Interaction to gauge theories

Panos Christakoglou[1,2]

$$\left(i\gamma^{\mu}\partial_{\mu}-im\right)\psi(x)=0.$$

[1] P.Christakoglou@uu.nl - Utrecht University, Leonard S. Ornsteinlaboratorium Princetonplein 1, 3584CC Utrecht, The Netherlands
[2] Panos.Christakoglou@nikhef.nl - Nikhef, Science Park 105, 1098XG Amsterdam, The Netherlands

Contents

Chapter 1
Introduction to gauge theories

In this chapter we will lay the mathematical foundations that will be used to develop the theories the Standard Model is based on, the ones that describe the electromagnetic, the strong and the weak interactions. These foundations rely in introducing the gauge theories that describe all interactions between elementary particles. We will start with a brief reminder of an essential topic, making the transition from classical to the quantum-mechanical representation thought the Schrodinger equation. We will then introduce the Klein-Gordon equation that describes spin-0 real and complex scalar fields. We then move to the Dirac equation that describes a system formed by spin-$\frac{1}{2}$ spinor fields. Finally, we will introduce the Proca equation that describes spin-1 vector fields.

1.1 Schrodinger equation

In quantum-mechanics free particles are described as waves that can be decomposed into a Fourier integral of plane waves according to

$$\Psi(\mathbf{x},t) \approx e^{i(\mathbf{kx}-\omega t)} = \cos(\mathbf{kx}-\omega t) + i\sin(\mathbf{kx}-\omega t),$$

where $\mathbf{k} = \mathbf{p}/\hbar$ the wave vector connected with the wavelength λ via $\lambda = h/p$. The angular frequency gives the energy of the wave according to $E = \hbar\omega$. The particle in quantum mechanics is not localised and the probability to find a particle described by the wave function $\Psi(\mathbf{x},t)$ in a volume V is given by

$$P(\mathbf{x},t)dV = |\Psi(\mathbf{x},t)|^2 dV$$

The normalisation condition gives rise to

$$P = \int P(\mathbf{x},t)dV = \int_{\text{all}-\text{space}} |\Psi(\mathbf{x},t)|^2 dV = 1$$

All physical quantities and time-dependent variables are described by operators according to

$$\hat{A}\Psi = \alpha\Psi,$$

where $\hat{A}$ is the operator, α the eigenvalue that corresponds to the operator and Ψ is the eigenfunction. For $\hat{A}$ to correspond to a physical observable the eigenvalues of the operator must be real. This means that the operator needs to be Hermitian. Table 1.1 gives some indicative examples for some of the basic "variables" in their classical representation with their quantum-mechanical counterparts.

In quantum mechanics we thus have in one dimension:

	Classical representation	QM representation
Scalar function	$f(x)$	$f(x)$
Momentum	$\mathbf{p}$	$-i\hbar\nabla$
Energy	E	$i\hbar\frac{\partial}{\partial t}$
Angular momentum (z-coordinate)	L_z	$-i\hbar\frac{\partial}{\partial\phi}$

Table 1.1: Association of some of basic "variables" in their classical representation with their quantum-mechanical counterparts.

$$\hat{P}_x\Psi(x,t) = -i\hbar\frac{\partial}{\partial x}(Ne^{i(kx-\omega t)}) = \hbar kNe^{i(kx-\omega t)} = p_x\Psi(x,t)$$

$$\hat{E}\Psi(x,t) = i\hbar\frac{\partial}{\partial t}(Ne^{i(kx-\omega t)}) = \hbar\omega Ne^{i(kx-\omega t)} = E\Psi(x,t)$$

In classical mechanics the totla energy of a non-relativistic particle is the sum of the kinetic (T) and the potential (U) energy, given by

$$E = H = T + U = \frac{p^2}{2m} + U$$

If one writes down the quantum-mechanical counterparts of the previous equation then we end up with

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) + \hat{U}\Psi(x,t)$$

The previous equation describes a particle moving in one dimension, along the x-axis. The generalisation in three dimensions gives rise to the time-dependent Schrodinger equation:

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t) = -\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{x},t) + \hat{U}\Psi(\mathbf{x},t) \tag{1.1.1}$$

The physical interpretation of the wave function is that the product $\Psi^*\Psi$ calculated over a volume element d^3x gives the probability of finding a particle in the volume element. This gives rise to the probability density expressed by

$$\rho(\mathbf{x},t) = \Psi^*(\mathbf{x},t)\Psi(\mathbf{x},t) \tag{1.1.2}$$

Assuming that a particle does not decay or interact, the total probability is constant. This can be reflected by a continuity equation, defining the probability curent density $\mathbf{j}(\mathbf{x},t)$, so that the flux of probability across an element surface $d\mathbf{S}$ is $\mathbf{j}(\mathbf{x},t)d\mathbf{S}$.

Taking the time derivative of the probability density over the entire volume reflects the amount of probability current density lines exiting the element surface $d\mathbf{S}$ (hence the minus sign below):

$$\frac{\partial}{\partial t}\int_V \rho dV = -\int_S \mathbf{j}d\mathbf{S} = -\int_V \nabla\mathbf{j}dV$$

The last part of the previous equation is given by the divergence theorem according to which

$$\int_V (\nabla F)dV = \int_S \mathbf{F}d\mathbf{S}$$

This is schematically given in fig. 1.1

Since the previous relation holds for an arbitrary volume, the continuity equation is written as

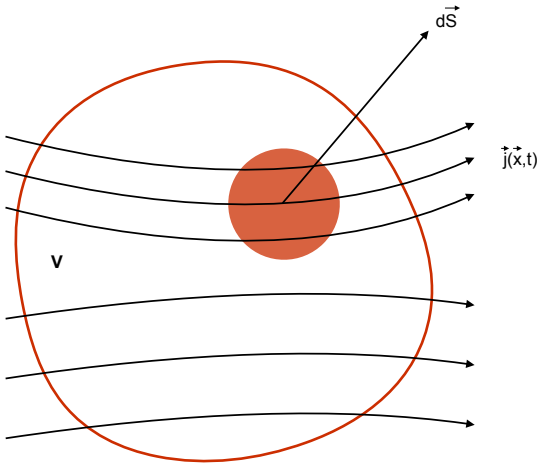

Fig. 1.1: The probability current density lines exiting from a surface $d\mathbf{S}$ of a volume V.

$$\nabla \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \tag{1.1.3}$$

The expression for the probability current can be extracted from the free particle time-dependent Schrodinger equation for Ψ and Ψ^*:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi$$

$$-i\hbar \frac{\partial \Psi^*}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi^*$$

Multiplying the first equation with Ψ^* and the second with Ψ gives

$$i\hbar \Psi^* \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \Psi^* \nabla^2 \Psi$$

$$-i\hbar \Psi \frac{\partial \Psi^*}{\partial t} = -\frac{\hbar^2}{2m} \Psi \nabla^2 \Psi^*$$

Subtracting the two equations gives:

$$i\hbar \left(\Psi^* \frac{\partial \Psi}{\partial t} + \Psi \frac{\partial \Psi^*}{\partial t} \right) = -\frac{\hbar^2}{2m} \left(\Psi^* \nabla^2 \Psi - \Psi \nabla^2 \Psi^* \right) \Rightarrow$$

$$-\frac{\hbar^2}{2m} \left(\Psi^* \nabla^2 \Psi - \Psi \nabla^2 \Psi^* \right) = i\hbar \frac{\partial}{\partial t} (\Psi^* \Psi) = i\hbar \frac{\partial \rho}{\partial t}$$

From the previous, the probability current density is given by

$$\mathbf{j} = \frac{\hbar}{2im}\Big(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*\Big) \tag{1.1.4}$$

1.1.1 Variable constant of motion

The time-dependent evolution of a quantum mechanical state is given by

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t) = \hat{H}\Psi(\mathbf{x},t) \Rightarrow \hat{H}\Psi(\mathbf{x},t) = E\Psi(\mathbf{x},t)$$

The time dependent eigenstate of $\hat{H}$ is then given by $\Psi(\mathbf{x},t) = \Psi(\mathbf{x})e^{-iEt/\hbar}$.

In general the expectation value of the operator $\hat{A}$ is given by

$$\langle\hat{A}\rangle = \langle\Psi|\hat{A}|\Psi\rangle = \int \Psi^\dagger\hat{A}\Psi d^3x,$$

where $\Psi^\dagger = (\Psi^*)^T$ the complex transpose conjugate.

The time-dependence of the expectation value of the operator $\hat{A}$, assuming that it is time independent and thus $\partial A/\partial t = 0$, is given by

$$\frac{d\langle\hat{A}\rangle}{dt} = \int\Big[\frac{\partial\Psi^\dagger}{\partial t}\hat{A}\Psi + \Psi^\dagger\hat{A}\frac{\partial\Psi}{\partial t}\Big]d^3x$$

If in the equation above we replace the time derivatives of Ψ and $\Psi^\dagger$ using their Schrodinger equations:

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi$$

and

$$-i\hbar\frac{\partial\Psi^\dagger}{\partial t} = (\hat{H}\Psi)^\dagger$$

we then get

$$\frac{d\langle\hat{A}\rangle}{dt} = \int\Big[\frac{i}{\hbar}(\hat{H}\Psi)^\dagger(\hat{A}\Psi) + \Psi^\dagger\hat{A}\Big(\frac{-i}{\hbar}\hat{H}\Psi\Big)\Big]d^3x =$$

$$\frac{i}{\hbar}\int\Big[\Psi^\dagger\hat{H}\hat{A}\Psi - \Psi^\dagger\hat{A}\hat{H}\Psi\Big]d^3x = \frac{i}{\hbar}\int\Big[\Psi^\dagger(\hat{H}\hat{A} - \hat{A}\hat{H})\Psi\Big]d^3x \Rightarrow$$

$$\frac{d\langle\hat{A}\rangle}{dt} = \frac{i}{\hbar}\langle[\hat{H},\hat{A}]\rangle$$

That means that if $[\hat{H},\hat{A}]$ then the observable represented by the operator $\hat{A}$ is a constant of motion.

1.1.2 Uncertainty principle

The operators that describe the x-spacial coordinate and the x-coordinate of the momentum of a particle are represented by operators that obey:

$$\hat{X}\Psi = x\Psi$$

and

$$\hat{p}_x\Psi = -i\hbar\frac{\partial}{\partial x}\Psi$$

The commutator $[\hat{x},\hat{p}_x]$ when acting on an eigenfunction gives

$$[\hat{x},\hat{p}_x]\Psi = \hat{x}\hat{p}_x\Psi - \hat{p}_x\hat{x}\Psi = x(-i\hbar)\frac{\partial\Psi}{\partial x} + i\hbar\frac{\partial}{\partial x}|(x\Psi) =$$

$$-i\hbar x\frac{\partial\Psi}{\partial x} + i\hbar x\frac{\partial\Psi}{\partial x} + i\hbar\Psi = i\hbar\Psi \Rightarrow$$

$$[\hat{x},\hat{p}_x] = i\hbar$$

The general uncertainty principle between two quantities A and B reads

$$(\Delta A)(\Delta B) \geq \frac{1}{2}|\langle i[\hat{A},\hat{B}]\rangle|,$$

where $(\Delta A)^2 = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2$

That brings the uncertainty principle for the position and momentum:

$$\Delta x\Delta p_x \geq \frac{\hbar}{2}$$

1.2 Klein-Gordon equation

The requirement of a relativistic formulation of a quantum mechanical system is that the equations that this system obey should be Lorentz invariant. The Schrodinger equation, however, of Eq. 1.1.1 is of first order in time and second order in space. It certainly does not lead to any Lorentz invariance! This non-invariance is a consequence of how it was constructed: we used the non-relativistic relationship of energy and momentum i.e. $E = p^2/2m$ and we replaced the variables with the quantum mechanical operators.

The first attempt to construct a relativistic theory of quantum mechanics came with the usage of the so-called Klein-Gordon equation. As a starting point we have the relativistic relation between energy and momentum (note the transition to natural units!!!):

$$E^2 = p^2 + m^2 \rightarrow \hat{E}^2\Psi = \hat{p}^2\Psi + m^2\Psi,$$

where $\hat{p} = -i\nabla$ and $\hat{E} = i\frac{\partial}{\partial t}$. If we replace the form of the relevant operators in the equation above, we get

$$\frac{\partial^2}{\partial t^2}\Psi(\mathbf{r},t) + \nabla^2\Psi(\mathbf{r},t) - m^2\Psi(\mathbf{r},t) \Rightarrow \partial_\mu\partial^\mu\Psi(\mathbf{r},t) = 0 \Rightarrow$$

$$(\partial_\mu \partial^\mu + m^2)\Psi(\mathbf{r},t) = 0 \tag{1.2.1}$$

In the previous equation we used the notation

$$\partial_\mu \partial^\mu = \frac{\partial^2}{\partial t^2} - \nabla^2 = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}$$

Equation 1.2.1 can be written in a more elegant way by introducing the D'Alembertian operator defined as

$$\Box \equiv \partial_\mu \partial^\mu = \frac{\partial^2}{\partial t^2} - \nabla^2$$

so that the Klein-Gordon equation can take its final form:

$$(\Box + m^2)\Psi(\mathbf{r},t) = 0 \tag{1.2.2}$$

The corresponding Lagrangian density can be written as

$$\mathscr{L}_{\text{K-G}} = \frac{1}{2}(\partial_\mu \Phi)(\partial^\mu \Phi) - \frac{1}{2}m^2\Phi^2 \tag{1.2.3}$$

Equation 1.2.3 describes a system consisting of a spin-0 scalar field. Applying the Euler-Lagrange equation of Eq. **??**, we can get back Eq. 1.2.2.

The Klein-Gordon equation has plane wave solutions of the form

$$\Psi(\mathbf{r},t) = Ne^{i(\mathbf{pr}-Et)},$$

with energies

$$E = \pm\sqrt{p^2 + m^2}$$

In classical mechanics the solutions with negative energy are dismissed as unphysical but in quantum mechanics all solutions are accepted. These negative energies do not create a real problem at this stage but this changes once we calculate the probability density and current starting from:

$$\frac{\partial^2 \Psi}{\partial t^2} = \nabla^2\Psi - m^2\Psi$$

and

$$\frac{\partial^2 \Psi^*}{\partial t^2} = \nabla^2\Psi^* - m^2\Psi$$

Multiplying the first equation with Ψ^* and the second equation with Ψ one gets

$$\Psi^* \frac{\partial^2 \Psi}{\partial t^2} = \Psi^*\nabla^2\Psi - \Psi^* m^2\Psi$$

$$\Psi \frac{\partial^2 \Psi^*}{\partial t^2} = \Psi\nabla^2\Psi^* - \Psi m^2\Psi$$

Subtracting the second from the first equation above one gets:

$$\Psi^* \frac{\partial^2 \Psi}{\partial t^2} - \Psi \frac{\partial^2 \Psi^*}{\partial t^2} = \Psi^* \nabla^2 \Psi - \Psi \nabla^2 \Psi^* - m^2 \Psi^* \Psi + \Psi m^2 \Psi \Rightarrow$$

$$\frac{\partial}{\partial t}\Big(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t}\Big) = \nabla\Big(\Psi^* \nabla \Psi - \Psi \nabla \Psi^*\Big) \Rightarrow$$

$$\rho = i\Big(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t}\Big)$$

$$\mathbf{j} = -i\Big(\Psi^* \nabla \Psi - \Psi \nabla \Psi^*\Big)$$

In the equation above, we include the factor "i" to make sure the probability density is real.

For a plane-wave solution we get:

$$\rho = 2|N|^2 E$$

$$\mathbf{j} = 2|N|^2 \mathbf{p}$$

If one wants to compactly add the above scalar and vector into one object, then we can introduce the 4–vector j^μ defined as

$$j^\mu = 2|N|^2 p^\mu$$

This clearly illustrates the next problem we are facing: the probability is proportional to the energy of a particle; however, the energy can take negative values as we have seen. This leads to the possibility of a negative probability density!

1.3 Dirac equation

The extension of the Klein-Gordon equation from Dirac solved the problem of the negative values of the probability density and also provided a natural description of probably the most important category of particles: the one containing spin-1/2 fermions.

Dirac looked for a first order equation in both space and time. What he tried was an equation of the form

$$\hat{E}\Psi = (\alpha \hat{p} + \beta m)\Psi \Rightarrow i\frac{\partial}{\partial t}\Psi = (-i\alpha_x \frac{\partial}{\partial x} - i\alpha_y \frac{\partial}{\partial y} - i\alpha_z \frac{\partial}{\partial z} + \beta m)\Psi,$$

where α and β are unknown for the time being mathematical entities. The solutions must satisfy the energy-momentum relation at its relativistic form and thus satisfy the Klein-Gordon equation. To do so, let us start by squaring the previous equation:

$$-\frac{\partial^2}{\partial t^2}\Psi = (i\alpha_x \frac{\partial}{\partial x} + i\alpha_y \frac{\partial}{\partial y} + i\alpha_z \frac{\partial}{\partial z} - \beta m)(i\alpha_x \frac{\partial}{\partial x} + i\alpha_y \frac{\partial}{\partial y} + i\alpha_z \frac{\partial}{\partial z} - \beta m) \Rightarrow$$

$$\frac{\partial^2 \Psi}{\partial t^2} = \alpha_x^2 \frac{\partial^2 \Psi}{\partial x^2} + \alpha_y^2 \frac{\partial^2 \Psi}{\partial y^2} + \alpha_z^2 \frac{\partial^2 \Psi}{\partial z^2} - \beta^2 m^2 \Psi$$

$$(\alpha_x \alpha_y + \alpha_y \alpha_x)\frac{\partial^2 \Psi}{\partial x \partial y} + (\alpha_y \alpha_z + \alpha_z \alpha_y)\frac{\partial^2 \Psi}{\partial y \partial z} + (\alpha_z \alpha_x + \alpha_x \alpha_z)\frac{\partial^2 \Psi}{\partial z \partial x} +$$

$$i(\alpha_x\beta+\beta\alpha_x)m\frac{\partial\Psi}{\partial x}+i(\alpha_y\beta+\beta\alpha_y)m\frac{\partial\Psi}{\partial y}+i(\alpha_z\beta+\beta\alpha_z)m\frac{\partial\Psi}{\partial z}$$

To reduce to the Klein-Gordon equation, α and β must satisfy:

- $\alpha_x^2=\alpha_y^2=\alpha_z^2=\beta^2=I$
- $\alpha_j\beta+\beta\alpha_j=0$
- $\alpha_j\alpha_k+\alpha_k\alpha_j=0$, for each $j\neq k$.

Let us now try to find the form of the unknown quantities α_i and β.

The variables α_i and β can not be real numbers. They are matrices!!!

In addition, from the fact that $\alpha_j\beta=-\beta\alpha_j$, that $\beta^2=I$ and that the trace[1] of the multiplication of three matrices (let's say A, B and C) is $\mathrm{Tr(ABC)}=\mathrm{Tr(BCA)}$ (i.e. cyclic identity), it can be seen that α_i and β are matrices with zero trace.

It can also be shown the the eigenvalues of α_i and β are ±1:

$$\alpha_iX=\lambda X\Rightarrow\alpha_i^2X=\lambda\alpha_iX\Rightarrow X=\lambda^2X\Rightarrow\lambda=\pm1,$$

since $\alpha_i^2=I$.

Furthermore, the sum of the eigenvalues of a matrix is equal to its trace. The only way that the trace can be 0 is if α_i and β have even dimensions.

Moreover, due to the fact that the Dirac Hamiltonian operator $\hat{H}=\alpha\hat{p}+\beta m$ is Hermitian, it follows that the matrices α_i and β should also be Hermitian: $\alpha_x^\dagger=\alpha_x$, $\alpha_y^\dagger=\alpha_y$, $\alpha_z^\dagger=\alpha_z$ and $\beta^\dagger=\beta$.

Since there are only three mutually anticommuting 2×2 matrices with zero trace, the Pauli matrices, α_i and β should be at least 4×4. The Dirac Hamiltonian is thus a 4×4 matrix of operators that acts on a 4–component wavefunction, the Dirac spinor:

$$\Psi=\begin{pmatrix}\Psi_1\\ \Psi_2\\ \Psi_3\\ \Psi_4\end{pmatrix}$$

The size of the spinor is driven by the existence of the β–term i.e. by the fact that particles have mass! Otherwise the world would be described just by the Pauli matrices.

Let us now define the form of the α and β matrices:

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

and

$$\alpha_i=\begin{pmatrix}0 & \sigma_i\\ \sigma_i & 0\end{pmatrix},$$

where $I=\begin{pmatrix}1&0\\0&1\end{pmatrix}$, $\sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix}$, $\sigma_y=\begin{pmatrix}0&-i\\i&0\end{pmatrix}$ and $\sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$.

Note that this is one of the many representations that these matrices can take. The physical predictions obtained using it do not depend on this choice of representation.

[1] For a $n\times n$ matrix A, the trace of the matrix is $\mathrm{tr(A)}=\sum_{i=1}^{n}\alpha_{ii}$, where α_{ii} are the diagonal elements of A

1.3.1 Probability density and current

To derive the probability density and current, since the wavefunctions are 4–component spinors, we use the transpose conjugate instead of just the complex one i.e. $\Psi^* \to \Psi^\dagger = (\Psi*)^T$. The Dirac equations for Ψ and for $\Psi^\dagger$ are:

$$i\frac{\partial}{\partial t}\Psi = (-i\alpha_x\frac{\partial}{\partial x} - i\alpha_y\frac{\partial}{\partial y} - i\alpha_z\frac{\partial}{\partial z} + \beta m)\Psi$$

and

$$-i\frac{\partial}{\partial t}\Psi^\dagger = i\frac{\partial\Psi^\dagger}{\partial x}\alpha_x^\dagger + i\frac{\partial\Psi^\dagger}{\partial y}\alpha_y^\dagger + i\frac{\partial\Psi^\dagger}{\partial z}\alpha_z^\dagger + m\Psi^\dagger\beta^\dagger$$

Multiplying the first with $\Psi^\dagger$ from the left and the second with Ψ from the right and then subtracting the two gives:

$$\Psi^\dagger\Big(-i\alpha_x\frac{\partial\Psi}{\partial x} - i\alpha_y\frac{\partial\Psi}{\partial y} - i\alpha_z\frac{\partial\Psi}{\partial z} + \beta m\Psi\Big) - \Big(i\frac{\partial\Psi^\dagger}{\partial x}\alpha_x^\dagger + i\frac{\partial\Psi^\dagger}{\partial y}\alpha_y^\dagger + i\frac{\partial\Psi^\dagger}{\partial z}\alpha_z^\dagger + m\Psi^\dagger\beta^\dagger\Big)\Psi = i\Psi^\dagger\frac{\partial}{\partial t}\Psi + i\frac{\partial}{\partial t}\Psi^\dagger\Psi$$

$\alpha_i = \alpha_i^\dagger$. The previous equation can be simplified if we write:

$$\Psi^\dagger\alpha_x\frac{\partial\Psi}{\partial x} + \frac{\partial\Psi^\dagger}{\partial x}\alpha_x\Psi = \frac{\partial}{\partial x}(\Psi^\dagger\alpha_x\Psi)$$

and

$$\Psi^\dagger\frac{\partial\Psi}{\partial t} + \frac{\partial\Psi^\dagger}{\partial t}\Psi = \frac{\partial}{\partial t}(\Psi^\dagger\Psi)$$

This gives:

$$\nabla(\Psi^\dagger\alpha\Psi) + \frac{\partial}{\partial t}(\Psi^\dagger\Psi) = 0$$

where $\Psi^\dagger = \begin{pmatrix}\Psi_1^* & \Psi_2^* & \Psi_3^* & \Psi_4^*\end{pmatrix}$.

That means that the probability density and current are now given by:

$$\rho = \Psi^\dagger\Psi$$

and

$$\mathbf{j} = \Psi^\dagger\alpha\Psi$$

The problem with the probability density turning negative is now solved:

$$\rho = \Psi^\dagger\Psi = |\Psi_1|^2 + |\Psi_2|^2 + |\Psi_3|^2 + |\Psi_4|^2 > 0$$

1.3.2 Covariant form of the Dirac equation

Let us first introduce a new set of matrices that will be used quite extensively from now on: they will be referred to as the γ-matrices. The are given by

$$\begin{pmatrix} \gamma^0 \equiv \beta \\ \gamma^1 \equiv \beta\alpha_x \\ \gamma^2 \equiv \beta\alpha_y \\ \gamma^3 \equiv \beta\alpha_z \end{pmatrix} \tag{1.3.1}$$

It is also useful to define now also the adjoint spinor $\overline{\Psi}$:

$$\overline{\Psi} \equiv \Psi^\dagger \gamma^0 \tag{1.3.2}$$

The Dirac equation can now be written as follows:

$$i\frac{\partial}{\partial t}\Psi = (-i\alpha_x\frac{\partial}{\partial x} - i\alpha_y\frac{\partial}{\partial y} - i\alpha_z\frac{\partial}{\partial z} + \beta m)\Psi \Rightarrow$$

$$i\beta\frac{\partial\Psi}{\partial t} + i\beta\alpha_x\frac{\partial\Psi}{\partial x} + i\beta\alpha_y\frac{\partial\Psi}{\partial y} + i\beta\alpha_z\frac{\partial\Psi}{\partial z} - \beta^2 m\Psi = 0 \Rightarrow$$

$$i\gamma^0\frac{\partial\Psi}{\partial t} + i\gamma^1\frac{\partial\Psi}{\partial x} + i\gamma^2\frac{\partial\Psi}{\partial y} + i\gamma^3\frac{\partial\Psi}{\partial z} - m\Psi = 0 \Rightarrow$$

If we now introduce the notation:

$$\gamma_\mu = (\gamma^0\ \gamma^1\ \gamma^2\ \gamma^3)$$

and

$$\partial_\mu = (\partial_0\ \partial_1\ \partial_2\ \partial_3) = \left(\tfrac{\partial}{\partial t}\ \tfrac{\partial}{\partial x}\ \tfrac{\partial}{\partial y}\ \tfrac{\partial}{\partial z}\right)$$

$$\partial^\mu = \begin{pmatrix} \partial_0 \\ -\partial_1 \\ -\partial_2 \\ -\partial_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial t} \\ -\frac{\partial}{\partial x} \\ -\frac{\partial}{\partial y} \\ -\frac{\partial}{\partial z} \end{pmatrix}$$

the previous equation takes its covariant form:

$$(i\gamma_\mu\partial^\mu - m)\Psi = 0 \tag{1.3.3}$$

1.3.3 Solutions of the Dirac equation

In this section we will see what the solutions of the Dirac equation represent. This is quite important to understand (not the way we reach the solution but what they physically mean)! Let us start from the free particle plane-wave solutions of the form

$$\Psi(\mathbf{x},t) = u(E,\mathbf{p})e^{i(\mathbf{px}-Et)}$$

where $u(E,\mathbf{p})$ is a 4–component Dirac spinor.

Since the spacial and time coordinates dependence of Ψ is contained in the exponent we can write:

$$\partial_0\Psi = -iE\Psi$$

$$\partial_1\Psi = ip_x\Psi$$

$$\partial_2\Psi = ip_y\Psi$$

$$\partial_3\Psi = ip_z\Psi$$

so that the Dirac equation gives:

$$(i\gamma_\mu\partial^\mu - m)\Psi = 0 \Rightarrow (\gamma^0 E - \gamma^1 p_x - \gamma^1 p_y - \gamma^1 p_z - m)u(E,\mathbf{p})e^{i(\mathbf{px}-Et)} = 0 \Rightarrow$$

$$(\gamma_\mu p^\mu - m)u = 0$$

1.3.3.1 Solutions for particles at rest

For a particle at rest $\mathbf{p} = 0$ and the free particle wavefunction can be written as

$$\Psi(\mathbf{x},t) = u(E,0)e^{-iEt}$$

$$(\gamma_\mu p^\mu - m)u = 0 \Rightarrow E\gamma^0 u = mu \Rightarrow E\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix} = m\begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix}$$

This leads to four orthogonal solution of the form:

$$u_1(E,0) = N\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$u_2(E,0) = N\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

which correspond to solutions with positive energy $E = +m$ and

$$u_3(E,0) = N \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$u_4(E,0) = N \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

which correspond to solutions with negative energy $E = -m$.

Let us make some remarks at this stage:

- The parameter N is the normalisation of Ψ
- The solution u_1 and u_2 represent two particles (note the positive energy) with spin-up and spin-down, respectively, states.
- The solution u_3 and u_4 represent two antiparticles (note the negative energy) with spin-up and spin-down, respectively, states.
- The time-dependent solutions of the Dirac equation for a particle at rest are given by

$$\Psi_1 = N \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} e^{-imt}$$

$$\Psi_2 = N \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} e^{-imt}$$

$$\Psi_3 = N \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{imt}$$

$$\Psi_4 = N \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{imt}$$

1.3.3.2 General solutions

The general solutions of the Dirac equation can be derived from the solutions for the case of a particle at rest, using the transformation properties of the Dirac spinors or using the following:

$$(\gamma_\mu p^\mu - m)u = 0 \Rightarrow (E\gamma^0 - p_x\gamma^1 - p_y\gamma^2 - p_z\gamma^3 - m)u = 0 \Rightarrow$$

$$\left[\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} E - \begin{pmatrix} 0 & \boldsymbol{\sigma}\mathbf{p} \\ -\boldsymbol{\sigma}\mathbf{p} & 0 \end{pmatrix} - m \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \right] u = 0$$

where

$$\sigma\mathbf{p} = \sigma_x p_x + \sigma_y p_y + \sigma_z p_z = \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}$$

We can now rewrite the spinor u as a two–component vector $u = (u_A, u_B)$ such that the previous equation can be written

$$\begin{pmatrix} (E-m)I & -\sigma\mathbf{p} \\ \sigma\mathbf{p} & -(E+m)I \end{pmatrix} \begin{pmatrix} u_A \\ u_B \end{pmatrix} = 0$$

The previous equation has the following solutions:

$$u_A = \frac{\sigma\mathbf{p}}{E-m} u_B$$

and

$$u_B = \frac{\sigma\mathbf{p}}{E+m} u_A$$

Two solution to the free Dirac equation can be found by taking the simplest orthogonal choices for u_A:

$$u_A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$u_A = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$u_B = \frac{\sigma\mathbf{p}}{E+m} u_A = \frac{1}{E+m} \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} u_A$$

The first two solutions of the free particle Dirac equation can thus be written as:

$$u_1(E, \mathbf{p}) = N_1 \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x + ip_y}{E+m} \end{pmatrix}$$

and

$$u_2(E, \mathbf{p}) = N_2 \begin{pmatrix} 0 \\ 1 \\ \frac{p_x - ip_y}{E+m} \\ -\frac{p_z}{E+m} \end{pmatrix}$$

Note that the choice for the representation of u_A was arbitrary: any choice would be equally fine!

The other two solutions can be found by writing

$$u_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$u_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

which leads to

$$u_3(E,\mathbf{p}) = N_3 \begin{pmatrix} \frac{p_z}{E-m} \\ \frac{p_x+ip_y}{E-m} \\ 1 \\ 0 \end{pmatrix}$$

and

$$u_4(E,\mathbf{p}) = N_4 \begin{pmatrix} \frac{p_x-ip_y}{E-m} \\ -\frac{p_z}{E-m} \\ 0 \\ 1 \end{pmatrix}$$

In terms of physical energy, the solutions of the Dirac equation are $\Psi_i = u_i e^{i(\mathbf{px}-Et)}$ and they represent again two particles with spin-up (i.e. u_1) and spin-down (i.e. u_3) and two antparticles with spin-up (i.e. u_3) and spin-down (i.e. u_4).

1.3.4 Dirac Lagrangian density

Finally, the Lagrangian density that describes systems consisting of spin-1/2 fields is:

$$\mathscr{L}_{\text{DIRAC}} = i\overline{\Psi}\gamma_\mu \partial^\mu \Psi - m\overline{\Psi}\Psi \tag{1.3.4}$$

from where with the usage of the Euler-Lagrange equations we get back the equation of motion of the system that obeys:

$$(i\gamma_\mu \partial^\mu - m)\Psi = 0$$

and

$$(i\gamma_\mu \partial^\mu + m)\overline{\Psi} = 0$$

1.4 Proca equation

The Lagrangian density for a massless spin-1 field is given by

$$\mathscr{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu},$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ the field strength tensor. The corresponding Lagrangian density for a massive spin-1 boson with mass MM is given by

$$\mathscr{L}_{\text{PROCA}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}M^2 A_\mu A^\mu \tag{1.4.1}$$

The equation of motion of the massive field can be extracted by the Euler-Lagrange equations:

$$\partial_\mu\Big(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\Phi)}\Big)-\frac{\partial\mathscr{L}}{\partial\Phi}=0\Rightarrow$$

$$(\Box+m^2)A^\mu-\partial^\mu(\partial_\nu A^\nu)=0$$

In case of a massless field of spin-1 (e.g. a photon), the previous equation takes the following form

$$\Box A^\mu-\partial^\mu(\partial_\nu A^\nu)=0$$

1.5 Yang-Mills theory

In 1954, Young and Mills extended the idea of applying local gauge invariance, a notion we first encountered in Chapter **??**, to higher order groups than the $\mathrm{U}(1)$. Local gauge invariance was first applied in $\mathrm{SU}(2)$ and then in $\mathrm{SU}(3)$ and derived the theory of the weak and the strong interactions. In the Standard Model all fundamental interactions are derived in this same way. We will talk about these in the following chapters.

Let us briefly review the first steps of the Young–Mills theory that will be used extensively in the chapters where we are going to encounter the strong and the weak interactions. Let us suppose that we have to spin-1/2 fields, Ψ_1 and Ψ_2. The Lagrangian density in the absence of any external field and thus interactions is given by:

$$\mathscr{L}=[i\overline{\Psi}_1\gamma_\mu\partial^\mu\Psi_1-m\overline{\Psi}_1\Psi_1]+[i\overline{\Psi}_2\gamma_\mu\partial^\mu\Psi_2-m\overline{\Psi}_2\Psi_2]$$

It is clear that the previous is just a sum of two Dirac Lagrangian densities. So far nothing special. But we can rewrite the previous equation in a more compact way, by introducing a two-component vector

$$\Psi\equiv\begin{pmatrix}\Psi_1\\ \Psi_2\end{pmatrix}$$

Note that both Ψ_1 and Ψ_2 are four component Dirac spinors! Things get a bit more complicated but we can get away from this complication by introducing an additional index so that one index corresponds to the particle (i.e. corresponds to either Ψ_1 or Ψ_2), while the second index refers to the component of each of the Dirac spinors. The adjoint spinor is then

$$\overline{\Psi}\equiv\begin{pmatrix}\overline{\Psi_1} & \overline{\Psi_2}\end{pmatrix}$$

The Lagrangian density of our system takes now the following form

$$\mathscr{L}=i\overline{\Psi}\gamma_\mu\partial^\mu\Psi-M\overline{\Psi}\Psi$$

where

$$M=\begin{pmatrix}m_1 & 0\\ 0 & m_2\end{pmatrix}$$

is the so-called “mass“ matrix. If the two masses are identical then the Lagrangian density takes a rather familiar form given by

$$\mathscr{L}_{\mathrm{Young-Mills}}=i\overline{\Psi}\gamma_\mu\partial^\mu\Psi-M\overline{\Psi}\Psi \tag{1.5.1}$$

Although this looks like Eq. 1.3.4, one has to point out that the main difference is that Ψ is now a two–component vector of Dirac spinors. The Lagrangian density of Eq. 1.5.1 takes a more general global invariance than a simple U(1), since now the transformation $\Psi \rightarrow \Psi' = U\Psi$ involves a 2×2 matrix (i.e. U).