Elements of Classical Field Theory
C6, HT 2016

ULI HAISCH[a]

[a]*Rudolf Peierls Centre for Theoretical Physics*
University of Oxford
OX1 3PN Oxford, United Kingdom

Please send corrections to u.haisch1@physics.ox.ac.uk.

1 Classical Field Theory

In this part of the lecture we will discuss various aspects of classical fields. We will cover only the bare minimum ground necessary before turning to the quantum theory, and will return to classical field theory at several later stages in the course when we need to introduce new concepts or ideas.

1.1 Lorentz Group

The Lorentz group L is of fundamental importance for the construction of relativistic field theories, since L is associated to the symmetry of 4-dimensional space-time.

Using the *Minkowski metric* $\eta = \text{diag}\,(1, -1, -1, -1)$, the Lorentz group L consists of the real 4×4 matrices Λ that satisfy

$$\eta = \Lambda^T \eta \Lambda \,, \tag{1.1}$$

which written in components $(\mu, \nu, \rho, \sigma = 0, 1, 2, 3)$ takes the form

$$\eta^{\mu\nu} = \eta^{\sigma\rho} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \,. \tag{1.2}$$

It is readily seen that the transformations

$$x^\mu \to \Lambda^\mu{}_\nu x^\nu \,, \tag{1.3}$$

with Λ satisfying (1.1) leave the distance ds^2 invariant. Setting for simplicity the speed of light c to 1, one has

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \to \eta_{\mu\nu} \Lambda^\mu{}_\rho dx^\rho \Lambda^\nu{}_\sigma dx^\sigma = \eta_{\rho\sigma} dx^\rho dx^\sigma = ds^2 \,. \tag{1.4}$$

The *Lorentz transformations* (LTs) (1.3) are therefore consistent with the postulate of special relativity that tells us that the speed of light is the same in all inertial frames.

$\det(\Lambda)$	$\Lambda^0{}_0$	name	contains	given by
$+1$	≥ 1	$L_+^\uparrow$	$\mathbf{1}_4$	$L_+^\uparrow$
$+1$	≤ -1	$L_+^\downarrow$	PT	$PTL_+^\uparrow$
-1	≥ 1	$L_-^\uparrow$	P	$PL_+^\uparrow$
-1	≤ -1	$L_-^\downarrow$	T	$TL_+^\uparrow$

Figure 1.1: The four disconnected components of the Lorentz group.

Let us examine the consequences of (1.1). First, we take its determinant

$$\det(\eta) = \det\left(\Lambda^T\right)\det(\eta)\det(\Lambda)\,, \tag{1.5}$$

from which we deduce that

$$\det(\Lambda) = \pm 1\,. \tag{1.6}$$

The case of $\det(\Lambda) = +1$ (-1) corresponds to *proper* (*improper*) LTs and the associated subgroup is L_+ (L_-). This implies that *parity* or *space-inversion* $P = \mathrm{diag}\,(1,-1,-1,-1)$ as well as *time-reversal* $T = \mathrm{diag}\,(-1,1,1,1)$ are improper LTs and as such part of L_-. Second, we look at the component η^{00}, in which case one finds from (1.2) the relation

$$1 = \eta^{\sigma\rho}\Lambda^0{}_\rho\Lambda^0{}_\sigma = \left(\Lambda^0{}_0\right)^2 - \sum_{i=1,2,3}\left(\Lambda^i{}_0\right)^2\,. \tag{1.7}$$

It follows that

$$|\Lambda^0{}_0| \geq 1\,. \tag{1.8}$$

When $\Lambda^0{}_0 \geq 1$ the LT is said to be *orthochronous* and part of $L^\uparrow$, while $\Lambda^0{}_0 \leq -1$ gives a *non-orthochronous* LT which belongs to $L^\downarrow$. In consequence, the Lorentz group consists out of four classes of LTs as illustrated in Figure 1.1

Let us have a look at some simple example of LTs. A rotation by the angle θ about the z-axis and a boost by $v < 1$ along the x-axis

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.9}$$

with $\gamma = (1-v^2)^{-1/2}$ are part of the Lorentz group. In fact, any 3-dimensional rotation

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 \\ 0 & O \end{pmatrix}, \tag{1.10}$$

leaves ds^2 invariant, since one has $O^T O = OO^T = \mathbf{1}_3$ by definition, if O is an element of the 3-dimensional rotation group $SO(3)$.[1]

---

[1]The group of 3-dimensional orthogonal matrices is denoted by $O(3)$, while its subgroup of matrices with determinant $+1$ is called the special orthogonal group $SO(3)$.

(j_+, j_-)	dimension	name	symbol
$(0,0)$	1	scalar	ϕ
$(1/2, 0)$	2	left-handed Weyl spinor	χ_L
$(0, 1/2)$	2	right-handed Weyl spinor	χ_R
$(1/2, 0) \oplus (0, 1/2)$	4	Dirac spinor	ψ
$(1/2, 1/2)$	4	vector	A_μ

Figure 1.2: Low-dimensional representations of the Lorentz group.

Any LT can be decomposed as the product of a rotation, a boost, space-inversion P, and time-reversal T. Let us concentrate on the continuous transformations. Since there are three rotations and three boosts, one for each space direction, the continuous LTs are described in terms of six parameters. To find the corresponding six *generators*, *i.e.*, a basis of transformation matrices that describes infinitesimal rotations and boosts, we write

$$\Lambda = \mathbf{1}_4 + iT\,, \tag{1.11}$$

where T are purely imaginary 4×4 matrices. Inserting this linearized LT into the defining relation (1.1), implies

$$T = -\eta T^T \eta\,. \tag{1.12}$$

This tells us that the generators T must be anti-symmetric in the space-space components, but symmetric in the space-time components. The space of such matrices has indeed dimension six and is spanned by the set $(i = 1, 2, 3)$

$$J_i = \begin{pmatrix} 0 & 0 \\ 0 & T_i \end{pmatrix}\,,$$

$$K_1 = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\,, \quad K_2 = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\,, \quad K_3 = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}\,. \tag{1.13}$$

Here $(T_i)_{jk} = -i\epsilon_{ijk}$ with ϵ_{ijk} the fully anti-symmetric Levi-Civita tensor ($\epsilon_{123} = +1$) are the generators of $SO(3)$. It follows that the matrices J_i (K_i) generate rotations (boosts).

It is a matter of simply algebra to work out the commutation relations of the generators J_i and K_i of the Lorentz group. One obtains

$$[J_i, J_j] = i\epsilon_{ijk}J_k\,, \qquad [K_i, K_j] = -i\epsilon_{ijk}J_k\,, \qquad [J_i, K_j] = i\epsilon_{ijk}K_k\,, \tag{1.14}$$

where $[a, b] = ab - ba$ is the usual commutator. These relations look very similar to the commutation relations of *angular momentum* or *spin* that you should know from quantum mechanics (QM). To make the analogy between the Lorentz group and the Lie group $SU(2)$ even more explicit, we introduce the following new basis of generators

$$J_i^\pm = \frac{1}{2}\left(J_i \pm iK_i\right)\,. \tag{1.15}$$

3

In terms of these generators, the *Lorentz Lie algebra* (1.14) takes the form

$$[J_i^\pm, J_j^\pm] = i\epsilon_{ijk} J_k^\pm\,, \qquad [J_i^\pm, J_j^\mp] = 0\,. \tag{1.16}$$

This means that J_i^+ and J_i^- independently obey the Lie algebra of $SU(2)$. By analogy to the spin quantum number, we can therefore introduce a pair (j_+, j_-) of two spins that characterize the possible *representations* of the Lorentz group. The states within a representation are further distinguished by the eigenvalues of J_3^+ and J_3^-, which can take the values $m_+ = -j_+, -j_+ + 1, \ldots, j_+ - 1, j_+$ and $m_- = -j_-, -j_- + 1, \ldots, j_- - 1, j_-$. The *dimension, i.e.*, the number of distinct states in a given representation (j_+, j_-) is hence $(2j_+ + 1)(2j_- + 1)$.

A list of a few low-dimensional Lorentz-group representations is provided in Figure 1.2. Field theories in Minkowski space usually require Lorentz invariance and, hence, the Lorentz group is of fundamental importance for such theories. Since it is related to the symmetries of space-time it is often also referred as *external symmetry* of the theory. The classification of Lorentz group representations provides us with objects which transform in a definite way under LTs and thus these objects are the main building blocks of such field theories. In what follows we will not deal with the spinor fields $\chi_{L,R}$ and ψ that describe fermions (quarks and leptons). Our focus will be on scalar ϕ and vector A_μ fields, which have very simple transformation properties under the Lorentz group:

$$\phi \to \phi\,, \qquad A_\mu \to \Lambda_\mu{}^\nu A_\nu\,. \tag{1.17}$$

1.2   Dynamics of Fields

A *field* is a quantity defined at every space-time point $x = (t, \boldsymbol{x})$. While classical particle mechanics deals with a finite number of *generalized coordinates* $q_a(t)$, indexed by a label a, in field theory we are interested in the dynamics of fields

$$\phi_a(t, \boldsymbol{x})\,, \tag{1.18}$$

where both a and $\boldsymbol{x}$ are considered as labels. We are hence dealing with an infinite number of degrees of freedom (dofs), at least one for each point $\boldsymbol{x}$ in space. Notice that the concept of position has been relegated from a dynamical variable in particle mechanics to a mere label in field theory.

Lagrangian and Action

The dynamics of the fields is governed by the Lagrangian. In all the systems we will study in this course, the Lagrangian is a function of the fields ϕ_a and their derivatives $\partial_\mu \phi_a$,[2] and given by

$$L(t) = \int d^3x\, \mathcal{L}(\phi_a, \partial_\mu \phi_a)\,, \tag{1.19}$$

---

[2]If there is no (or only little) room for confusion, we will often drop the arguments of functions and write $\phi_a = \phi_a(x)$ *etc.* to keep the notation short.

where the official name for $\mathcal{L}$ is *Lagrangian density*. Like everybody else we will, however, simply call it Lagrangian from now on. For any time interval $t \in [t_1, t_2]$, the *action* corresponding to (1.19) reads

$$S = \int_{t_1}^{t_2} dt \int d^3x \, \mathcal{L} = \int d^4x \, \mathcal{L} \,. \tag{1.20}$$

Recall that in classical mechanics L depends only on q_a and $\dot{q}_a$, but not on the second time derivatives of the generalized coordinates. In field theory we similarly restrict to Lagrangians $\mathcal{L}$ depending on ϕ_a and $\dot{\phi}_a$. Furthermore, with an eye on *Lorentz invariance*, we will only consider Lagrangians depending on $\nabla \phi_a$ and not higher derivatives.

Notice that employing $c = \hbar = 1$, *i.e.*, working with *natural units*, the dimension of the action is $[S] = 0$. With (1.20) and $[d^4x] = -4$, it follows that the Lagrangian must necessarily have $[\mathcal{L}] = 4$. Other objects that we will use frequently to construct Lagrangians are derivatives, masses, couplings, and most importantly fields. The dimensions of the former two objects are $[\partial_\mu] = 1$ and $[m] = 1$, while the dimensions of the latter two quantities depend on the specific type of coupling and field one considers. We therefore postpone the discussion of the mass dimension of couplings and fields to the point when we meet the relevant building blocks.

Principle of Least Action

The dynamical behavior of fields can be determined by the *principle of least action*. This principle states that when a system evolves from one given configuration to another between times t_1 and t_2 it does so along the "path" in configuration space for which the action is an extremum (usually a minimum) and hence satisfies $\delta S = 0$. This condition can be rewritten, using partial integration, as follows

$$\begin{aligned}
\delta S &= \int d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi_a} \delta \phi_a + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \delta (\partial_\mu \phi_a) \right\} \\
&= \int d^4x \left\{ \left[\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) \right] \delta \phi_a + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \delta \phi_a \right) \right\} = 0 \,.
\end{aligned} \tag{1.21}$$

The last term is a total derivative and vanishes for any $\delta \phi_a$ that decays at spatial infinity and obeys $\delta \phi_a(t_1, \boldsymbol{x}) = \delta \phi_a(t_2, \boldsymbol{x}) = 0$. For all such paths, we obtain the Euler-Lagrange *equations of motion* (EOMs) for the fields ϕ_a, namely

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_a} = 0 \,. \tag{1.22}$$

Hamiltonian Formalism

The link between the Lagrangian formalism and the quantum theory goes via the *path integral*. While this is a powerful formalism, we will for the time being use *canonical quantization*, since it makes the transition to QM easier. For this we need the Hamiltonian formalism of field

theory. We start by defining the *momentum density* $\pi^a(x)$ conjugate to $\phi_a(x)$,

$$\pi^a = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a} \,. \tag{1.23}$$

In terms of π^a, $\dot{\phi}_a$, and $\mathcal{L}$ the *Hamiltonian density* is given by

$$\mathcal{H} = \pi^a \dot{\phi}_a - \mathcal{L} \,, \tag{1.24}$$

where, as in classical mechanics, we have eliminated $\dot{\phi}_a$ in favor of π^a everywhere in $\mathcal{H}$. The Hamiltonian then simply takes the form

$$H = \int d^3x \, \mathcal{H} \,. \tag{1.25}$$

1.3   Noether's Theorem

The role of symmetries in field theory is possibly even more important than in particle mechanics. There are Lorentz symmetry, internal symmetries, gauge symmetries, supersymmetries, *etc.* We start here by recasting Noether's theorem in a field theoretic framework.

Currents and Charges

Noether's theorem states that every *continuous symmetry* of the Lagrangian gives rise to a *conserved current* $J^\mu(x)$, so that the EOMs (1.22) imply

$$\partial_\mu J^\mu = 0 \,, \tag{1.26}$$

or in components $dJ^0/dt + \boldsymbol{\nabla} \cdot \boldsymbol{J} = 0$. To every conserved current there exists also a *conserved (global) charge* Q, *i.e.*, a physical quantity which stays the same value at all times, defined as

$$Q = \int_{\boldsymbol{R}^3} d^3x \, J^0 \,. \tag{1.27}$$

The latter statement is readily shown by taking the time derivative of Q,

$$\frac{dQ}{dt} = \int_{\boldsymbol{R}^3} d^3x \, \frac{dJ^0}{dt} = -\int_{\boldsymbol{R}^3} d^3x \, \boldsymbol{\nabla} \cdot \boldsymbol{J} \,, \tag{1.28}$$

which is zero, if one assumes that $\boldsymbol{J}$ falls off sufficiently fast as $|\boldsymbol{x}| \to \infty$. Notice, however, that the existence of the conserved current $\boldsymbol{J}$ is much stronger than the existence of the (global) charge Q, because it implies that charge is in fact conserved *locally*. To see this, we define the charge in a finite volume V by

$$Q_V = \int_V d^3x \, J^0 \,. \tag{1.29}$$

Repeating the above analysis, we find

$$\frac{dQ_V}{dt} = -\int_V d^3x \, \boldsymbol{\nabla} \cdot \boldsymbol{J} = -\oint_S d\boldsymbol{S} \cdot \boldsymbol{J} \,, \tag{1.30}$$

6

where S denotes the area bounding V, $d\boldsymbol{S}$ is a shorthand for $\boldsymbol{n}\,dS$ with $\boldsymbol{n}$ being the outward pointing unit normal vector of the boundary S, and we have used Gauss' theorem. In physical terms the result means that any charge leaving V must be accounted for by a flow of the current 3-vector $\boldsymbol{J}$ out of the volume. This kind of local conservation law of charge holds in any local field theory.

Proof of Theorem

In order to prove Noether's theorem, we'll consider *infinitesimal* transformations. This is always possible in the case of a continuous symmetry. We say that $\delta\phi_a$ is a symmetry of the theory, if the Lagrangian changes by a total derivative

$$\delta\mathcal{L}(\phi_a) = \partial_\mu \mathcal{J}^\mu(\phi_a)\,, \tag{1.31}$$

for a set of functions $\mathcal{J}^\mu$. We then consider the transformation of $\mathcal{L}$ under an *arbitrary* change of field $\delta\phi_a$. Glancing at (1.21) tells us that in this case

$$\delta\mathcal{L} = \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\right]\delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\delta\phi_a\right)\,. \tag{1.32}$$

When the EOMs are satisfied than the term in square bracket vanishes so that we are simply left with the total derivative term. For a symmetry transformation satisfying (1.30) and (1.31), the relation (1.32) hence takes the form

$$\partial_\mu\mathcal{J}^\mu = \delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\delta\phi_a\right)\,, \tag{1.33}$$

or simply $\partial_\mu J^\mu = 0$ with

$$J^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\delta\phi_a - \mathcal{J}^\mu\,, \tag{1.34}$$

which completes the proof. Notice that if the Lagrangian is *invariant* under the infinitesimal transformation $\delta\phi_a$, *i.e.*, $\delta\mathcal{L} = 0$, then $\mathcal{J}^\mu = 0$ and J^μ contains only the first term on the right-hand side of (1.34).

We stress that that our proof only goes through for continuous transformations for which there exists a choice of the transformation parameters resulting in a unit transformation, *i.e.*, no transformation. An example is a Lorentz boost with some velocity $\boldsymbol{v}$, where for $\boldsymbol{v} = 0$ the coordinates x remain unchanged. There are examples of symmetry transformations where this does not occur. *E.g.*, a parity transformation P does not have this property, and Noether's theorem is not applicable then.

Energy-Momentum Tensor

Recall that in classic particle mechanics, *spatial translation* invariance gives rise to the conservation of momentum, while invariance under *time translations* is responsible for the conservation of energy. What happens in classical field theory? To figure it out, let's have a look at infinitesimal translations

$$x^\nu \to x^\nu - \epsilon^\nu \implies \phi_a(x) \to \phi_a(x+\epsilon) = \phi_a(x) + \epsilon^\nu\partial_\nu\phi_a(x)\,, \tag{1.35}$$

where the sign in the field transformation is plus, instead of minus, because we are doing an active, as opposed to passive, transformation. If the Lagrangian does not explicitly depend on x but only through $\phi_a(x)$ (which will always be the case in the Lagrangians discussed in the course), the Lagrangian transforms under the infinitesimal translation as

$$\mathcal{L} \to \mathcal{L} + \epsilon^\nu \partial_\nu \mathcal{L} \,. \tag{1.36}$$

Since the change in $\mathcal{L}$ is a total derivative, we can invoke Noether's theorem which gives us four conserved currents $T^\mu{}_\nu = (J^\mu)_\nu$ one for each of the translations ϵ^ν ($\nu = 0, 1, 2, 3$). From (1.34) and (1.35) we readily read off the explicit expressions for $T^\mu{}_\nu$,

$$T^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \, \partial_\nu \phi_a - \delta^\mu{}_\nu \mathcal{L} \,. \tag{1.37}$$

This quantity is called the *energy-momentum* (or *stress-energy*) *tensor*. It has dimension $[T^\mu{}_\nu] = 4$ and satisfies

$$\partial_\mu T^\mu{}_\nu = 0 \,. \tag{1.38}$$

The four "conserved charges" are ($\mu = 0, 1, 2, 3$)

$$P^\mu = \int d^3x \, T^{0\mu} \,, \tag{1.39}$$

Specifically, the "time component" of P^μ is

$$P^0 = \int d^3x \, T^{00} = \int d^3x \left(\pi^a \dot{\phi}_a - \mathcal{L} \right) \,, \tag{1.40}$$

which (looking at (1.24) and (1.25)) is nothing but the Hamiltonian H. We thus conclude that the charge P^0 is the total energy of the field configuration, and it is conserved. In fields theory, energy conservation is thus a pure consequence of time translation symmetry, like it was in particle mechanics. Similarly, we can identify the charges P^i ($i = 1, 2, 3$),

$$P^i = \int d^3x \, T^{0i} = -\int d^3x \, \pi^a \partial_i \phi_a \,, \tag{1.41}$$

as the momentum components of the field configuration in the three space directions, and they are of course also conserved.

1.4 Scalar Field Theories

In the following we will apply the formalism described above to the simplest class of Lorentz invariant field theories, involving only fields that transform trivially under the Lorentz group. We start by considering the case of a single real scalar field.

A Single Real Scalar Field

The Lagrangian density of our real scalar field ϕ is given by

$$\mathcal{L} = \frac{1}{2} \left(\partial_\mu \phi \right)^2 - V(\phi) \,, \tag{1.42}$$

where the first term is referred to as *kinetic energy* (or *kinetic term*), while the second is known as the *scalar potential*. Notice that ϕ corresponds to the $(j_+, j_-) = (0, 0)$ representation of the Lorentz group. The dimension of the real scalar field is obviously $[\phi] = 1$.

The scalar potential can be written as

$$V = \frac{1}{2} m^2 \phi^2 + \sum_{n=3}^{\infty} \frac{\lambda_n}{n!} \phi^n \,. \tag{1.43}$$

Here m is the mass of ϕ (this will become clear later on) and the coefficients λ_n are called *coupling constants*. Dimensional analysis tells us that

$$[m] = 1 \,, \qquad [\lambda_n] = 4 - n \,. \tag{1.44}$$

The coupling terms in (1.43) fall into three different categories. First, dimension-three terms with $[\lambda_3] = 1$. For such terms, we can define a dimensionless parameter λ_3/E, where E has dimension of mass and represents the energy scale of the process of interest. This means that $\Delta \mathcal{L}_3 = \lambda_3 \phi^3/(3!)$ is a small perturbation for high energies, *i.e.*, $E \gg \lambda_3$, but a big one at low energies, *i.e.*, $E \ll \lambda_3$. Such terms are called *relevant*, because they become and are most relevant at low energies which, after all, is where most of the physics that we experience lies. In a relativistic quantum field theory (QFT), we have $E > m$, which means that we can always make this sort of perturbations small by taking $\lambda_3 \ll m$. Second, terms of dimension four with $[\lambda_4] = 0$. *E.g.*, $\Delta \mathcal{L}_4 = \lambda_4 \phi^4/(4!)$. Such terms are small if $\lambda_4 \ll 1$ and are called *marginal*. Third, operators with dimension of higher than four, having $[\lambda_n] < 0$. In this case the appropriate dimensionless parameters is $(\lambda_n E^{n-4})$ and terms $\Delta \mathcal{L}_n = \lambda_n \phi^n/(n!)$ with $n \geq 5$ are small (large) at low (high) energies. Such contributions are called *irrelevant*, since in daily life, meaning $E^{n-4} \ll \lambda_n$, these terms do not matter.

As we will see later, it is typically impossible to avoid high-energy processes in QFT. We hence might expect problems with irrelevant terms (or operators) that become important at high energies. Indeed, these operators lead to *non-renormalizable* QFTs in which one cannot make sense of the infinities at arbitrarily high energies. This does not mean that these theories are useless, it just means that they become incomplete at some energy scale and need to be embedded into an appropriate complete theory *aka* an ultraviolet (UV) completion. Let me also add that the above naive assignment of relevant, marginal, and irrelevant operators is not always carved in stone, since quantum corrections can sometimes change the character of an operator.

Low-Energy Description

In typical applications of QFT only the relevant and marginal couplings are important. This is due to the fact that the irrelevant couplings become small at low energies, as we have seen

above. In practice this saves us, since instead of considering the infinite number of coupling terms in (1.43), only a handful are actually needed. *E.g.*, in the case of the real scalar field ϕ described earlier, we only have to take into account two operators, namely $\Delta\mathcal{L}_3 = \lambda_3\,\phi^3/(3!)$ and $\Delta\mathcal{L}_4 = \lambda_4\,\phi^4/(4!)$, in the low-energy limit.

Let us have a closer look at this issue. Suppose that at some day we discover the true superduper theory *aka* the TOE that describes the world at very high energy scales. Whatever this scale is, let's call it Λ. Since it is an energy scale, we obviously have $[\Lambda] = 1$. What we want to understand are the laws of physics at energy scales E that we can probe directly in a laboratory, which given today's standards, means $E \ll \Lambda$. Let us further suppose that at energies of order E, the laws of physics are described by a real scalar field.[3] This scalar field will have some complicated coupling terms (1.43), where the precise form is dictated by all the stuff that is going on in the TOE. Can we get an idea about the interactions? Well, we can write our dimensionful coupling constants λ_n in terms of dimensionless couplings g_n, multiplied by a suitable power of the relevant scale Λ,

$$\lambda_n = \frac{g_n}{\Lambda^{n-4}}\,. \tag{1.45}$$

The exact values of the dimensionless couplings g_n depend on the details of the TOE,[4] so we have to do some guesswork. Since the couplings g_n are dimensionless, 1 looks like a pretty good and somehow a natural guess. Since we are not completely sure, let's say $g_n = \mathcal{O}(1)$. This means that in a laboratory with $E \ll \Lambda$ the coupling terms $\Delta\mathcal{L}_n = \lambda_n\phi^n/(n!)$ of (1.43) will be suppressed by powers of $(E/\Lambda)^{n-4}$ if $n \geq 5$. Given the CERN Large Hadron Collider (LHC) energy of around $1\,\text{TeV}$, this is a suppression by many orders of magnitude. *E.g.*, for $\Lambda = M_{\text{P}} = 10^{16}\,\text{TeV}$ corresponding to the *Planck mass* M_P, one has $E/\Lambda = 10^{-16}$. It is this simple argument based on dimensional analysis that ensures that we need to focus only on the first few terms in the interactions, namely those that are relevant and marginal. It also means that if we only have access to low-energy experiments, it is going to be very difficult to figure out the precise nature of the TOE, because its effects are highly diluted except for the relevant and marginal interactions. Some people therefore call the superduper theory that everybody is looking for, not TOE, but TOENAIL, which stands for "theory of everything not accessible in laboratories".

Klein-Gordon Equation

Applying the Euler-Lagrange equations to (1.42) leads to

$$\partial^\mu\partial_\mu\phi + \frac{\partial V}{\partial\phi} = 0\,. \tag{1.46}$$

The Laplacian in Minkowski space is sometimes denoted by $\Box$ and $\partial V/\partial\phi$ is commonly written as V'. In this notation, the above equation reads $\Box\phi + V' = 0$.

---

[3]Of course, we know that this assumption is plain wrong, since the standard model (SM) of particle physics is a non-abelian gauge theory with chiral fermions, but the same argument applies in that case.

[4]If we would know the precise structure of the TOE we could, in fact, calculate the couplings g_n.

Since for non-vanishing coefficients λ_n the solutions to (1.46) are difficult to find, let's simply ignore them for the time being and only consider the free case. The dynamics of such a *non-interacting* massive real scalar field is encoded in the famous *Klein-Gordon equation*:

$$\left(\Box + m^2\right)\phi = 0\,. \tag{1.47}$$

The corresponding Lagrangian is of course

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\phi\right)^2 - \frac{1}{2}m^2\phi^2\,. \tag{1.48}$$

In order to find the solutions to the classical Klein-Gordon equation (1.48), we only have to Fourier transform the field ϕ,

$$\phi(t,\boldsymbol{x}) = \int \frac{d^3p}{(2\pi)^3}\,e^{i\,\boldsymbol{p}\cdot\boldsymbol{x}}\,\phi(t,\boldsymbol{p})\,. \tag{1.49}$$

In momentum space (1.48) simply reads

$$\left[\frac{\partial^2}{\partial t^2} + \left(\boldsymbol{p}^2 + m^2\right)\right]\phi(t,\boldsymbol{p}) = 0\,, \tag{1.50}$$

which tells us that for each value of $\boldsymbol{p}$, the Fourier transform $\phi(t,\boldsymbol{p})$ solves the equation of a *harmonic oscillator* with frequency

$$\omega_{\boldsymbol{p}} = \sqrt{|\boldsymbol{p}|^2 + m^2}\,. \tag{1.51}$$

It is then not difficult to see that the most general solution of the classical Klein-Gordon equation is a linear superposition of simple harmonic oscillators,[5]

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\boldsymbol{p}}}\left(a(p)e^{-ipx} + a^*(p)e^{ipx}\right)\,, \tag{1.52}$$

each vibrating at a different frequency with a different amplitude (the appearance of the specific combination of coefficients $a(p)$ and $a^*(p)$ is dedicated by the reality of the Klein-Gordon field ϕ). The 4-vector p in the exponents is understood as $p_\mu = (w_{\boldsymbol{p}},\boldsymbol{p})$. While the decomposition (1.52) has not actually been derived here, the result looks plausible after noticing that the measure $d^3p/(2\pi)^3\,1/(2\omega_{\boldsymbol{p}})$ is a Lorentz-invariant quantity. This follows from "reverse engineering"

$$\int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\boldsymbol{p}}} = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2p_0}\bigg|_{p_0=\omega_{\boldsymbol{p}}} = \int \frac{d^4p}{(2\pi)^3}\,\delta(p_0^2 - \boldsymbol{p}^2 - m^2)\,\big|_{p_0>0}\,. \tag{1.53}$$

From the latter result we can also figure out that the Lorentz-invariant delta function for 3-momenta is

$$2\omega_{\boldsymbol{p}}\,\delta^{(3)}(\boldsymbol{p} - \boldsymbol{q})\,, \tag{1.54}$$

---

[5]The actual derivation of this result can be found on page 47 of the script by John Chalker and Andre Lukas. It is not repeated here.

since

$$\int \frac{d^3p}{2\omega_{\boldsymbol{p}}} 2\omega_{\boldsymbol{p}}\,\delta^{(3)}(\boldsymbol{p}-\boldsymbol{q}) = 1\,. \tag{1.55}$$

The Lorentz invariance of the measure $d^3p/(2\pi)^3\,1/(2\omega_{\boldsymbol{p}})$ and its consequences, should be kept in mind, because these features will be important at several occasions later on.

Let me add that in order to quantize the field ϕ, we must hence only quantize a infinite number of harmonic oscillators (as Sidney Coleman once said: "The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction."). You should already know how this is done in QM and we will see later on how it is done in QFT.

Lorentz Invariance

Let us first check that a LT

$$\phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x)\,, \tag{1.56}$$

leaves the Lagrangian (1.48) and its action invariant. Notice that the inverse Λ^{-1} appears here because we are dealing with an active transformation, in which the field is truly shifted. To see why this means that the inverse appears, it will suffice to consider a non-relativistic example such as a temperature field. Suppose we start with an initial field $\phi(\boldsymbol{x})$ which has a hotspot at, say, $\boldsymbol{x} = (1,0,0)$. Let's now make a rotation $\boldsymbol{x} \to O\boldsymbol{x}$ about the z-axis so that the hotspot ends up at $\boldsymbol{x} = (0,1,0)$. If we want to express the new field $\phi'(\boldsymbol{x})$ in terms of the old field $\phi(\boldsymbol{x})$, we have to place ourselves at $\boldsymbol{x} = (0,1,0)$ and ask what the old field looked like at the point $O^{-1}\boldsymbol{x} = (1,0,0)$ we came from. This O^{-1} is the origin of the Λ^{-1} factor in the argument of the transformed field in (1.56).

According to the LT (1.56), the transformation of the mass term is $1/2\,m^2\phi^2(x) \to 1/2\,m^2\phi^2(x')$ with $x' = \Lambda^{-1}x$. The transformation of $\partial_\mu\phi$ is

$$\partial_\mu\phi(x) \to \partial_\mu(\phi(x')) = (\Lambda^{-1})^\nu{}_\mu (\partial_\nu\phi)(x')\,. \tag{1.57}$$

Here we have used that

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial(x')^\nu}\frac{\partial(x')^\nu}{\partial x^\mu} = (\Lambda^{-1})^\nu{}_\mu\,\partial'_\nu\,. \tag{1.58}$$

and wrote $\partial'_\mu\phi(x') = (\partial_\mu\phi)(x')$. Below we will also use the notation $\partial_\mu\phi(x')$ to denote the derivative at $x = x'$. Using the defining property $(\Lambda^{-1})^\rho{}_\mu(\Lambda^{-1})^\sigma{}_\nu\,\eta^{\mu\nu} = \eta^{\rho\sigma}$ of the LTs, we thus find that the derivative term in the Klein-Gordon Lagrangian behaves as

$$\frac{1}{2}\left(\partial_\mu\phi(x)\right)^2 \to \frac{1}{2}(\Lambda^{-1})^\rho{}_\mu(\partial_\rho\phi)(x')(\Lambda^{-1})^\sigma{}_\nu(\partial_\sigma\phi)(x')\,\eta^{\mu\nu}$$

$$= \frac{1}{2}(\partial_\rho\phi)(x')(\partial_\sigma\phi)(x')\,\eta^{\rho\sigma} \tag{1.59}$$

$$= \frac{1}{2}\left(\partial_\mu\phi(x')\right)^2\,.$$

Putting things together, we see that the action of the Klein-Gordon theory is indeed Lorentz invariant,

$$S = \int d^4x \, \mathcal{L}(x) \to \int d^4x \, \mathcal{L}(x') = \int d^4x' \, \mathcal{L}(x') = S \,. \tag{1.60}$$

Notice that changing the integration variables from d^4x to d^4x', in principle introduces an Jacobian factor $\det(\Lambda)$. This factor is, however, equal to 1 for LT connected to the identity, that we are dealing with (remember that $\mathbf{1}_4$ is part of the *proper, orthochronous* or *restricted* Lorentz group $L_+^\uparrow$).

A similar calculation also shows that, as promised, also EOM of the Klein-Gordon field ϕ, as given in (1.47), is invariant,

$$\left(\partial^2 + m^2\right)\phi(x) \to \left(\partial^2 + m^2\right)\phi'(x)$$
$$= \left[(\Lambda^{-1})^\nu{}_\mu \partial_\nu (\Lambda^{-1})^{\rho\mu}\partial_\rho + m^2\right]\phi(x') \tag{1.61}$$
$$= \left(\eta^{\nu\rho}\partial_\nu\partial_\rho + m^2\right)\phi(x') = 0 \,.$$

Here we have used another common notation for the Laplacian, $\partial^\mu\partial_\mu = \partial^2$.

Notice that the above explicit example shows that the Lagrangian formulation of field theory makes it especially easy to discuss Lorentz invariance, since an EOM is automatically Lorentz invariant if it follows from a Lagrangian that is a *Lorentz scalar*. This is an immediate consequence of the principle of least action. If a LT leaves the Lagrangian unchanged, the transformation of an extremum in the action will be another extremum.

Conservation Laws

Let us derive the Hamiltonian density and the energy-momentum tensor for (1.42). The conjugate momentum of ϕ is simply $\pi = \dot{\phi}$, and therefore

$$\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V \,. \tag{1.62}$$

For the stress tensor we find from (1.37) the expression

$$T^{\mu\nu} = (\partial^\mu\phi)(\partial^\nu\phi) - \frac{1}{2}\eta^{\mu\nu}(\partial_\rho\phi)^2 + \eta^{\mu\nu}V \,. \tag{1.63}$$

Notice that the energy-momentum tensor of (1.42) immediately comes out symmetric, *i.e.*, $T_{\mu\nu} = T_{\nu\mu}$. This is not always the case. In accordance with our general formula (1.40), we hence have

$$P^0 = \int d^3x \, T^{00} = \int d^3x \left(\frac{1}{2}\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V\right) = \int d^3x \, \mathcal{H} = H \,. \tag{1.64}$$

The energy is hence conserved in our real scalar theory. Analog relations hold for the space components P^i, signalling 3-momentum conservation.

In classical particle mechanics, *rotational invariance* gives rise to conservation of angular momentum. What is the analogy in field theory? What happens the remaining three LTs,

13

namely the *boosts*. What conserved quantity do they correspond to? In order to address these questions, we first need the infinitesimal form of the LTs

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu \,, \tag{1.65}$$

where $\omega^\mu{}_\nu$ is infinitesimal. The condition (1.2) for Λ to be a LT becomes in infinitesimal form

$$\eta^{\mu\nu} = \eta^{\rho\sigma} \left(\delta^\mu{}_\rho + \omega^\mu{}_\rho \right) \left(\delta^\nu{}_\sigma + \omega^\nu{}_\sigma \right) = \eta^{\mu\nu} + \omega^{\mu\nu} + \omega^{\nu\mu} + \mathcal{O}(\omega^2) \,. \tag{1.66}$$

This shows that $\omega^{\mu\nu}$ must be an anti-symmetric matrix (a feature that I have stated before without proof),

$$\omega^{\mu\nu} = -\omega^{\nu\mu} \,. \tag{1.67}$$

Notice that an anti-symmetric 4×4 matrix has six independent parameters, which agrees with the number of different Lorentz transformations, *i.e.*, three rotations and three boosts.

Applying the infinitesimal LT to our real scalar field ϕ, we have

$$\phi(x) \to \phi(x - \omega x) = \phi(x) - \omega^\mu{}_\nu x^\nu \partial_\mu \phi(x) \,, \tag{1.68}$$

where the minus sign arises from the factor Λ^{-1} in (1.56). The variation of the field ϕ under an infinitesimal LT is hence given by

$$\delta\phi = -\omega^\mu{}_\nu x^\nu \partial_\mu \phi \,. \tag{1.69}$$

By the same line of reasoning, one shows that the variation of the Lagrangian is

$$\delta\mathcal{L} = -\omega^\mu{}_\nu x^\nu \partial_\mu \mathcal{L} = -\partial_\mu \left(\omega^\mu{}_\nu x^\nu \mathcal{L} \right) \,, \tag{1.70}$$

where in the last step we used the fact that $\omega^\mu{}_\mu = 0$ due to its anti-symmetry. The Lagrangian changes by a total derivative, so we can apply Noether's theorem (1.34) with $\mathcal{J}^\mu = -\omega^\mu{}_\nu x^\nu \mathcal{L}$ to find the conserved current,

$$\begin{aligned} J^\mu &= -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \omega^\rho{}_\nu x^\nu \partial_\rho\phi + \omega^\mu{}_\nu x^\nu \mathcal{L} \\ &= -\omega^\rho{}_\nu \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \partial_\rho\phi - \delta^\mu{}_\rho \mathcal{L} \right] x^\nu = -\omega^\rho{}_\nu T^\mu{}_\rho x^\nu \,. \end{aligned} \tag{1.71}$$

Stripping off $\omega^\rho{}_\nu$, we obtain six different currents, which we write as

$$(\mathcal{I}^\lambda)^{\mu\nu} = x^\mu T^{\lambda\nu} - x^\nu T^{\lambda\mu} \,. \tag{1.72}$$

These currents satisfy

$$\partial_\lambda (\mathcal{I}^\lambda)^{\mu\nu} = 0 \,, \tag{1.73}$$

and give (as usual) rise to six conserved charges . For $\mu, \nu \neq 0$, the LT is a rotation and the three conserved charges give the *total angular momentum* of the field ($i, j = 1, 2, 3$):

$$Q^{ij} = \int d^3x \left(x^i T^{0j} - x^j T^{0i} \right) \,. \tag{1.74}$$

What's about the boosts? In this case, the conserved charges are

$$Q^{0i} = \int d^3x \left(x^0 T^{0i} - x^i T^{00} \right) .$$ (1.75)

The fact that these are conserved tells us that

$$0 = \frac{dQ^{0i}}{dt} = \int d^3x \, T^{0i} + t \int d^3x \, \frac{dT^{0i}}{dt} - \frac{d}{dt} \int d^3x \, x^i T^{00}$$

$$= P^i + t \frac{dP^i}{dt} - \frac{d}{dt} \int d^3x \, x^i T^{00} .$$ (1.76)

Yet, also the momentum P^i is conserved, *i.e.*, $dP^i/dt = 0$, and we conclude that

$$\frac{d}{dt} \int d^3x \, x^i T^{00} = \frac{d}{dt} \int d^3x \, x^i \mathcal{H} = \text{const.}$$ (1.77)

This is the statement that the center of energy of the field travels with a constant velocity. In a sense it's a field theoretical version of Newton's first law but, rather surprisingly, appearing here as a conservation law. Notice that after restoring the label a our results for $(\mathcal{I}^\lambda)^{\mu\nu}$ *etc.* also apply in the case of multicomponent fields.

Poincaré Invariance

We now require that a physical system possesses both space-time translation (1.35) and LT symmetry. The symmetry group that includes both transformations is called the *Poincaré group*. Notice that for any Poincaré-invariant theory the two charge conservation equations (1.38) and (1.73) should hold. This is only possible if the energy-momentum tensor $T^{\mu\nu}$ is symmetric. Indeed,

$$0 = \partial_\lambda (\mathcal{I}^\lambda)^{\mu\nu} = \partial_\lambda \left(x^\mu T^{\lambda\nu} - x^\nu T^{\lambda\mu} \right)$$

$$= x^\mu \partial_\lambda T^{\lambda\nu} + T^{\lambda\nu} \partial_\lambda x^\mu - x^\nu \partial_\lambda T^{\lambda\mu} - T^{\lambda\mu} \partial_\lambda x^\nu$$ (1.78)

$$= T^{\lambda\nu} \delta_\lambda{}^\mu - T^{\lambda\mu} \delta_\lambda{}^\nu = T^{\mu\nu} - T^{\nu\mu} .$$

This general result tells us that the expression of the energy-momentum tensor in any theory can be made symmetric without changing physics. The key to actually do it, lies in making use of the conservation law (1.38) in an appropriate way.

Discrete Internal Symmetries

So far we have only imposed *external* symmetries, such as Lorentz or Poincaré symmetry, on our scalar theory. Let us also have a brief look at *internal* symmetries. An interesting internal symmetry to be considered is a $\mathbb{Z}_2$ symmetry which acts as

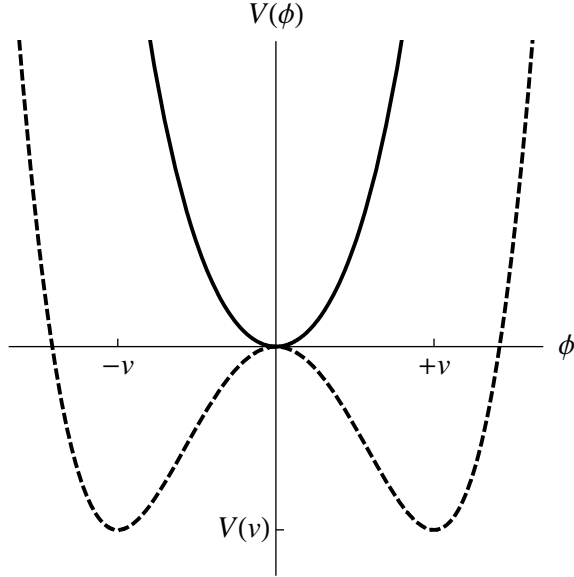$$\phi(x) \to -\phi(x) .$$ (1.79)

Figure 1.3: Shape of the scalar potential for $m^2 \geq 0$ (solid line) and $m^2 < 0$ (dashed line). In the latter case the positions of the minima are at $\phi = \pm v = \pm\sqrt{(-6m^2)/\lambda}$. The value of the potential at the minima is $V(v) = V_0 - (\lambda v^4)/24$.

This transformation leaves (1.42) invariant if the scalar potential V contains only terms with an even number of ϕ fields (but it does not lead to a conserved quantity since it is not a continuous symmetry). Restricting ourselves to relevant and marginal couplings, we write

$$V = V_0 + \frac{1}{2}\,m^2\phi^2 + \frac{\lambda}{4!}\,\phi^4\,. \tag{1.80}$$

where we have added a constant term V_0 to the potential. Notice that such a shift leaves the EOM (1.46) unaltered.

Let us now study the simplest type of solutions to the real scalar theory with the potential (1.80). A trivial solution of (1.46) is $\phi(x) = v = \text{const.}$, provided that

$$\left.\frac{\partial V}{\partial \phi}\right|_{\phi=v} = 0\,. \tag{1.81}$$

This relation implies that we should look for the extrema of the potential (1.80). In fact, to minimize the energy (1.64) we should focus on the minima of V. Such solutions of the classic theory are called *vacua* (or *vacuum* if there is a single one). Note that if the quartic coupling λ is negative, the potential is not bounded from below and the energy of a constant field configuration tends to minus infinity for large field values. To avoid such an unphysical situation, we simply assume $\lambda > 0$ in what follows.

The shape of the scalar potential V is shown in Figure 1.3. We see that one has to distinguish two cases. First, the case where $m^2 \geq 0$ (solid curve). In this situation there is a single minimum located at the origin $\phi = 0$. This solution is mapped into itself under the transformation (1.79) and we say that the symmetry is *unbroken* in this vacuum. The second

16

case with $m^2 < 0$ (dashed curve) is more interesting. In this case $\phi = 0$ is a maximum, while there are two minima. These are located at

$$\phi = \pm v = \pm \sqrt{\frac{-6m^2}{\lambda}} \, . \tag{1.82}$$

At the minima, the potential takes the value

$$V(v) = V_0 + \frac{1}{4} m^2 v^2 = V_0 - \frac{\lambda v^4}{24} \, . \tag{1.83}$$

Notice that neither minimum is invariant under the $\mathbb{Z}_2$ symmetry. In fact, the minima are interchanged under (1.79), and we say that the symmetry is *spontaneously broken.* In general, spontaneous symmetry breaking refers to the case where a symmetry of the theory is partially or fully broken by the vacuum solution of the theory. We will see later on that the concept of spontaneous symmetry breaking is a rather important one.

Continous Internal Symmetries

Besides discrete also continuous internal symmetries play an important role in particle physics. We start our discussion by considering a scalar theory that involves two real scalar fields $\phi_{1,2}$ and combine these fields into a field vector

$$\boldsymbol{\phi} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \, . \tag{1.84}$$

In the following we will be interested in theories that are invariant under $SO(2)$ transformations (*aka* 2-dimensional rotations) that take the form

$$\boldsymbol{\phi} \to \boldsymbol{\phi}' = O\boldsymbol{\phi} \, . \tag{1.85}$$

Since this transformation is the same at every space-time point, symmetries of this type are called *global.* The length of $\boldsymbol{\phi}$ should be $SO(2)$ invariant, which implies that

$$O^T O = OO^T = \mathbf{1}_2 \, , \tag{1.86}$$

and we also require that

$$\det(O) = +1 \, . \tag{1.87}$$

Proper 2-dimensional rotation matrices can hence be written as

$$O = e^{i\alpha T} \, , \tag{1.88}$$

where α is a real parameter (the rotation angle) and T denotes the 2×2 matrix that generates the 2-dimensional rotations. A suitable choice of the *generator* reads

$$T = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \, . \tag{1.89}$$

17

Notice that $SO(2)$ is said to be an *abelian* group, since

$$e^{i\alpha T}e^{i\beta T} = e^{i\beta T}e^{i\alpha T} = e^{i(\alpha+\beta)T}\,, \tag{1.90}$$

which means that the ordering of two successive group transformations does not matter.

In the case of 3-dimensional rotations things become slightly more complicated. Since there is an independent rotation corresponding to each plane in three dimensions, the $SO(3)$ group has $3\,(3-1)\,/2 = 3$ parameters α^i (*Euler angles*) and three generators T_i $(i=1,2,3)$. Infinitesimal $SO(3)$ transformations thus act on $\boldsymbol{\phi} = (\phi_1,\phi_2,\phi_3)^T$ as $\boldsymbol{\phi} \to \boldsymbol{\phi}' = \boldsymbol{\phi} + i\alpha^i T_i\,\boldsymbol{\phi}$. From (1.86) and (1.87) it follows that the generator T_i are hermitian and traceless:

$$(T_i)^\dagger = T_i\,, \qquad \mathrm{Tr}\,(T_i) = 0\,. \tag{1.91}$$

A possible set of 3×3 matrices with these properties is

$$T_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}\,, \qquad T_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}\,, \qquad T_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,, \tag{1.92}$$

which can be written in a more compact way as $(T_i)_{jk} = -i\epsilon_{ijk}$ $(j,k=1,2,3)$. By a straight-forward calculation, one can prove that the latter matrices satisfy the following *Lie algebra*

$$[T_i, T_j] = i\epsilon_{ijk}T_k\,. \tag{1.93}$$

These relations imply that two 3-dimensional rotations do in general not "commute", which causes the *non-abelian* character of $SO(3)$.

Other examples of non-abelian groups are provided by the special unitary groups $SU(n)$, which consist of all complex $n \times n$ matrices U satisfying

$$U^\dagger U = U U^\dagger = \mathbf{1}_n\,, \qquad \det\,(U) = +1\,. \tag{1.94}$$

The simplest non-trivial example of a special unitary group is $SU(2)$. It has $2^2 - 1 = 3$ generators τ_i and close to the identity one can write the group transformations as $U = \mathbf{1}_2 + it^i\tau_i$. From (1.94) it follows that the generators obey $\tau_i^\dagger = \tau_i$ and $\mathrm{Tr}\,(\tau_i) = 0$. A convenient choice of the generators τ_i is

$$\tau_i = \frac{\sigma_i}{2}\,, \tag{1.95}$$

where σ_i are the usual Pauli matrices given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\,, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,. \tag{1.96}$$

Recalling that $\sigma_i\sigma_j = \delta_{ij}\mathbf{1}_2 + i\epsilon_{ijk}\sigma_k$, it is easy to show that $\mathrm{Tr}\,(\tau_i\tau_j) = \delta_{ij}/2$ and that the $SU(2)$ generators fulfil the commutation relations

$$[\tau_i, \tau_j] = i\epsilon_{ijk}\tau_k\,. \tag{1.97}$$

Interestingly, (1.93) and (1.97) are the same commutation relations. This means that the matrices τ_i and T_i form two different *representations* of the $SU(2)$ Lie algebra. The representation that is classified by the 2×2 matrices (1.95) is called *fundamental*. It acts on a complex vector of size two and hence the dimension of the representation space is two. The representation where the generators are 3×3 matrices of the form $(T_i)_{jk} = -i\epsilon_{ijk}$ (as in (1.92)) is instead called the *adjoint* representation of $SU(2)$. This representation acts on 3-dimensional real vectors and therefore it has the dimension of $SU(2)$ that is equal to three.

A Complex Scalar Field With $U(1)$ Symmetry

For the further discussion it will prove convenient to arrange the two real scalar fields that we have meet at the beginning of the previous subsection into a single *complex* Klein-Gordon field defined as

$$\varphi = \frac{1}{\sqrt{2}} \left(\phi_1 + i\phi_2 \right) . \tag{1.98}$$

On this complex representation, the transformation (1.85) acts as follows

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \to \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \implies \begin{pmatrix} \phi_1 \\ i\phi_2 \end{pmatrix} \to \begin{pmatrix} \cos\alpha & i\sin\alpha \\ i\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \phi_1 \\ i\phi_2 \end{pmatrix} \tag{1.99}$$

$$\implies \quad \phi_1 + i\phi_2 \to e^{i\alpha} \left(\phi_1 + i\phi_2 \right) .$$

This means that all terms in the Lagrangians we are looking for should be invariant under a field phase-redefinition *aka* a global $U(1)$ transformation,

$$\varphi \to \varphi' = e^{i\alpha}\varphi , \qquad \varphi^* \to (\varphi')^* = e^{-i\alpha}\varphi^* . \tag{1.100}$$

The different transformation properties of φ and φ^* tell us that these fields carry charge 1 and -1, respectively.

Since φ and φ^* carry a non-zero charge not all terms polynomial in the fields are allowed in the Lagrangian if we want to respect the symmetry. *E.g.*, a term φ^2 has charge $+2$ and is hence not invariant under (1.100). In general, we can only allow terms with the same number of φ and φ^* field, so that the sought $U(1)$ invariant Lagrangian takes the form

$$\mathcal{L} = (\partial_\mu \varphi^*)(\partial^\mu \varphi) - V(\varphi^*\varphi) , \tag{1.101}$$

with

$$V(\varphi^*\varphi) = V_0 + m^2 \varphi^*\varphi + \frac{\lambda}{4} \left(\varphi^*\varphi \right)^2 . \tag{1.102}$$

Note that it is essential for the invariance of the kinetic term that the group parameter α is space-time independent. For the EOM for φ, we find from the Euler-Lagrange equation (1.22),

$$\partial^2 \varphi + \frac{\partial V}{\partial \varphi^*} = \partial^2 \varphi + m^2 \varphi + \frac{\lambda}{2} \left(\varphi^*\varphi \right) \varphi = 0 . \tag{1.103}$$

19

A similar equation holds in the case of φ^*. The conjugate momenta are $\pi = \dot{\varphi}^*$ and $\pi^* = \dot{\varphi}$, so that the Hamiltonian reads

$$\mathcal{H} = \pi \dot{\varphi}^* + \pi^* \dot{\varphi} - \mathcal{L} = \pi^* \pi + (\boldsymbol{\nabla} \varphi^*) \cdot (\boldsymbol{\nabla} \varphi) + V(\varphi^* \varphi) \,. \tag{1.104}$$

Being translation and Lorentz invariant the above theory has conserved stress and angular momentum tensors which can be obtained in complete analogy with the single scalar field case discussed before. We leave the derivation of the explicit expressions for $T^{\mu\nu}$ and Q^{ij} as an useful exercise.

According to Noether's theorem the presence of the internal symmetry (1.100) leads to a new type of conserved current which we will now derive. The variations of the fields φ and φ^* (treated as independent) under the global $U(1)$ symmetry are

$$\delta\varphi = i\alpha\varphi \,, \qquad \delta\varphi^* = -i\alpha\varphi^* \,. \tag{1.105}$$

By plaguing these variations into the general formula (1.34) and realizing that in our case $\mathcal{J}^\mu = 0$ since $\delta\mathcal{L} = 0$, the conserved current and charge is readily written down:

$$J^\mu = i\left(\varphi^* \partial^\mu \varphi - \varphi \partial^\mu \varphi^*\right) \,, \qquad Q = i\int d^3x \left(\varphi^* \dot{\varphi} - \varphi \dot{\varphi}^*\right) = i\int d^3x \left(\varphi^* \pi^* - \varphi \pi\right) \,. \tag{1.106}$$

There is an ambiguity worth noting, when applying Noether's theorem to find the conserved charge under the transformation (1.100). Obviously, if Q is conserved, then so is every other operator $c_1 Q + c_2$ with $c_{1,2}$ constant numbers. The expression for Q in (1.106) is therefore unique up to a multiplicative and an additive constant. The multiplicative factor essentially denotes the units in which we measure the charge. Notice that we have already used this ambiguity in (1.106) and simply ignored a factor $-\alpha$. Later we will also learn how the ambiguity on the additive constant is removed.

Spontaneous Symmetry Breaking

We would now like to discuss the vacua of the theory, *i.e.*, solutions to the EOM (1.103) with $\varphi = v = \text{const}$. Like in the real scalar case for $m^2 \geq 0$ there is a single minimum at $\varphi = 0$. This solution is left invariant under the transformations (1.100) and hence the $U(1)$ symmetry is unbroken. For $m^2 < 0$ the shape of the potential is illustrated in Figure 1.4. We can see from the depicted red line that in this case there is a whole circle of minima

$$\varphi = v = \frac{v_0}{\sqrt{2}} e^{i\delta} \,, \qquad v_0 = \sqrt{\frac{-4m^2}{\lambda}} \,, \tag{1.107}$$

where δ is an arbitrary phase. Clearly, the existence of this one-dimensional space of vacua is not an accident, but originates from the $U(1)$ invariance of the scalar potential (1.102):

$$V(\varphi^* \varphi) \to V((\varphi')^* \varphi') = V(e^{-i\alpha} \varphi^* e^{i\alpha} \varphi) = V(\varphi^* \varphi) \,. \tag{1.108}$$

In fact, this invariance implies that for every minimum φ of V also $\varphi' = \exp(i\alpha)\varphi$ is a minimum for arbitrary α. We can use this freedom to choose a particular minimum that
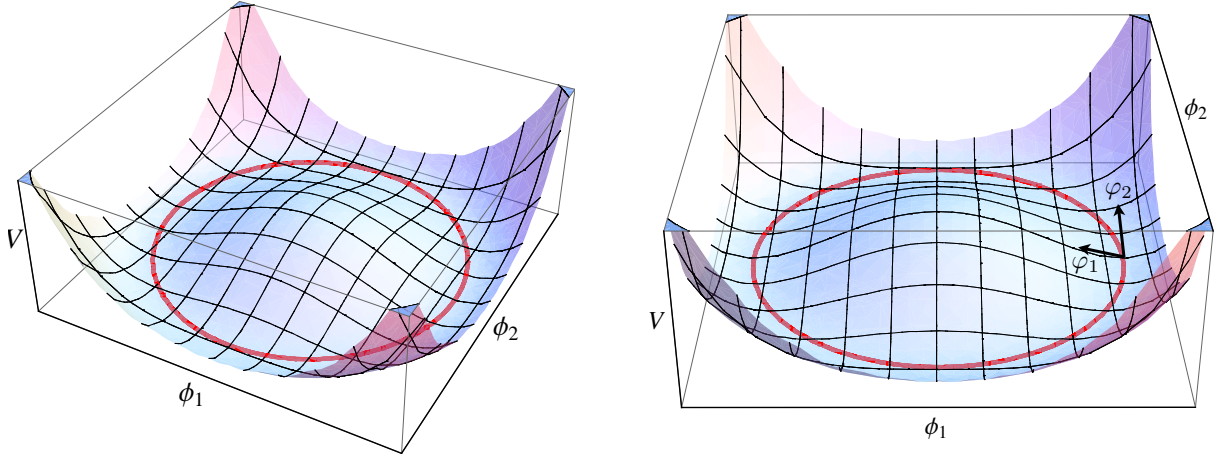
Figure 1.4: Left: Shape of the scalar potential of the complex Klein-Gordon field for $m^2 < 0$. The positions of the minima are indicated by the red curve. Right: Field excitations $\varphi_{1,2}$ (black arrows) around the minimum of the scalar potential at $\varphi = v_0\sqrt{2}$.

breaks the $U(1)$ symmetry spontaneously. Let's take the minimum on the ϕ_1 axis, meaning $\varphi = v_0/\sqrt{2}$. Around this point we then can expand the field as

$$\varphi = \frac{1}{\sqrt{2}} \left(v_0 + \varphi_1 + i\varphi_2 \right) , \tag{1.109}$$

where $\varphi_{1,2}$ are understood to be small fluctuations. Inserting the above relation into (1.102), we find

$$V = V_0 + \frac{1}{4} m^2 v_0^2 - m^2 \varphi_1^2 + \mathcal{O}(\varphi_1^3, \varphi_2^3) . \tag{1.110}$$

This is an interesting result, because it tells us that while the excitation φ_1 is massive with mass $\sqrt{2}m$ (remember $m^2 < 0$) the excitation φ_2 remains massless. Looking at Figure 1.4 this is not a completely unexpected feature. While φ_2 corresponds to an fluctuation around the circle of minima (a flat direction in the potential), φ_1 represents an excitation perpendicular to φ_1 (a direction of the potential with a curvature). The appearance of massless scalars for spontaneously broken global symmetries is a general feature known as *Goldstone's theorem* and the corresponding massless scalars are also called *Goldstone bosons*. We will now study this phenomenon in a more general setting.

Goldstone's Theorem

To generalize the above findings, we consider a field theory involving a set of scalar fields $\boldsymbol{\phi} = (\phi_1, \dots, \phi_n)^T$ with corresponding potential $V(\boldsymbol{\phi})$. We assume that the potential has a minimum at $\boldsymbol{v} = (v_1, \dots, v_n)^T$ that satisfies

$$\left. \frac{\partial V}{\partial \phi_a} \right|_{\boldsymbol{\phi}=\boldsymbol{v}} = 0 . \tag{1.111}$$

After defining $\boldsymbol{\varphi} = \boldsymbol{\phi} - \boldsymbol{v}$, we can Taylor expand the potential around such a minimum in the following way

$$V = V(\boldsymbol{v}) + \frac{1}{2} M_{ab}\,\varphi^a \varphi^b + \mathcal{O}(\boldsymbol{\varphi}^3)\,, \tag{1.112}$$

where $(a, b = 1, \ldots, n)$

$$M_{ab} = \left.\frac{\partial^2 V}{\partial \phi_a \partial \phi_b}\right|_{\phi=v}. \tag{1.113}$$

Clearly, the eigenvalues of the mass matrix M are the mass squares of the fields (around the vacuum $\boldsymbol{v}$).

Now let us assume that our scalar theory is invariant under a continuous (global) symmetry group G, under which

$$\boldsymbol{\phi} \to R(g)\boldsymbol{\phi}\,, \tag{1.114}$$

where $R(g)$ denotes the representation of $\boldsymbol{\phi}$. The invariance of the scalar theory in particular means that the scalar potential is invariant under such a transformation, namely

$$V(\boldsymbol{\phi}) = V(R^{-1}(g)\boldsymbol{\phi})\,, \tag{1.115}$$

for all $g \in G$. The vacuum will in general not respect the full symmetry group G, but will spontaneously break it into a subgroup $H \subset G$, so that

$$R(g)\boldsymbol{v} = \boldsymbol{v} \ \text{ for } g \in H\,, \qquad R(g)\boldsymbol{v} \neq \boldsymbol{v} \ \text{ for } g \notin H\,. \tag{1.116}$$

Now introducing infinitesimal group transformations

$$R(g) = \mathbf{1} + it^\alpha T_\alpha\,, \tag{1.117}$$

with generators T_α (in the representation R) and small coefficients t^α. The set of generators $\{T_\alpha\}$ can be split into two subsets T_i and T_j, where T_i denotes the generators of the unbroken subgroup H, while T_j are the remaining generators corresponding to the broken part of G. By definition, one has

$$T_i\boldsymbol{v} = 0\,, \qquad T_j\boldsymbol{v} \neq 0\,. \tag{1.118}$$

Now let us see how (1.117) acts on our potential. We have

$$V(\boldsymbol{\phi}) = V(\boldsymbol{\phi} - it^\alpha T_\alpha \boldsymbol{\phi}) = V(\boldsymbol{\phi}) - i\left(\frac{\partial V}{\partial \boldsymbol{\phi}}\right)^T t^\alpha T_\alpha \boldsymbol{\phi}\,. \tag{1.119}$$

Differentiating this expression again and evaluating it at $\boldsymbol{\phi} = \boldsymbol{v}$, one finds using (1.111) and (1.113), that

$$MT_\alpha \boldsymbol{v} = 0\,. \tag{1.120}$$

From this relation we see that every broken generator T_j (with $T_j\boldsymbol{v} \neq 0$) leads necessarily to an eigenvector of M with eigenvalue zero. Or in more physical terms, every broken generator leads to a massless scalar excitation. We have proven Goldstone's theorem!

Scalar Field Theory With $SU(2) \times U(1)$ Symmetry

Let me further illustrate Goldstone's theorem by a still simple but educated example. We consider the symmetry $SU(2) \times U(1)$ and a complex scalar field Φ that is a doublet under $SU(2)$ and has a charge of $1/2$ under the $U(1)$. Our scalar transforms as

$$\Phi \to e^{i\alpha Y} e^{it^i \tau_i} \Phi \simeq \left(1 + i\alpha Y + it^i \tau_i\right) \Phi \,, \tag{1.121}$$

with the generators

$$Y = \frac{1}{2} \,, \qquad \tau_i = \frac{\sigma_i}{2} \,. \tag{1.122}$$

Notice that our notation is on purpose sloppy here, since the 1 on the right-hand side of (1.121) should be in fact $\mathbf{1}_2$. A similar statement also applies to the generator Y. Such a sloppiness in notation is common in textbooks, and it is good to get used to it early.

The Lagrangian we are interested in, can be written as

$$\mathcal{L} = (\partial_\mu \Phi)^\dagger (\partial^\mu \Phi) - V(\Phi^\dagger \Phi) \tag{1.123}$$

with

$$V(\Phi^\dagger \Phi) = V_0 + \mu^2 \, \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2 \,. \tag{1.124}$$

Provided that $\lambda > 0$ and $\mu^2 < 0$ (again a change in notation), the scalar potential (1.124) is bounded from below and minimized for

$$\Phi^\dagger \Phi = v_0^2 = -\frac{\mu^2}{2\lambda} \,. \tag{1.125}$$

Using (1.121) allows us to choose a particular simple vacuum configuration. We take

$$\Phi = \boldsymbol{v} = \begin{pmatrix} 0 \\ v_0 \end{pmatrix} \,. \tag{1.126}$$

For this choice, one finds readily

$$\tau_1 \boldsymbol{v} \neq 0 \,, \qquad \tau_2 \boldsymbol{v} \neq 0 \,, \qquad (\tau_3 - Y) \, \boldsymbol{v} \neq 0 \,, \tag{1.127}$$

and

$$(\tau_3 + Y) \, \boldsymbol{v} = 0 \,. \tag{1.128}$$

Hence three of the four generators of $SU(2) \times U(1)$ are broken, which corresponds to the breaking pattern

$$SU(2) \times U(1) \to U(1) \,, \tag{1.129}$$

where the $U(1)$ on the right-hand side is the diagonal subgroup of $SU(2) \times U(1)$ corresponding to the generator $Q = \tau_3 + Y$.

So why is this example relevant? The relevance of the example stems from the fact that the breaking pattern (1.129) is precisely that of the electroweak gauge group of the SM of particle physics: the SM is in fact based on $SU(3)_c \times SU(2)_L \times U(1)_Y \to SU(3)_c \times U(1)_{\text{em}}$,

where the subscript c refers to color, L indicates that the $SU(2)$ only acts on left-handed fields, Y represents hypercharge, and $U(1)_{\text{em}}$ corresponds to the unbroken electromagnetism associated to the generator Q, *i.e.*, the *electric charge*. Yet, there is an important difference. While our theory predicts three massless Goldstone bosons, in Nature we only observe a single weakly-interacting massless state, *i.e.*, the *photon*. This difference is explained by the fact that the electroweak symmetry in the SM is realized as a *local* or *gauge(d)* symmetry, while in our simple example we dealt with a global symmetry. It turns out that in the case of the spontaneously broken electroweak gauge symmetry our three massless Goldstone bosons are absorbed (or eaten) by three vector bosons (the electrical neutral Z boson and the charged $W^{\pm}$ bosons) that receive their mass from the breaking mechanism. This is the famous *Higgs mechanism*, which also gives rise to a (physical) massive scalar state the *Higgs* (sometimes also referred to as the "god particle"). Let me mention that last year both the ATLAS and the CMS collaborations (situated at the CERN LHC in Geneva) have announced the existence of a new bosonic state with a mass of around $125\,\text{GeV}$. This discovery defines a turning point in the history of elementary-particle physics: the almost 50 year-long hunt for the Higgs boson has come to a dazzling end. This year the Nobel committee gave the Nobel prize to Francois Englert and Peter W. Higgs for their theoretical work on the mechanism of electroweak symmetry breaking. In order to fully understand the physics of the Higgs mechanism, we are still missing an important ingredient, namely vector fields. These 4-component fields will be discussed now.

1.5   Vector Field Theories

In this section we will examine the basic features of another important ingredient of realistic (quantum) field theories, *i.e.*, vector fields. These fields carry spin one and are called vector fields, since they transform like a vector under the Lorentz group.

Maxwell's Theory Of Electromagnetism

As another simple application of the formalism we have developed, let us try to derive Maxwell's equations using the field theory formulation. In terms of the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$ and the charge density ρ and 3-vector current $\boldsymbol{j}$, these equations take the well-known form

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0\,, \tag{1.130}$$

$$\boldsymbol{\nabla} \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = 0\,, \tag{1.131}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = \rho\,, \tag{1.132}$$

$$\boldsymbol{\nabla} \times \boldsymbol{B} - \frac{\partial \boldsymbol{E}}{\partial t} = \boldsymbol{j}\,. \tag{1.133}$$

The $\boldsymbol{E}$ and $\boldsymbol{B}$ fields are spatial 3-vectors and can be expressed in terms of the components

of the 4-vector field $A^\mu = (\phi, \boldsymbol{A})$ by

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \qquad \boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}. \tag{1.134}$$

This definition ensures that the first two homogeneous Maxwell equations (1.130) and (1.131) are automatically satisfied,

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{A}) = \epsilon_{ijk}\partial_i\partial_j A_k = 0, \tag{1.135}$$

$$\boldsymbol{\nabla} \times \left(-\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t}\right) + \frac{\partial}{\partial t}(\boldsymbol{\nabla} \times \boldsymbol{A}) = -\boldsymbol{\nabla} \times (\boldsymbol{\nabla}\phi) = -\epsilon_{ijk}\partial_j\partial_k\phi = 0. \tag{1.136}$$

Here the indices $i, j, k = 1, 2, 3$ are summed over if they appear twice.

The remaining two inhomogeneous Maxwell equations (1.132) and (1.133) follow from the Lagrangian

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2}(\partial_\mu A_\nu)(\partial^\nu A^\mu) - A_\mu J^\mu, \tag{1.137}$$

with $J^\mu = (\rho, \boldsymbol{j})$. From the rules presented before, we gather that the dimension of the vector field and current is $[A_\mu] = 1$ and $[J^\mu] = 3$, respectively. The funny minus sign of the first term on the right-hand side is required to ensure that the kinetic term $1/2\dot{A}_i^2$ is positive using the Minkowski metric. Notice also that the Lagrangian (1.137) has no kinetic term $1/2\dot{A}_0^2$ and hence A_0 is not dynamical. Why this is and necessarily has to be the case will become clear pretty soon.

Enough said, let's do serious business and compute something. To see that the statement made before (1.137) is indeed correct, we first evaluate

$$\frac{\partial \mathcal{L}}{\partial A_\nu} = -J^\nu, \qquad \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\partial^\mu A^\nu + \partial^\nu A^\mu, \tag{1.138}$$

from which we derive the EOMs,

$$0 = \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)}\right) - \frac{\partial \mathcal{L}}{\partial A_\nu} = \partial_\mu(-\partial^\mu A^\nu + \partial^\nu A^\mu) + J^\nu. \tag{1.139}$$

Introducing now the *field-strength tensor*

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{1.140}$$

we can write (1.137) and (1.139) quite compact,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - J_\mu A^\mu. \tag{1.141}$$

$$\partial_\mu F^{\mu\nu} = J^\nu, \tag{1.142}$$

Does this look familiar? I hope so. Notice that we have $[F_{\mu\nu}] = 2$ and that the field-strength tensor satisfies the *Bianchi identity*:

$$\partial_\rho F_{\mu\nu} + \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} = 0. \tag{1.143}$$

In order to see that (1.142) indeed captures the physics of (1.132) and (1.133), we compute the components of $F^{\mu\nu}$. We find

$$F^{0i} = -F^{i0} = \partial^0 A^i - \partial^i A^0 = \left(\boldsymbol{\nabla}\phi + \frac{\partial \boldsymbol{A}}{\partial t} \right)^i = -E^i \,,$$

$$F^{ij} = -F^{ji} = \partial^i A^j - \partial^j A^i = -\epsilon^{ijk} B^k \,,$$

(1.144)

while all other components are zero. With this in hand, we then obtain from $\partial_\mu F^{\mu 0} = \rho$ and $\partial_\mu F^{\mu 1} = j^1$,

$$\partial_\mu F^{\mu 0} = \partial_0 F^{00} + \partial_i F^{i0} = \boldsymbol{\nabla} \cdot \boldsymbol{E} = \rho \,,$$

$$\partial_\mu F^{\mu 1} = \partial_0 F^{01} + \partial_i F^{i1} = -\frac{\partial E^1}{\partial t} + \frac{\partial B^3}{\partial x_2} - \frac{\partial B^2}{\partial x_3} = \left(\boldsymbol{\nabla} \times \boldsymbol{B} - \frac{\partial \boldsymbol{E}}{\partial t} \right)^1 = j^1 \,.$$

(1.145)

Similar relations hold for the remaining components $i = 2, 3$. Taken together this proves the second inhomogeneous Maxwell equation (1.133).

Let me also derive the energy-momentum tensor $T^{\mu\nu}$ of electrodynamics, ignoring for the moment the source term $J_\mu A^\mu$. Using (1.138) one finds

$$T^{\mu\nu} = -F^{\mu\rho} \partial^\nu A_\rho + \frac{1}{4} \eta_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} \,.$$

(1.146)

Notice that the first term in (1.146) is not symmetric, which implies that $T^{\mu\nu} \neq T^{\nu\mu}$. In fact, this is not really surprising since the definition of the energy-momentum tensor (1.37) does not exhibit an explicit symmetry in the indices μ and ν. Nevertheless, there is always a way to massage the energy-momentum tensor of any theory into a symmetric form.[6] To learn how this can be done in the case under consideration is the objective of a homework problem.

Physical Degrees Of Freedom

Under LTs, A_μ transforms as a vector, *i.e.*, $A_\mu \to \Lambda_\mu{}^\nu A_\nu$. The same applies of course also to the current J_μ. The field-strength tensor $F_{\mu\nu}$, on the other hand, carries two indices and hence transforms as a tensor, *i.e.*, $F_{\mu\nu} \to \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma F_{\rho\sigma}$. Equipped with these transformation properties it is an easy exercise to show that the Lagrangian (1.141) is Lorentz invariant (at the end it is sufficient to observe that in the expression for $\mathcal{L}$ all indices are contracted, so that (1.141) is a Lorentz scalar). The explicit calculation is left as an exercise to the interested reader.

After (1.137) we have already mentioned that not all components of A_μ are dynamical, which signals that A_μ contains unphysical dofs. In fact, this is not really a surprise because it is $F_{\mu\nu}$ that is directly related to the physical fields $\boldsymbol{E}$ and $\boldsymbol{B}$ not A_μ (see (1.144)). Formally, this is expressed by the fact that a *gauge transformation*

$$A_\mu(x) \to A_\mu(x) + \partial_\mu f(x) \,,$$

(1.147)

---

[6]One (but not the only) reason that you might want to have a symmetric energy-momentum tensor $T^{\mu\nu}$ is to make contact with general relativity, since such an object sits on the right-hand side of Einstein's field equations.

with arbitrary function $f(x)$ leaves $F_{\mu\nu}$ invariant:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \rightarrow \partial_\mu(A_\nu + \partial_\nu f) - \partial_\nu(A_\mu + \partial_\mu f)$$

$$= \partial_\mu A_\nu - \partial_\nu A_\mu + \partial_\mu \partial_\nu f - \partial_\nu \partial_\mu f = F_{\mu\nu} \,. \tag{1.148}$$

This is an important result, because it gives us a recipe on how to "derive" the covariant formulation of electrodynamics without ever talking about Maxwell's equations. The idea is simply to find the most general Lagrangian (including terms up to a certain mass dimension) that is Lorentz invariant and unchanged under (1.147). Notice that gauge invariance in particular implies that $\mathcal{L}$ should only depend on A_μ through $F_{\mu\nu}$. Up to dimension four this leaves basically only one term[7] (in the case of a vanishing source term), namely $\mathcal{L} = -1/4\, F_{\mu\nu}F^{\mu\nu}$. The associated EOM is of course $\partial_\mu F^{\mu\nu} = 0$, or written in terms of A_μ,

$$\partial^2 A_\mu - \partial_\mu \partial_\nu A^\nu = 0 \,. \tag{1.149}$$

In fact, this equation can be further simplified by exploiting the gauge symmetry (1.147) by imposing a *gauge condition* on A_μ through a suitable choice of gauge parameter f. There are various possibilities for such a gauge choice and here we consider the *Lorenz gauge* defined by

$$\partial_\mu A^\mu = 0 \,. \tag{1.150}$$

This gauge has the salient benefit that it is covariant (in contrast to other gauges such as *temporal gauge* $\dot{A}_0 = 0$ or *Coloumb gauge* $\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$) and it simplifies (1.149) to

$$\partial^2 A_\mu = 0 \,. \tag{1.151}$$

It is however important to realize that (1.150) does not fix the gauge entirely but leaves a *residual gauge freedom*, since all choices of f with

$$\partial^2 f = 0 \,, \tag{1.152}$$

are equivalent in the sense that they all lead to (1.151).

The solutions of $\partial^2 A_\mu = 0$ are immediately written done after noticing that the latter equation is nothing but a Klein-Gordon equation for a massless vector field. Adding vector indices to (1.52), we arrive at

$$A_\mu(x) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2\omega_{\boldsymbol{p}}} \left(a_\mu(p) e^{-ipx} + a_\mu^*(p) e^{ipx} \right) \,, \tag{1.153}$$

where $\omega_{\boldsymbol{p}} = |\boldsymbol{p}|$ and $p_\mu = (\omega_{\boldsymbol{p}}, \boldsymbol{p})$. Fourier transforming the Lorenz gauge condition (1.150), we find in addition

$$p^\mu a_\mu(p) = 0 \,. \tag{1.154}$$

---

[7] The term $F_{\mu\nu}\tilde{F}^{\mu\nu}$ with $\tilde{F}^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$ denoting the *dual field-strength tensor* also mets all requirements. It can however be written as a total derivative $4\partial_\mu(\epsilon^{\mu\nu\rho\sigma}A_\nu\partial_\rho A_\sigma)$ and therefore does not contribute to the *classic* EOMs. We hence ignore such a term.

To exploit this constraint one conveniently introduces a set of polarization vectors $\epsilon_\mu^{(\alpha)}(p)$ with $\alpha = 0, 1, 2, 3$ and the following properties. The vectors $\epsilon_\mu^{(1)}$ and $\epsilon_\mu^{(2)}$ are orthogonal to both p and a vector n with $n^2 = 1$ and $n_0 > 0$. They furthermore are chosen such that they satisfy $\epsilon^{(\alpha)} \cdot \epsilon^{(\beta)} = -\delta^{\alpha\beta}$ for $\alpha, \beta = 1, 2$. The polarization vector $\epsilon_\mu^{(3)}$ is taken to be in the p–n plane, orthogonal to n and normalized to -1, i.e., $n \cdot \epsilon^{(3)} = 0$ and $(\epsilon^{(3)})^2 = -1$. Finally, one defines $\epsilon^{(0)} = n$. With these conventions one has an orthogonal set of vectors satisfying

$$\epsilon^{(\alpha)} \cdot \epsilon^{(\beta)} = \eta^{\alpha\beta} \,, \tag{1.155}$$

for $\alpha, \beta = 0, 1, 2, 3$. This basis can be used to write

$$a_\mu(p) = \sum_{\alpha=0}^{3} a^{(\alpha)}(p) \, \epsilon_\mu^{(\alpha)}(p) \,. \tag{1.156}$$

In fact, the basic idea behind all this "mumbo-jumbo" is that this specific choice of basis of polarization vectors, allows to easily separate the directions transversal to p (corresponding to $\epsilon_\mu^{(1)}$ and $\epsilon_\mu^{(2)}$) from the other two directions (corresponding to $\epsilon_\mu^{(0)}$ and $\epsilon_\mu^{(3)}$) that are longitudinal to p. As an example, if we choose a spatial momentum $\boldsymbol{p}$ pointing in the z-directions, the components of above vectors are explicitly given by

$$\epsilon_\mu^{(\alpha)} = \delta^{\alpha\mu} \,. \tag{1.157}$$

In the general case, one has instead

$$\epsilon_\mu^{(3)} = \frac{p_\mu}{p_0} - n_\mu \,. \tag{1.158}$$

Clearly,

$$n \cdot \epsilon^{(3)} = \frac{np}{p_0} - n^2 = \frac{np}{p_0} - 1 = 0 \implies \frac{np}{p_0} = 1 \,, \tag{1.159}$$

and thus

$$(\epsilon^{(3)})^2 = \frac{p^2}{p_0^2} + n^2 - 2\frac{np}{p_0} = 1 - 2\frac{np}{p_0} = -1 \,, \tag{1.160}$$

where we have used that our vector field is massless, i.e., $p^2 = 0$. So (1.158) is indeed the sought solution. Notice also that

$$\epsilon_\mu^{(0)} + \epsilon_\mu^{(3)} = \frac{p_\mu}{p_0} \,. \tag{1.161}$$

We now return to (1.154) and evaluate this sum inserting the expansion (1.156). By definition the transversal components $\epsilon_\mu^{(1)}$ and $\epsilon_\mu^{(2)}$ do not contribute and we find

$$p^\mu a_\mu(p) = (p \cdot \epsilon^{(0)}) a^{(0)}(p) + (p \cdot \epsilon^{(3)}) a^{(3)}(p) = p_0 \left(a^{(0)}(p) - a^{(3)}(p) \right) = 0 \,. \tag{1.162}$$

Here we have used that

$$p \cdot \epsilon^{(0)} = -p \cdot \epsilon^{(3)} = p_0 \,. \tag{1.163}$$

Note that the first equality follows from (1.161) after contracting it with p^μ, while the second equality is a consequence of (1.159). From (1.162) it follows that $a^{(0)}(p) = a^{(3)}(p)$. The expansion (1.156) can thus be written as

$$a_\mu(p) = a^{(3)}(p) \frac{p_\mu}{p_0} + \sum_{\alpha=1}^{2} a^{(\alpha)}(p) \, \epsilon_\mu^{(\alpha)}(p) \,, \tag{1.164}$$

and we are left with two transversal modes and a longitudinal one along the p direction (remember that after (1.137) we already mentioned the fact that the time component A_0 is not dynamical, which reduces the number of independent dofs in A_μ from four to three). This is still one more dof than a physical massless field, such as the photon, ought to have.

The trick to get rid of the remaining longitudinal component is to make use of the residual gauge freedom (1.152) and to "gauge away" the term proportional to $a^{(3)}(p)$ in (1.164). Let see how this works. Given that f also fulfils a Klein-Gordon equation, the most general decomposition obviously reads

$$f(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\boldsymbol{p}}} \left(\xi(p) e^{-ipx} + \xi^*(p) e^{ipx} \right) \,. \tag{1.165}$$

A gauge transformation (1.147) with such a f corresponds to

$$a_\mu(p) \to a'_\mu(p) = a_\mu(p) - ip_\mu \xi(p) \,. \tag{1.166}$$

But such a shift is exactly what is needed to remove the remaining longitudinal component from (1.164) and we arrive at the correct physical answer. Needless to say that the reduction from apparent four dofs to actually two is intimately related to the gauge invariance of (quantum) electrodynamics.

A Massive Vector Field

What happens if we try to give our vector field A_μ a mass? We add a term

$$\frac{1}{2} m^2 A_\mu A^\mu \,, \tag{1.167}$$

to our Lagrangian $\mathcal{L} = -1/4 \, F_{\mu\nu} F^{\mu\nu}$, but immediately realize that such a term is not allowed by gauge invariance (1.147). This is of course stupid, but let us ignore this unwanted feature for a moment. The EOMs for the massive vector field are not $\partial_\mu F^{\mu\nu} = 0$, but

$$\partial_\mu F^{\mu\nu} + m^2 A^\nu = 0 \,. \tag{1.168}$$

Applying ∂_ν to this relation we conclude that $\partial_\nu A^\nu = 0$, since $\partial_\nu \partial_\mu F^{\mu\nu}$ is trivially zero. We can hence split the latter equation into two

$$(\partial^2 + m^2) A_\mu = 0 \,, \qquad \partial_\mu A^\mu = 0 \,. \tag{1.169}$$

The first equation is a massive Klein-Gordon equation for A_μ with the general solution (1.153) but $\omega_{\boldsymbol{p}} = (|\boldsymbol{p}|^2 + m^2)^{1/2}$ instead of $\omega_{\boldsymbol{p}} = |\boldsymbol{p}|$. In order to satisfy the second equation we need to impose the condition (1.154), which reduces the number of dofs from four to three. Can we scrap another dof? No we can't, because the mass term (1.167) breaks gauge invariance and thus we do not have the freedom to use (1.152) and gauge it away. This means that a massive vector field has three physical dofs, one more than a massless one.

Current Sources

In the discussion of the last two subsections we have always ignored the possible presence of external sources by simply setting $J_\mu = 0$ by hand. The Lagrangian leading to Maxwell's theory including a source term $J_\mu A^\mu$ has already been given in (1.141) and by applying (1.147), we see that a gauge transformation of such a term introduces an additional piece into the action:

$$S \to S - \int d^4x \, (\partial_\mu f) \, J^\mu = S + \int d^4x \, f \, (\partial_\mu J^\mu) \,. \qquad (1.170)$$

As usual we have used integration by parts to arrive at the final answer. This result tells us that the apparent breaking of gauge invariance by the term $J_\mu A^\mu$ can be avoided, if we require that

$$\partial_\mu J^\mu = 0 \,. \qquad (1.171)$$

In other words the vector field A_μ should couple to a conserved current J_μ. In a fundamental theory, the current J_μ should arise from fields, rather than put in by hand as an external source. We conclude the chapter on classical field theory by studying the simplest fundamental theory of such a kind.

Scalar Electrodynamics And Abelian Higgs Mechanism

The goal of this final subsection is to write down an interesting theory involving a complex scalar φ and a vector field A_μ. By interesting we mean that the two fields should interact, which tells us that we want to have a Lagrangian that involves terms with both φ and A_μ. The problem is now that in general φ and A_μ do not transform trivially under a given symmetry. E.g., the complex scalar φ transforms as (1.100) under a global $U(1)$, while A_μ obeys the gauge transformation (1.147). The important point to notice is that both transformations involve only a single parameter, namely α and f. However, $\alpha = \text{const.}$ is space-time independent, while $f = f(x)$ is a function of the space-time coordinate x. What happens if we generalize (1.100) by allowing the $U(1)$ symmetry to be *local* rather than global? Meaning that our fields φ and φ^* should transform as

$$\varphi \to \varphi' = e^{i\alpha(x)}\varphi \,, \qquad \varphi^* \to (\varphi')^* = e^{-i\alpha(x)}\varphi^* \,. \qquad (1.172)$$

Since the scalar potential is polynomial in $\varphi^*\varphi$ it is clearly invariant under these local $U(1)$ transformations. But as already noted after (1.102), for the kinetic term to be invariant, we need $\alpha = \text{const.}$ Explicitly, we have

$$\partial_\mu \varphi \to \partial_\mu \left(e^{i\alpha(x)}\varphi \right) = e^{i\alpha(x)} \left(\partial_\mu + i\partial_\mu \alpha(x) \right) \varphi \,. \qquad (1.173)$$

Obviously, the term involving $\partial_\mu \alpha(x)$ will cause $(\partial_\mu \varphi^*)(\partial^\mu \varphi)$ to transform non-trivially under (1.172). But not all hope is lost since we still have the gauge transformation (1.147) of the vector field, which also involves a term with a derivative. So let us see how $(\partial_\mu - iA_\mu)\varphi$ transforms. One finds

$$(\partial_\mu - iA_\mu)\varphi \to e^{i\alpha} \left[(\partial_\mu + i\partial_\mu \alpha) - i(A_\mu + \partial_\mu f) \right] \varphi = e^{i\alpha} \left[(\partial_\mu - iA_\mu) + i\partial_\mu(\alpha - f) \right] \varphi \,. \qquad (1.174)$$

"Eureka!" if we identify $\alpha(x) = f(x)$ then

$$D_\mu \varphi = (\partial_\mu - iA_\mu)\varphi \,, \tag{1.175}$$

has the correct properties so that

$$(D_\mu \varphi)^* (D^\mu \varphi) \to (D_\mu \varphi)^* (D^\mu \varphi) \,, \tag{1.176}$$

under

$$\varphi \to e^{i\alpha} \varphi \,, \qquad A_\mu \to A_\mu + \partial_\mu \alpha \,. \tag{1.177}$$

In other words the new kinetic term $(D_\mu \varphi)^* (D^\mu \varphi)$ with ∂_μ replaced by the *covariant derivative* D_μ is gauge invariant. Notice that the covariant derivative (1.175) has dimension $[D_\mu] = 1$ and involves a "minimal" coupling $A_\mu \varphi$ between the vector field A_μ and the scalar φ.

We can now combine our scalar and gauge field Lagrangian to obtain

$$\mathcal{L} = (D_\mu \varphi)^* (D^\mu \varphi) - V(\varphi^* \varphi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \,, \quad V(\varphi^* \varphi) = V_0 + m^2 \varphi^* \varphi + \frac{\lambda}{4} (\varphi^* \varphi)^2 \,, \tag{1.178}$$

which is the Lagrangian of *scalar electrodynamics*. Notice that in this new theory the role of A_μ is to facilitate the invariance of the scalar field theory under local gauge transformations. In fact, if we would have started from the global $U(1)$ invariant scalar theory with the task of finding its locally $U(1)$ invariant version we would have necessarily been led to introducing a gauge field A_μ.

We will now use (1.178) as a toy model to study the spontaneous symmetry breaking of a local $U(1)$ symmetry. Since the scalar potential is identical to the one in the global case, we can reuse all our old results. The case $m^2 \geq 0$ is again not very interesting since the symmetry is unbroken, so let's focus on the situation with $m^2 < 0$. In the latter case the potential takes the form of a "Mexican hat" as in Figure 1.4. We parametrize φ around the minima in the following way

$$\varphi = \frac{1}{\sqrt{2}} (v_0 + h) \, e^{i\chi} \,, \tag{1.179}$$

with v_0 given in (1.107). It is straightforward to see that h corresponds to a real massive field, while χ is a Goldstone mode (also a real field). We can now eliminate the Goldstone mode from φ by choosing the specific gauge

$$\varphi \to \varphi' = e^{-i\chi} \varphi = \frac{1}{\sqrt{2}} (v_0 + h) \,. \tag{1.180}$$

Under this transformation, we see from (1.177) that

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu \chi \,. \tag{1.181}$$

Since the Lagrangian (1.178) is gauge invariant, we can as well write it in terms of primed rather than unprimed fields. For the covariant derivative, one finds

$$(D_\mu \varphi)' = \frac{1}{\sqrt{2}} \left(\partial_\mu h - i \, (v_0 + h) \, A'_\mu \right) \,, \tag{1.182}$$

31

and inserting this expression into (1.178) leads to

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu h\right)^2 + \frac{1}{2}A'_\mu A'^\mu\left(v_0^2 + 2v_0 h + h^2\right) - V(h) - \frac{1}{4}F'_{\mu\nu}F'^{\mu\nu}\,, \tag{1.183}$$

with

$$V(h) = V_0 + \frac{1}{4}m^2 v_0^2 - m^2 h^2 + \frac{\lambda}{16}\left(4v_0 h^3 + h^4\right)\,. \tag{1.184}$$

The most striking feature of this result is that the Goldstone mode χ has (magically) disappeared and we are only left with the fields h and A'_μ. So did we loose a dof? No, because the vector field is now no longer massless, but has a mass

$$m_{A'} = v_0\,. \tag{1.185}$$

It has hence three dofs (as opposed to just two in the massless case), which explains the disappearance of χ: the Goldstone boson has been "eaten" by the vector field and provides the additional longitudinal component for the massive A'_μ. This feature can also be inferred from (1.181). Notice furthermore that the mass of h is given by

$$m_h = m\,, \tag{1.186}$$

and that (1.183) and (1.184) have a rich structure of interactions: there are couplings of two gauge fields to one and two massive scalars ($A'A'h$ and $A'A'h^2$) as well as cubic and quartic self-couplings of the massive scalar (h^3 and h^4).

So let's briefly summarize our main findings: we have learned that a spontaneously broken local (or gauge) symmetry leads to a mass for the associated vector boson and the conversion of the Goldstone boson into the longitudinal component of the gauge field. This is the famous Higgs mechanism that we already alluded to. The massive scalar h is hence called *Higgs scalar*. The phenomenon described above (but in its generalization to *non-abelian* gauge groups) is at work in the SM of particle physics. In fact, since the SM symmetry breaking pattern is $SU(2)_L \times U(1)_Y \to U(1)_{\text{em}}$, one has three Goldstone bosons that give mass to three electroweak gauge bosons. The masses of these particles are all proportional to the symmetry breaking scale (or *vacuum expectation value*) v_0. A detailed discussion of the breaking of the electroweak symmetry, while worthwhile, is beyond the scope of this lecture, as it requires the introduction of non-abelian gauge symmetries. You will learn a little bit more in an exercise.