$$\bar{\psi}'M'\psi' = \psi^\dagger\gamma'_0 M'\psi' = \begin{pmatrix} u' \\ v' \end{pmatrix}^\dagger \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} u' \\ v' \end{pmatrix} = (u')^\dagger m_1 u' + (v')^\dagger m_2 v', \tag{8.79}$$

whereas at the moment we are dealing with

$$\bar{\psi}M\psi = \psi^\dagger\gamma_0 M\psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}^\dagger \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = m\chi_L^\dagger \xi_R + m\xi_R^\dagger \chi_L. \tag{8.80}$$

The latter basis, which we worked with so far, makes it easy to interpret things in terms of chirality, whereas it's easier for Dirac spinors in the mass basis to make the connection to **physical** propagating **particles**.

To find the connection between the second and the first form, we need to diagonalize the matrix $M = \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} = m \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. The matrix is diagonalized through the matrix $N = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$:

$$N^{-1} \underbrace{\begin{pmatrix} -m & 0 \\ 0 & m \end{pmatrix}}_{\equiv M'} N = M \tag{8.81}$$

$$\rightarrow m \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} = m \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{8.82}$$

and therefore we redefine the Dirac spinors accordingly

$$\begin{aligned} \bar{\psi} M \psi &= \bar{\psi} \underbrace{N N^{-1}}_{=1} M \underbrace{N N^{-1}}_{=1} \psi \\ &= \underbrace{\bar{\psi} N}_{\equiv \bar{\psi}'} \underbrace{N^{-1} M N}_{\equiv M'} \underbrace{N^{-1} \psi}_{\equiv \psi'} \\ &= \bar{\psi}' M' \psi'. \end{aligned} \tag{8.83}$$

It is instructive to have a look at the chiral projection operators $P_L = \frac{1-\gamma_5}{2}$ in this basis. We need to calculate how the matrix γ_5 looks like in this new basis

$$\begin{aligned} \tilde{\gamma_5} = N^{-1} \gamma_5 N &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \end{aligned} \tag{8.84}$$

The corresponding eigenvectors are $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. This means a chiral eigenstate is now described by a Dirac spinor with upper and lower components. For example, a left-chiral state is in this basis of the form $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. In contrast, in the chiral basis γ_5

was diagonal and a left-chiral eigenstate was given by a Dirac spinor with upper components only $\psi_L = \begin{pmatrix} \chi_L \\ 0 \end{pmatrix}$, and a right-chiral Dirac spinor with lower components only $\psi_R = \begin{pmatrix} 0 \\ \xi_R \end{pmatrix}$.

The chiral projection operator is in this basis

$$P_L = \frac{1-\gamma_5}{2} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{8.85}$$

8.10.1 Solutions of the Dirac Equation in the Mass Basis

We can solve the Dirac equation in the mass basis

$$\big(i\gamma_\mu \partial^\mu - m\big)\Psi = 0 \tag{8.86}$$

by making the ansatz $\psi = u\mathrm{e}^{-ipx}$, which yields

$$\begin{aligned} &\big(\gamma_\mu p^\mu - m\big)u\mathrm{e}^{-ipx} = 0 \\ &\rightarrow \big((\gamma_\mu p^\mu - m\big)u = 0\,. \end{aligned} \tag{8.87}$$

Equivalently we can make the ansatz $\psi = v\mathrm{e}^{ipx}$, which yields

$$\begin{aligned} &\big(-\gamma_\mu p^\mu - m\big)v\mathrm{e}^{ipx} = 0 \\ &\rightarrow \big(-\gamma_\mu p^\mu - m\big)v = 0. \end{aligned}$$

Analogously to our solution in the chiral basis, we work in the rest frame, i.e. $\vec{p} = 0$. We are allowed to make such a choice, because physics is the same in all frames of reference and therefore we can pick one that fits our needs best. In this frame of reference, with $p_i = 0$, we have

$$\begin{aligned} &\rightarrow \big(\gamma_0 p^0 - m\big)u = 0 \\ &\rightarrow \big(-\gamma_0 p^0 - m\big)v = 0. \end{aligned}$$

In addition, we have $p_0 = E$ and we can use the relativistic energy-momentum relation, which we derived at the beginning of this chapter (Eq. 8.2). In the rest frame we have $E = \sqrt{(p_i)^2 + m^2} = m$. We now use the explicit form of γ_0 in the mass basis, which can be computed using the matrix N from above and the transformation law $\gamma_0' = N^{-1}\gamma_0 N$. Remember that we have an implicit unit matrix behind m and therefore

$$\rightarrow \Big(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} m - m \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\Big)u = 0$$

$$\rightarrow \Big(- \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} m - m \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Big) v = 0$$

$$\rightarrow \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} u = 0$$

$$\rightarrow \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} v = 0.$$

Recalling that each Dirac spinor consists of two two-component objects, we conclude that the lower two-component object of u and the upper two-component object of v must be zero:

$$\rightarrow \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2u_2 \end{pmatrix} = 0 \rightarrow u_2 = 0$$

$$\rightarrow \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -2v_1 \\ 0 \end{pmatrix} = 0 \rightarrow v_1 = 0.$$

We can see that in this basis the physical propagating particles (=solutions of the Dirac equation) are described by spinors with upper components only or equivalently for antiparticles with lower components only. We therefore have again four linearly independent solutions

$$\Psi_1' = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} e^{-imt} \quad \Psi_2' = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} e^{-imt} \tag{8.88}$$

and

$$\tilde{\Psi}_1' = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{imt} \quad \tilde{\Psi}_2' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{imt}. \tag{8.89}$$

A general solution in this frame, in this basis, is a linear combination

$$\psi = u e^{-ipx} + v e^{ipx} = \begin{pmatrix} u_1 \\ 0 \end{pmatrix} e^{-ipx} + \begin{pmatrix} 0 \\ v_1 \end{pmatrix} e^{ipx} \tag{8.90}$$

and we get the solution in an arbitrary frame by transforming this solution with a Lorentz boost. In addition, the most general solution is a superposition of all possible momenta and spin configurations[44]

$$\Psi = \sum_r \sqrt{\frac{m}{(2\pi)^3}} \int \frac{d^3p}{\sqrt{E_p}} \left(c_r(p) u_r(p) e^{-ipx} + d_r^\dagger(p) v_r(p) e^{+ipx} \right) \tag{8.91}$$

[44] Recall that the two components of a Weyl spinor represent different spin states.