6

Free Theory

In this chapter we will derive the basic equations for a physical theory of **free** (=non-interacting) fields[1] from symmetry. We will

- derive the Klein-Gordon equation using the $(0,0)$ representation of the Lorentz group,
- derive the Dirac equation using the $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$ representation of the Lorentz group,
- derive the Proca equations from the vector $(\frac{1}{2},\frac{1}{2})$ representation of the Lorentz group which in the massless limit become the famous Maxwell equations.

[1] Although we specify here for fields we will see in a later chapter how the equations we derive here can be used to describe particles, too.

6.1 Lorentz Covariance and Invariance

In the following sections, we will derive the fundamental equations of motion of the standard model of particle physics, which is the best physical theory that we have. We want that these equations are the same in all inertial frames, because otherwise would have a different equation for each possible frame of reference. This would be useless because there is no preferred frame of reference in special relativity. The technical term for an equation that looks the same in all inertial frames of reference is **Lorentz covariant equation**. An object is Lorentz covariant if it transforms under a given representation of the Lorentz group. For example, a vector A_μ, transforms according to the $(\frac{1}{2},\frac{1}{2})$ representation and is therefore Lorentz covariant. This means $A_\mu \to A'_\mu$, but not something completely different. On the other hand, for example, a term of the form $A_1 + A_3$ is not Lorentz covariant, because it does not transform according to a representation of the Lorentz group. This does not mean that we do not know how it transforms. The transformation properties can be easily derived from the transformation laws for A_μ, but nevertheless this term looks

J. Schwichtenberg, *Physics from Symmetry*, Undergraduate
Lecture Notes in Physics, https://doi.org/10.1007/978-3-319-66631-0_6

completely different in different inertial frames. In a boosted frame, it may look like $A_2 + A_4$. An equation that involves only Lorentz covariant objects is called a Lorentz covariant equation. For example,

$$A_\mu + 7B_\mu + C_\nu A^\nu D_\mu = 0$$

is a Lorentz covariant equation, because in another coordinate system it reads

$$\Lambda_\mu^\nu A_\nu + 7\Lambda_\mu^\nu B_\nu + \Lambda_\nu^\rho C_\rho \Lambda_\eta^\nu A^\eta \Lambda_\mu^\sigma D_\sigma = A'_\mu + 7B'_\mu + C'_\nu A'^\nu D'_\mu = 0.$$

We see that it looks the same. An equation containing only some components of such objects is, in general, not Lorentz covariant and therefore looks completely different in each inertial frame.

To make sure we only end up with Lorentz **covariant** equations, we require the action S to be Lorentz **invariant**. This means it should only contain terms that stay exactly the same when changing the frame of reference. In other words: The action is only allowed to contain terms that do not change under Lorentz transformations. We get the equations of motion from the action[2] S. Now if S depends on the frame of reference, so would the terms in the equations that follow from it and therefore these equations can not be Lorentz covariant.

[2] Recall that we minimize the action and the result of this minimization procedure is the Euler-Lagrange equation, which yields the equation of motion for our system.

As already discussed in the last chapter, we can use the more restrictive requirement that the Lagrangian should be invariant, because if the Lagrangian is invariant the action is, too.

6.2 Klein-Gordon Equation

We now start with the simplest possible case: scalars, which transform according to the $(0,0)$ representation of the Lorentz group. To specify the equation of motion for scalars we need to find the corresponding Lagrangian. A general Lagrangian that is compatible with our restrictions[3] is

[3] Discussed in Section 4.2: We only consider terms of order $0, 1$ and 2 in Φ. The term with the lowest possible derivative will become clear in a moment.

$$\mathscr{L} = A\Phi^0 + B\Phi + C\Phi^2 + D\partial_\mu\Phi + E\partial_\mu\Phi\partial^\mu\Phi + F\Phi\partial_\mu\Phi\,. \qquad (6.1)$$

Firstly, take note that we are considering the Lagrangian density $\mathscr{L}$ and not L itself and we get our physical theory from the action

$$S = \int dx \mathscr{L}, \qquad (6.2)$$

where dx is to be understood as the integral over space and time. Therefore a term like $\Phi\partial_\mu\partial^\mu\Phi$ would be redundant, because it is equivalent to the term $\partial_\mu\Phi\partial^\mu\Phi$, as we can see if we integrate by parts[4].

[4] The boundary terms, as usual, vanish, because fields vanish at infinity. Recall that this follows, because field configurations that do not vanish at infinity correspond to infinite field energy, which is non-physical (Section 2.3).

In addition, Lorentz **invariance** restricts the Lagrangian to be a scalar. Therefore, all odd powers in ∂_μ, like in $\partial_\mu \Phi$ are forbidden. What about the constants i.e. a and c etc. having a Lorentz index? This would mean that a, c are four-vectors, specifying a direction in spacetime and therefore violating the assumption of isotropy of space. We can neglect the constant term, i.e. $A = 0$, because we get our physical theory from the Euler-Lagrange equation and a constant in the Lagrangian has no influence on the equation of motion[5]. In addition, we ignore the term linear in Φ, i.e. $B = 0$, because it leads, using the Euler-Lagrange equation, to a constant in our equations of motion[6]. At least in a free theory, the absolute value of the field has no relevance and thus we can always perform a redefinition $\Phi \to \Phi' = \Phi + \text{const.}$ such that the constant term vanishes from the equation of motion[7]. What remains is

$$\mathscr{L} = C\Phi^2 + E\partial_\mu \Phi \partial^\mu \Phi \,. \tag{6.3}$$

As the heading of this chapter indicates we want to develop a free theory, which means there is just one Φ and no terms of the form $\Phi_1 \Phi_2$. Terms like this will be investigated in the next chapter, when we develop a theory describing interactions.

There is one last thing to note: we are left with only two constants C and E. By using variational calculus we are able to combine these into just one constant, because an overall constant in the Lagrangian has no influence on the physics[8]. Nevertheless, it is conventional to include an overall factor $\frac{1}{2}$ into the Lagrangian and call the remaining constant[9] $-m^2$. Therefore, we are finally left with

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2) \,. \tag{6.4}$$

If we now use the variational calculus machinery, which means putting this Lagrangian into the Euler-Lagrange equation (Eq. 4.10), we get the equation of motion

$$\begin{aligned}
&0 = \frac{\partial \mathscr{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)} \right) \\
&\to 0 = \frac{\partial}{\partial \Phi} \left(\frac{1}{2}(\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2) \right) - \partial_\mu \left(\frac{\partial}{\partial(\partial_\mu \Phi)} \left(\frac{1}{2}(\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2) \right) \right) \\
&\to 0 = (\partial_\mu \partial^\mu + m^2)\Phi \,. \qquad (6.5)
\end{aligned}$$

This is the famous **Klein-Gordon equation**, which is the correct equation of motion to describe free spin 0 fields and particles.

[5] See Eq. 4.10: $\frac{\partial \mathscr{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)} \right) = 0$ and therefore $\mathscr{L} \to \mathscr{L} + A$ with some constant A does not change anything: $\frac{\partial(\mathscr{L}+A)}{\partial \Phi} - \partial_\mu \left(\frac{\partial(\mathscr{L}+A)}{\partial(\partial_\mu \Phi)} \right) = \frac{\partial \mathscr{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)} \right)$

[6] $\mathscr{L} \to \mathscr{L} + B\Phi$ yields $\frac{\partial(\mathscr{L}+B\Phi)}{\partial \Phi} - \partial_\mu \left(\frac{\partial(\mathscr{L}+B\Phi)}{\partial(\partial_\mu \Phi)} \right) = \frac{\partial \mathscr{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)} \right) + B = 0$, which is just an additional constant in the equation of motion.

[7] However take note that if the field interacts with other fields, this is no longer true, because such a field shift then also affects the terms that describe the coupling between the fields. This will be discussed in detail in Section. 7.3 in the context of spontaneous symmetry breaking.

[8] $\mathscr{L} \to C\mathscr{L}$ yields
$$\frac{\partial(C\mathscr{L})}{\partial \Phi} - \partial_\mu \left(\frac{\partial(C\mathscr{L})}{\partial(\partial_\mu \Phi)} \right) = 0$$
$$\updownarrow$$
$$\frac{\partial(\mathscr{L})}{\partial \Phi} - \partial_\mu \left(\frac{\partial(\mathscr{L})}{\partial(\partial_\mu \Phi)} \right) = 0$$

[9] The suggestive name of this constant will become clear later, because we will see that it coincides with the mass of particles described by this Lagrangian.

6.2.1 Complex Klein-Gordon Field

For spin 0 fields, we are able to construct a Lorentz invariant Lagrangian without using the complex conjugate of the scalar field. This will not be the case for spin $\frac{1}{2}$ fields Ψ and this curious fact will have very interesting consequences[10]. Nevertheless, nothing prevents us from investigating the equally Lorentz invariant Lagrangian

$$\mathscr{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi$$

as many textbooks do. This is simply the same as investigating **two** scalar fields of equal mass at the same time

$$\mathscr{L} = \frac{1}{2}\partial_\mu \phi_1 \partial^\mu \phi_1 - \frac{1}{2} m^2 \phi_1^2 + \frac{1}{2}\partial_\mu \phi_2 \partial^\mu \phi_2 - \frac{1}{2} m^2 \phi_2^2 \,,$$

because we have

$$\phi \equiv \frac{1}{\sqrt{2}} \left(\phi_1 + i\phi_2\right).$$

Again Lorentz symmetry dictates the form of the Lagrangian. Elementary scalar (=spin 0) particles are very rare. In fact, only one is experimentally verified: **the Higgs boson**. However this Lagrangian can also be used to describe composite systems like mesons. We will not investigate this Lagrangian any further and most textbooks use it only for "training purposes".

[10] To spoil the surprise: there is an antiparticle for each spin $\frac{1}{2}$ particle. Using complex fields is the same as considering two fields at the same time as explained in the text below. Therefore we are forced by Lorentz invariance to use **two** (closely connected) fields at the same time, which are commonly interpreted as particle and antiparticle fields.

6.3 Dirac Equation

> Another story told of Dirac is that when he first met Richard Feynman, he said after a long silence "I have an equation. Do you have one too?"
>
> **- Anthony Zee**[11]

In this section we want to find the equation of motion for free spin $\frac{1}{2}$ fields/particles. We will use the $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ representation of the Lorentz group, because a theory that respects symmetry under parity transformations must include the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations at the same time[12]. The objects transforming under this representation are called Dirac spinors and combine right-chiral and left-chiral Weyl spinors into one object[13]:

$$\Psi \equiv \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \begin{pmatrix} \chi_a \\ \xi^{\dot{a}} \end{pmatrix}. \tag{6.6}$$

Now, we need to search for Lorentz invariant objects constructed from Dirac spinors, which we can then put into the Lagrangian. The first step is to search for invariants constructed from our left-chiral and right-chiral Weyl spinors.

[11] Anthony Zee. *Quantum Field Theory in a Nutshell*. Princeton University Press, 1st edition, 3 2003. ISBN 9780691010199

[12] We showed in Section 3.7.9 that a parity transformation transforms the $(\frac{1}{2}, 0)$ representation, into the $(0, \frac{1}{2})$ representation.

[13] This was discussion in Section 3.7.9.

We will use the Van-der-Waerden notation, which was introduced in Section 3.7.7. Two possibly invariants are[14]

$$I_1 := \chi_{\dot{a}}^T \xi^{\dot{a}} = (\chi_a^\star)^T \xi^{\dot{a}} = (\chi_a)^\dagger \xi^{\dot{a}} = (\chi_L)^\dagger \xi_R \tag{6.7}$$

and

$$I_2 := (\xi^a)^T \chi_a = ((\xi^{\dot{a}})^\star)^T \chi_a = (\xi_R)^\dagger \chi_L, \tag{6.8}$$

because we always need to combine a lower dotted with an upper dotted and a lower undotted with an upper undotted index, in order to get a Lorentz invariant term. This was shown explicitly in Section 3.7.7. Here we see again that right-chiral and left-chiral spinors are needed in pairs.

Furthermore, we can construct two Lorentz-invariant combinations involving first order derivatives, as we will see in a moment. But first we need to understand how we can write the derivative of a spinor. We learned in Section 3.7.8 how we can construct four-vectors from spinors

$$v_{a\dot{b}} = v_\nu \sigma^\nu_{a\dot{b}} \,,$$

where v_ν transforms like a four-vector. The differentiation operator is therefore in the spinor formalism

$$\partial_{a\dot{b}} = \partial_\nu \sigma^\nu_{a\dot{b}} \,. \tag{6.9}$$

It is conventional to define $\bar{\sigma}^0 = I_{2\times 2}$, $\bar{\sigma}^i = -\sigma^i$ and using this we can construct the following Lorentz invariant terms

$$I_3 := (\chi_{\dot{a}})^T \partial_\mu (\sigma^\mu)^{\dot{a}b} \chi_b = (\chi_L)^\dagger \partial_\mu \bar{\sigma}^\mu \chi_L \tag{6.10}$$

and

$$I_4 := (\xi^a)^T \partial_\mu (\sigma^\mu)_{a\dot{b}} \xi^{\dot{b}} = (\xi_R)^\dagger \partial_\mu \sigma^\mu \xi_R \,. \tag{6.11}$$

In addition to $(\sigma^\mu)_{a\dot{b}}$, we need here also $(\sigma^\mu)^{\dot{a}b}$. The first index must be dotted and the second index undotted to combine properly with the other spinor indices. We get $(\sigma^\mu)^{\dot{a}b}$ by using the spinor metric twice:

$$\begin{aligned}(\sigma^\mu)^{\dot{a}b} &= ((\sigma^\mu)^T)^{b\dot{a}} = \epsilon^{bc}((\sigma^\mu)^T)_{c\dot{d}}(\epsilon^{\dot{a}\dot{d}})^T \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\sigma^\mu)^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^T \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\sigma^\mu)^T \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \bar{\sigma}_\mu \,. \end{aligned} \tag{6.12}$$

[14] There are two other possibilities, which go by the name Majorana mass terms. We already know how we can move spinor indices up and down by using the spinor "metric" ϵ. We can write down a Lorentz invariant term of the form $\epsilon(\chi_L)^\dagger \chi_L$, because $\epsilon(\chi_L)^\dagger = \chi^a$ has an upper-undotted index. $\epsilon(\chi_L)^\dagger$ transforms like a right-chiral spinor, but by writing a term like this we have less degrees of freedom. ξ_R and χ_L both have two components, and therefore we have four degrees of freedom in a term like $(\xi_R)^\dagger \chi_L$. In the term $\epsilon(\chi_L)^\dagger \chi_L = \chi^a \chi_a$, the object that transforms like a right-chiral spinor is not independent of χ_L and we therefore only have two degrees of freedom here. There is a lot more one can say about Majorana spinors and it is currently under (experimental) investigation which type of term is the correct one for neutrinos. One more thing is worth noting: a Majorana spinor is a "real" Dirac spinor. I put real into quotation marks, because usually real means $\Psi^\star \overset{!}{=} \Psi$. For spinors this condition is not Lorentz invariant (because the Lorentz transformations are complex in this representation). If we impose the standard condition ($\Psi^\star \overset{!}{=} \Psi$) in one frame, it will, in general, not hold in another frame. Instead, it is possible to derive a Lorentz invariant "reality" condition for Dirac spinors: $\begin{pmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{pmatrix} \Psi^\star \overset{!}{=} \Psi$, which is commonly interpreted as charge conjugation. This interpretation will be explained in Section 7.1.5. Therefore, a Majorana spinor describes a particle which is equivalent to its charge conjugated particle, commonly called anti-particle. Majorana particles are their own anti-particles and a Majorana spinor is a Dirac spinor with an extra-condition: $\Psi_M \equiv \begin{pmatrix} \chi_L \\ \epsilon \chi_L^\star \end{pmatrix}$ or $\Psi_M \equiv \begin{pmatrix} -\epsilon \xi_R^\star \\ \xi_R \end{pmatrix}$

For example, for σ_3, we have

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{=\sigma_3^T} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}}_{=\bar{\sigma}_3=-\sigma_3}.$$

Don't let yourself get confused why ∂_μ acts only on one spinor. We are going to use these invariants in Lagrangians, which we always evaluate inside of integrals[15] and therefore, we can always integrate by parts to get the other possibility. Therefore, our choice here is no restriction[16].

[15] Remember that we get the equations of motion from the action, which is the integral over the Lagrangian.

[16] We will see this more clearly when we derive the corresponding equations of motion. We get the same equations regardless of where we put ∂_μ and we could put both possibilities into the Lagrangian. This would be longer, but doesn't gives us anything new.

When we introduce the matrices[17]

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma_\mu \\ \bar{\sigma}_\mu & 0 \end{pmatrix} \to \gamma_\mu = \begin{pmatrix} 0 & \bar{\sigma}_\mu \\ \sigma_\mu & 0 \end{pmatrix}, \tag{6.13}$$

[17] We get the matrix with lowered index by using the metric: $\gamma_\mu = \eta_{\mu\nu}\gamma^\nu = \eta_{\mu\nu}\begin{pmatrix} 0 & \sigma^\nu \\ \bar{\sigma}^\nu & 0 \end{pmatrix} = \begin{pmatrix} 0 & \eta_{\mu\nu}\sigma^\nu \\ \eta_{\mu\nu}\bar{\sigma}^\nu & 0 \end{pmatrix} = \begin{pmatrix} 0 & \bar{\sigma}_\mu \\ \sigma_\mu & 0 \end{pmatrix}$, because $\eta_{\mu\nu}\sigma^\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} \sigma^0 \\ \sigma^1 \\ \sigma^2 \\ \sigma^3 \end{pmatrix} = \begin{pmatrix} \sigma^0 \\ -\sigma^1 \\ -\sigma^2 \\ -\sigma^3 \end{pmatrix} = \bar{\sigma}_\mu$.

we can write the invariants we just found using the Dirac spinor formalism. Using the matrices γ_μ and Dirac spinors our invariants can be written as

$$\Psi^\dagger\gamma_0\Psi \quad \text{and} \quad \Psi^\dagger\gamma_0\gamma^\mu\partial_\mu\Psi, \tag{6.14}$$

because

$$\Psi^\dagger\gamma_0\Psi = \begin{pmatrix} (\chi_L)^\dagger & (\xi_R)^\dagger \end{pmatrix} \begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \underbrace{(\chi_L)^\dagger\bar{\sigma}_0\xi_R}_{=I_1} + \underbrace{(\xi_R)^\dagger\sigma_0\chi_L}_{=I_2}.$$

These are exactly the first two invariants we found earlier and[18]

[18] Remember that σ_0 is just the unit matrix and we have $\bar{\sigma}_0 = \sigma_0$.

$$\Psi^\dagger\gamma_0\gamma^\mu\partial_\mu\Psi = \begin{pmatrix} (\chi_L)^\dagger & (\xi_R)^\dagger \end{pmatrix} \begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^\mu\partial_\mu \\ \bar{\sigma}^\mu\partial_\mu & 0 \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}$$

$$= \underbrace{(\chi_L)^\dagger\bar{\sigma}_0\bar{\sigma}^\mu\partial_\mu\chi_L}_{=I_3} + \underbrace{(\xi_R)^\dagger\sigma_0\sigma^\mu\partial_\mu\xi_R}_{=I_4}$$

gives the other two invariants[19], as promised. To avoid writing γ_0 all the time it is conventional to introduce the notation

[19] Take note that $\sigma^\mu\partial_\mu = \partial_\mu\sigma^\mu$, because σ^μ are constant matrices.

$$\bar{\Psi} := (\Psi)^\dagger\gamma_0. \tag{6.15}$$

Now we have everything we need to construct a Lorentz-invariant Lagrangian using Dirac spinors that is in agreement with the restrictions[20] discussed in Section 4.2:

[20] Recall that these were: maximum order two in Ψ and only the lowest, non-trivial order in ∂_μ, which is here order 1.

$$\mathscr{L} = A\Psi^\dagger\gamma_0\Psi + B\Psi^\dagger\gamma_0\gamma^\mu\partial_\mu\Psi = A\bar{\Psi}\Psi + B\bar{\Psi}\gamma^\mu\partial_\mu\Psi.$$

Putting in the constants ($A = -m$, $B = i$) gives us the final **Dirac-Lagrangian**

$$\mathscr{L}_{\text{Dirac}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma^\mu\partial_\mu\Psi = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi. \tag{6.16}$$

Take note that what appears here in our Lagrangian are two distinct fields, because Ψ is complex[21]. This is a requirement, because otherwise we can't get something Lorentz invariant. More explicitly, we have

$$\Psi = \Psi_1 + i\Psi_2$$

with two real fields Ψ_1 and Ψ_2. Instead of working with two real fields it is conventional to work with two complex fields Ψ and $\bar{\Psi}$, as two distinct fields.

[21] Therefore the left-chiral and right-chiral spinors inside each Dirac spinor, are complex, too.

Now, if we put this Lagrangian into the Euler-Lagrange equation, which we recite here for convenience

$$\frac{\partial\mathscr{L}}{\partial\Psi} - \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\Psi)}\right) = 0$$

we get

$$-m\bar{\Psi} - i\partial_\mu\bar{\Psi}\gamma^\mu = 0 \to (i\partial^\mu\bar{\Psi}\gamma_\mu + m\bar{\Psi}) = 0\,. \tag{6.17}$$

With the Euler-Lagrange equation for the field $\bar{\Psi}$

$$\frac{\partial\mathscr{L}}{\partial\bar{\Psi}} - \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\bar{\Psi})}\right) = 0$$

we get the equation of motion for Ψ

$$(i\gamma_\mu\partial^\mu - m)\Psi = 0\,. \tag{6.18}$$

This is the famous **Dirac equation**, which is the equation of motion for spin $\frac{1}{2}$ particles and fields. Take note that this is exactly what we get if we integrate the Lagrangian by parts

$$-m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi \underbrace{=}_{\text{Integrate by parts. Just imagine here the action integral.}} -m\bar{\Psi}\Psi - (i\partial^\mu\bar{\Psi})\gamma_\mu\Psi$$

and then use the Euler-Lagrange equation,

$$\to -m\Psi + i\partial_\mu\gamma^\mu\Psi = 0\,.$$

So it really makes no difference and the way we wrote the Lagrangian is no restriction, despite its asymmetry[22].

[22] You are free to use the longer Lagrangian that includes both possibilities, but the results are the same.

6.4 Proca Equation

Now, we want to find the equation of motion for an object transforming according to the $(\frac{1}{2}, \frac{1}{2})$ representation of the Lorentz group. We already saw that this representation **is** the vector representation and therefore this task is easy. We simply take an arbitrary vector field A_μ and construct all possible Lorentz-invariants from it that are in agreement with the restrictions from Section 4.2. We must combine an upper with a lower index, because we defined the scalar product of Minkowski space in Section 2.4 this way and scalars are what we need in the Lagrangian. The possible invariants are[23]

[23] Again, a term of the form $\partial_\mu \partial^\mu A^\nu A_\nu$ is redundant, because we can integrate by parts and get $\partial^\mu A^\nu \partial_\mu A_\nu$.

$$I_1 = \partial^\mu A^\nu \partial_\mu A_\nu \quad , \quad I_2 = \partial^\mu A^\nu \partial_\nu A_\mu$$
$$I_3 = A^\mu A_\mu \quad , \quad I_4 = \partial^\mu A_\mu$$

and the Lagrangian reads

$$\mathscr{L}_{\text{Proka}} = C_1 \partial^\mu A^\nu \partial_\mu A_\nu + C_2 \partial^\mu A^\nu \partial_\nu A_\mu + C_3 A^\mu A_\mu + C_4 \partial^\mu A_\mu . \quad (6.19)$$

We can neglect the term $\partial^\mu A_\mu$, because has no influence on the equations of motion, as can be seen by looking at the Euler-Lagrange equation[24]. Therefore order 1 in ∂_μ is trivial.

[24] $\partial_\sigma \left(\frac{\partial (C_4 \partial^\mu A_\mu)}{\partial(\partial_\sigma A_\rho)} \right) = 0.$

If we now want to compute the equations of motion using the Euler-Lagrange equation for each field component independently

$$\frac{\partial \mathscr{L}}{\partial A_\rho} = \partial_\sigma \left(\frac{\partial \mathscr{L}}{\partial(\partial_\sigma A_\rho)} \right),$$

we need to be very careful about the indices. Let us take a look at the right-hand side of the Euler-Lagrange equation and pick the term involving C_1:

$$\begin{aligned}
&\partial_\sigma \left(\frac{\partial}{\partial(\partial_\sigma A_\rho)} (C_1 \partial^\mu A^\nu \partial_\mu A_\nu) \right) \\
&\underbrace{=}_{\text{product rule}} C_1 \partial_\sigma \left((\partial_\mu A_\nu) \frac{\partial(\partial^\mu A^\nu)}{\partial(\partial_\sigma A_\rho)} + (\partial^\mu A^\nu) \frac{\partial(\partial_\mu A_\nu)}{\partial(\partial_\sigma A_\rho)} \right) \\
&\underbrace{=}_{\text{lowering indices with the metric}} C_1 \partial_\sigma \left((\partial_\mu A_\nu) g^{\mu\kappa} g^{\nu\lambda} \frac{\partial(\partial_\kappa A_\lambda)}{\partial(\partial_\sigma A_\rho)} + (\partial^\mu A^\nu) \frac{\partial(\partial_\mu A_\nu)}{\partial(\partial_\sigma A_\rho)} \right) \\
&= C_1 \partial_\sigma \left((\partial_\mu A_\nu) g^{\mu\kappa} g^{\nu\lambda} \delta^\sigma_\kappa \delta^\rho_\lambda + (\partial^\mu A^\nu) \delta^\sigma_\mu \delta^\rho_\nu \right) \\
&= C_1 \partial_\sigma \left(\partial^\sigma A^\rho + \partial^\sigma A^\rho \right) \\
&= 2C_1 \partial_\sigma \partial^\sigma A^\rho . \qquad (6.20)
\end{aligned}$$

Following similar steps we can compute

$$\partial_\sigma \left(\frac{\partial}{\partial(\partial_\sigma A_\rho)} (C_2 \partial^\mu A^\nu \partial_\nu A_\mu) \right) = 2C_2 \partial^\rho (\partial_\sigma A^\sigma) .$$

Therefore the equation of motion, following from the Lagrangian in Eq. 6.19 reads

$$2C_3 A^\rho = 2C_1 \partial_\sigma \partial^\sigma A^\rho + 2C_2 \partial^\rho (\partial_\sigma A^\sigma) ,$$

which gives us, when we put in the conventional constants[25]

$$\rightarrow m^2 A^\rho = \partial_\sigma(\partial^\sigma A^\rho - \partial^\rho A^\sigma). \quad (6.21)$$

This is called the **Proca equation**, which is the equation of motion for massive spin 1 particles and fields. For massless ($m = 0$) spin 1 particles, e.g. photons, the equation reads

$$\rightarrow 0 = \partial_\sigma(\partial^\sigma A^\rho - \partial^\rho A^\sigma). \quad (6.22)$$

This is the **inhomogeneous Maxwell equation** in absence of electric currents. To unclutter the notation it is conventional to define the electromagnetic tensor

$$F^{\sigma\rho} := \partial^\sigma A^\rho - \partial^\rho A^\sigma. \quad (6.23)$$

Then the inhomogeneous Maxwell-equations read

$$\partial_\sigma F^{\sigma\rho} = 0 \quad (6.24)$$

and we can rewrite the Lagrangian for massless spin 1 fields

$$\mathscr{L}_{\text{Maxwell}} = \frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu)$$

as

$$\begin{aligned}
\mathscr{L}_{\text{Maxwell}} &= \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \\
&= \frac{1}{4}(\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) \\
&= \frac{1}{4}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu - \partial^\nu A^\mu \partial_\mu A_\nu + \partial^\nu A^\mu \partial_\nu A_\mu) \\
&\underbrace{=}_{\text{renaming dummy indices}} \frac{1}{4}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu - \partial^\mu A^\nu \partial_\nu A_\mu + \partial^\mu A^\nu \partial_\mu A_\nu) \\
&= \frac{1}{4}(2\partial^\mu A^\nu \partial_\mu A_\nu - 2\partial^\mu A^\nu \partial_\nu A_\mu) \\
&= \frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu) \qquad \checkmark.
\end{aligned} \quad (6.25)$$

$\mathscr{L}_{\text{Maxwell}} = \frac{1}{4} F^{\mu\nu} F_{\mu\nu}$ is the conventional way to write the Lagrangian. Equivalently, the Lagrangian for a massive spin 1 field can be written as

$$\mathscr{L}_{\text{Proca}} = \frac{1}{4} F^{\mu\nu} F_{\mu\nu} + m^2 A_\mu A^\mu = \frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu) + m^2 A_\mu A^\mu. \quad (6.26)$$

In this chapter we derived the equations of motion that describe free fields and particles. We want to understand what we can do with these equations in order to get predictions for experiments. However

[25] Maybe you wonder why we use $C_1 = -C_2$. The reason for this it that with $C_1 = -C_2$ the Lagrangian in Eq. 6.19 has a special internal symmetry. This symmetry will be crucial in our discussion of how different fields or particles interact with each other. This will be discussed in Section 7.1.2 ff.

first, we need to derive some more equations, because experiments always work through interactions. For example, we are only able to detect an electron if we use another particle, like a photon. Therefore, in the next chapter we will derive Lagrangians that describe the interaction between different fields and particles.