Lecture notes on Quantum Field Theory

Astrid Eichhorn,
notes typeset by F. Homburg*

Wintersemester 2025/26

Contents

*This is a draft version of the lecture notes. If you notice typos, points that remain unclear, mistakes or similar, please send an email to f.homburg@thphys.uni-heidelberg.de and a.eichhorn@thphys.uni-heidelberg.de.

0 Preliminaries

Literature

These lecture notes combine material from several sources. Some topics in these lecture notes are treated in all books listed below; for some, a particular book or selection of books is better suited. In this case, this will be indicated in the corresponding chapter.
In particular, you might find it useful to also read up on the topics in the following books as well as lecture notes (available online):

- Srednicki, Quantum Field Theory
- Ryder, Quantum Field Theory
- Gelis, Quantum Field Theory
- Schwartz, Quantum Field Theory and the Standard Model
- Peskin/Schröder, Quantum Field Theory
- Nastase, Quantum Field Theory
- Fradkin, Quantum Field Theory
- There are many other books on QFT and it is often a matter of personal taste, which one is most useful.
- Lecture notes on QFT by D. Tong (Cambridge University), A. Hebecker (Heidelberg University), T. Weigand (from the QFT courses at Heidelberg University)

Many topics are treated to the greatest level of depth in the QFT books by Weinberg. However, for a first encounter with a topic, the books are usually not useful, but rather become helpful later on, when one has already learned about a topic and wants to come back to it to learn more about it.
There is also the book "Quantum field theory in a nutshell" by Zee, which focuses more on some conceptual aspects rather than technical points and it can be a useful addition to the above list of literature.

Mini-exercises

The best way to learn quantum field theory is to do calculations yourself, and think and discuss about concepts yourself. Therefore, each lecture has at least one "mini-exercise", which you will work on during the lecture. This gives you the opportunity to engage more actively with the material and notice when you have questions. You will likely not always have time to finish the mini-exercise during class. Therefore, solutions will not only be provided on the blackboard, but are also available in the back of the lecture notes. They will be made available in the update of the lectures notes that will be made online after each lecture.

1 Introduction

1.1 Motivation: Why quantum field theory?

Quantum mechanics is a non-relativistic theory. This results in a question, namely:

→ What happens to systems in which quantum effects and relativistic effects are important? There is a heuristic argument that points us towards how relativistic quantum physics differs from quantum mechanics. From the standard Heisenberg uncertainty principle, one can motivate an uncertainty relation between energy E and time t, namely $\Delta E \Delta t \geqslant \frac{\hbar}{2}$. In a relativistic setting, we can combine this with $E = mc^2$, which we know from special relativity.
$\implies$ We expect that particle number is never fixed in a system, because, for short enough time durations, energy is not constant, but fluctuates and these fluctuations in energy translate into fluctuations in particle number. We call these fluctuations "virtual" particles.
$\implies$ We cannot work with a wavefunction for a fixed number of particles, as we did in quantum mechanics. Instead, we need a formalism in which the particle number can change in a system over time, and in which the presence of virtual particles is accounted for.

We can also see the incompatibility between special relativity and quantum mechanics in a different way:

→ Special relativity requires that two measurements that are done at spacelike separation, must be independent in order not to violate causality. In Quantum Mechanics, independence of measurements is encoded in commuting operators. However, the notion that spacelike separated operators commute is not naturally built into QM.

$\implies$ We need to adapt our formalism.

How should the new formalism look like?
To go beyond wavefunctions for fixed numbers of particles, we need a (mathematical) quantity that is more fundamental than particles, i.e., particles should be a *derived* notion.
We take inspiration from electrodynamics, because electrodynamics can be formulated in a relativistic way. At the same time, we know from the photo-electric effect, that there are particles in electrodynamics, namely photons. Thus, it is a useful guide to point us to the type of formalism that we should develop. Electrodynamics is a field theory, i.e., the fundamental quantity is a *field*, i.e., a quantity that takes on values at each spacetime point.
From experiments, we already know that photons (the corresponding particles) are derived from the field, in fact, they correspond to (quantized) excitations of the field. This can, e.g., be seen in laser experiments, in which the power incident on a screen is recorded. As the intensity of the laser is lowered, the power arrives in discrete, "quantized packages", the photons.
In order to be compatible with special relativity, we need to build a theory which has Lorentz invariance built into it, just like the relativistic formulation of electrodynamics has.

What will this type of theory be able to describe?

- elementary particles and their interactions, in particular the Standard Model of particle physics.

- any setting in which particle number is not conserved, e.g., condensed-matter-systems in which we are interested in effective (not fundamental) excitations, such as, e.g., phonons, or Cooper-pairs in superconductivity.

- if the energy of the system is low enough, the formalism that we are developing is even sufficient to understand the quantum properties of gravity.

Note: in our current understanding of cosmology, the origin of all structures in the universe (galaxies, galaxy clusters ...) are quantum fluctuations of the fundamental fields in the early universe. Ultimately, we thus owe our existence to the physics of QFT!

1.2 Why learning quantum field theory is hard

Quantum field theory is not an easy subject. This has several reasons. First, the quantities that we are dealing with are often abstract and more difficult to develop an intuitive understanding of than, for instance, systems in classical mechanics. Second, we need to develop an entirely new formalism to describe quantum fields, in which we bring together classical field theory and quantum theory. In other words, we are learning a (mathematical) language in which to describe the systems that we are interested in, and, just like with any other new language, learning it can be hard and it takes some time until the concepts start to feel familiar and intuitive.
However, you should not feel discouraged by this or think about giving up. Rather, if you have questions and/or doubts, bring them up with the lecturer (either after the lecture, or by email to eichhorn@thphys.uni-heidelberg.de) or to your tutor, or to the head tutor, Zois Gyftopolous (gyftopolous@thphys.uni-heidelberg.de). The whole team of lecturer and tutors is here to support you in learning and understanding quantum field theory!

1.3 Why learning quantum field theory is absolutely worth it

Quantum field theory provides the framework for the most advanced and deepest understanding of fundamental physics that we have. Therefore, it is like a key with which we can unlock fascinating insights into elementary particles and their properties. Thus, some of the highlights that await us this term are:

- understanding how powerful symmetries are and how we can deduce properties of elementary particles from an understanding of the Lorentz group and how we can deduce the existence of the electromagnetic field from thinking about symmetries

- understanding were the Pauli principle for Spin-1/2-particles comes from

- understanding why antiparticles must exist in order for causality to not be violated

- understanding that the vacuum is not a boring state of “nothing”, but is a highly non-trivial state which results in a force between conducting plates (“Casimir force”) or the scattering of photons off each other (unlike in classical electrodynamics, in which the equations of motion for the gauge field are linear and electromagnetic waves do not interact)

- and much more!

1.4 Classical field theory

We have already emphasized the role and importance of *symmetries*, so we will spend some more time developing the *mathematics of symmetries*, namely groups and their representations. First, however, we need to establish some of the notions that form the basis of this course, namely fields and their classical description.

A field takes a value at each spacetime point. Examples that you may already know include

- the E- and B-field, $E(\vec{x},t), B(\vec{x},t)$, which are 3-vectors.
- the density in hydrodynamics, $\rho(\vec{x},t)$, which is just a one-component function.
- the gauge field $A_\mu(\vec{x},t)$ in electrodynamics, which is a 4-vector.

To describe their dynamics, we start from an action S, which is a *functional*, i.e., its argument is a function (and it maps to the real numbers).
For instance, in the relativistic way of writing electrodynamics, we have

$$\begin{aligned} S[A] &= \frac{1}{4}\int d^4x F_{\mu\nu}F^{\mu\nu}, \\ &= \int d^4x \mathcal{L}_{\mathrm{ED}}, \qquad\qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \end{aligned} \tag{1}$$

We denote functionals with square brackets around their arguments, which are functions. $\mathcal{L}_{\mathrm{ED}}$ is the *Lagrange density.* It is not a functional, because it does not depend on the full function (in this case, the field at all spacetime points), but is just a function of the spacetime-coordinates, through its dependence on the field at a point.
To establish some of the key notions, we will use a scalar field, conventionally denoted by $\phi(\vec{x},t)$. An example for scalar fields relevant in nature is the Higgs field in the Standard Model; hypothetical scalar fields include the inflaton field (that drives the (conjectured) inflationary phase in the early universe), and proposals for dark matter (e.g., the axion, which is, to be more precise, a pseudoscalar). Scalars that can be collective degrees of freedom also play a role in many condensed-matter systems, starting from the Ising model.
The Lagrange density $\mathcal{L}$ depends on the field $\phi(\vec{x},t)$ and its derivatives, $\partial_\mu\phi(\vec{x},t)$, $\partial_\mu\partial_\nu\phi(\vec{x},t)$ ... and is a priori completely arbitrary. We will make two assumptions:

- the Lagrange density is *local*, i.e., it depends on fields and their derivatives at one point and it only depends on a finite number of derivatives. (We call this local, because a derivative always compares a field at a point to its (infinitesimally removed) neighboring point. An infinitely high power of derivatives thus involves fields a finite distance apart.)
  This has two motivations: First, observationally, local interactions seem to describe nature very well; e.g., in the LHC detectors, one can see that particles interact locally. Second, non-local interaction may get into conflict with causality, because non-localities may mean interactions at spacelike distances.
- We assume that the Lagrange density does not have higher than second derivatives in time. The reason is Ostrogradsky's theorem, which is a theorem in classical mechanics and states that, (under a non-degeneracy condition), a Hamiltonian that contains higher-than-second-order time derivatives is unbounded from below. This may- but need not!- make the the-

ory dynamically unstable. Because this theorem implicitly underlies the formulation of Lagrangians in many settings (classical mechanics, classical field theory, quantum field theory), we will take a closer look at it in the exercises.[1]

The Lagrangian

$$L = \int d^3x \mathcal{L}, \tag{2}$$

is the spatial integral of the Lagrange density. We will often work with $\mathcal{L}$, because it makes the equal treatment of space and time, that we want in a relativistic theory, manifest. It is often called "the Lagrangian" in a slight abuse of naming conventions.

$\mathcal{L}$ consists of two parts, a kinetic part, T, that depends on derivatives, and a potential, V,

$$\mathcal{L} = T - V. \tag{3}$$

We will often focus on

$$T = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi, \tag{4}$$

and

$$V = \frac{1}{2}m^2\phi^2 + \lambda\phi^4, \tag{5}$$

where in V we assumed that we can Taylor-expand $V(\phi)$ around a minimum ϕ_0 and we can set $\phi_0 = 0$ and $V(\phi_0) = 0$ without loss of generality. We further assume a symmetry $\phi \to -\phi$, so that there is no ϕ^3 present, which would render $V(\phi)$ unbounded from below. We call m the mass, because we will see that the equations of motion imply $p^2 = m^2$ for the square of the four-momentum, if the term $m^2\phi^2$ is present in the Lagrangian. The quartic term, $\lambda\phi^4$ leads to non-linear equations of motion, i.e., it describes *interactions* of the field (and the corresponding particles) with itself. The strength of these interactions is parameterized by the coupling λ. In the next few lectures, we focus on just the mass term.

Our choice of T requires a bit more justification: The kinetic part describes how the field changes in space and time, thus it must contain a derivative, and $\partial_\mu \phi$ is the building block to use. In order to have a Lorentz-invariant expression, we must contract the open index and the only other 4-vector we have is another derivative. Thus, up to rescalings of the term, we have a unique lowest order action in ϕ

$$S = \int d^4x \left(\frac{1}{2}\partial_\mu \phi\, \partial_\nu \phi\, \eta^{\mu\nu} - \frac{m^2}{2}\phi^2\right), \tag{6}$$

where $\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$ in our conventions, which most QFT books use. Many GR books use $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$. The overall sign is pure convention; the difference in signs between the time part and the spatial part is physics.

Mini-Exercise 1.1. We made the statement that we can set a constant and a linear term in $\mathcal{L}$ to zero without loss of generality. For the constant term, this is because the equations of motion follow from minimizing the action and the field value that minimizes S does not depend on whether or not there is a constant shift in S.

[1] You have probably encountered or will encounter many examples where the Lagrangian does not have higher than second order time derivatives. Electrodynamics is one example, General Relativity another, and classical mechanics is full of examples. Note however that there are subtleties and there are counterexamples to the intuition that a Hamiltonian that is unbounded from below leads to instabilities.

For the linear term, we can always remove it by a change of our field variable (which you can think of as analogous to a change in coordinates in class. mech.)
Show this! Start with

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - C\phi - \frac{1}{2}m^2\phi^2. \tag{7}$$

Define $\varphi = \phi + \gamma$. What is the choice of γ, such that

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 + const\,? \tag{8}$$

The equations of motion follow from extremizing the action, i.e., we perform a variation of the action (i.e., a variation of the field, $\phi \to \phi + \delta\phi$, by some arbitrary amount $\delta\phi$). We set the variation of the action to zero, just like, when we are searching for the minimum of a function, we are setting its first derivative (analogous to the variation of the argument of the function) to zero:

$$\begin{aligned} 0 = \delta S = \delta \int d^4x \left(\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2\right) \\ = \int d^4x \big((\partial_\mu\phi)\eta^{\mu\nu}(\partial_\nu\delta\phi) - m^2\phi\delta\phi\big) \\ = \int d^4x \big(-(\partial_\nu\partial_\mu\phi)\eta^{\mu\nu}\delta\phi - m^2\phi\delta\phi\big) \\ = \int d^4x \big(-(\partial_\nu\partial_\mu\phi)\eta^{\mu\nu} - m^2\phi\big)\delta\phi, \end{aligned} \tag{9}$$

where in the second-to-last step we used partial integration and where we assume that $\delta\phi = 0$ at $x \to \pm\infty$. Because $\delta\phi$ is an arbitrary variation, to satisfy Eq. (9), the factor $-\partial_\nu\partial_\nu\phi\eta^{\mu\nu} - m^2\phi$ must be zero.
This is the Klein-Gordon equation,

$$\boxed{\partial^2\phi + m^2\phi = 0,} \tag{10}$$

with $\partial^2 = \partial_\mu\partial_\nu\eta^{\mu\nu}$. The Klein-Gordon equation is a relativistic, massive wave equation.
For the Lagrangian, $\delta S = 0$ translates into the Euler-Lagrange equations

$$\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial_\mu\phi}\right) = 0. \tag{11}$$

The solutions to the equations of motion are spanned by plane waves,

$$\phi(x) = \phi_0\cos(kx), \qquad \text{(assuming } \phi(x) = \phi(-x)\text{)} \tag{12}$$

with the shorthand $kx = k_\mu x^\mu$ and the relativistic, massive dispersion-relation $k_\mu k^\mu = k^2 = m^2$.

Later on, a starting point for one quantization scheme (path-integral quantization) will be the action, but the starting point for another quantization scheme (canonical quantization) will be the Hamiltonian.
Just as in classical mechanics, where we define $p = \frac{\partial L}{\partial \dot{q}}$, the canonically conjugate momentum, and $H(p,q) = p\dot{q} - L$, in quantum field theory we define $\pi(\vec{x})$, the canonically conjugate field. (Note: it is the canonically conjugate field to ϕ, but has nothing to do with the momentum of the particles that we will describe. It is sometimes called the (canonically conjugate) momentum field, because it arises in the generalization of the Hamiltonian formalism to QFT and it generalizes the

momentum of a particle, which is the canonically conjugate variable to the position.) Its definition is

$$\pi(\vec{x}) = \frac{\delta L}{\delta \dot{\phi}(\vec{x})}, \tag{13}$$

which is a functional derivative, i.e., a derivative with respect to a function. Just like $\frac{\partial x}{\partial x} = 1$, we have

$$\frac{\delta \phi(\vec{x})}{\delta \phi(\vec{y})} = \delta^3(\vec{x} - \vec{y}). \tag{14}$$

Thus, for the Lagrangian in Eq. (6), we obtain

$$\begin{aligned}\pi(\vec{x}) &= \frac{\delta}{\delta \dot{\phi}(\vec{x})} \int d^3y \left(\underbrace{\frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\left(\vec{\nabla}\phi\right)^2}_{\partial_\mu \phi \partial^\mu \phi} - \frac{1}{2} m^2 \phi^2 \right) \\ &= \int d^3y \left(\dot{\phi}\, \delta^3(\vec{x} - \vec{y}) \right) = \dot{\phi}(\vec{x}).\end{aligned} \tag{15}$$

Thus, to calculate the Hamiltonian, we can use that $\dot{\phi}$ can be substituted by π. We obtain the Hamiltonian of the system as

$$\begin{aligned}H &= \left(\int d^3x\, \pi \dot{\phi} \right) - L \\ &= \int d^3x \left(\pi^2 - \left(\frac{\pi^2}{2} - \frac{1}{2}\left(\vec{\nabla}\phi\right)^2 - \frac{m^2}{2}\phi^2 \right) \right) \\ &= \frac{1}{2} \int d^3x \left(\underbrace{\pi^2 + \left(\vec{\nabla}\phi\right)^2 + m^2 \phi^2}_{2\mathcal{H}} \right) \\ &= \int d^3x\, \mathcal{H},\end{aligned} \tag{16}$$

where we defined the Hamiltonian density $\mathcal{H}$.

2 The importance and the mathematics of symmetries

Useful literature for this chapter is the following: There are books on group theory in physics and more specifically particle physics, e.g., "Group theory in physics" by Wu-Ki Tung and "Lie algebras in particle physics" by Howard Georgi.

QFT books also cover discussions of symmetry groups, for instance: Schwartz, chapter 2, covers the basics of Lorentz transformations, and group theory basics for the Lorentz group are discussed in 10.1. The Lorentz group and its Lie algebra generators are also discussed in Srednicki, chapter 2. Gelis (chapter 7.1) summarizes Lie groups and Lie algebras.

Symmetries are one of the most important foundational elements in QFT. This becomes obvious from many examples:

i) In particle physics, the various mesons and baryons are organized into sets, e.g., the eight lightest mesons are grouped into the meson octet according to the "eightfold way" which is based on a so-called "$SU(3)$ flavor symmetry". Historically, this type of organization into sets according to symmetries was central in predicting new particles.

ii) You might have heard that the Standard Model is an $SU(3) \times SU(2) \times U(1)$ gauge theory. Specifying this symmetry already fixes a large part of the Standard Model particle content and the allowed interactions between particles.

iii) In condensed matter, phase transitions are associated with spontaneous *breaking of symmetries*. For instance, in a ferromagnet, at high enough temperature, there is no macroscopic magnetization, which means that there is full rotational symmetry for each of the microscopic spin vectors. At low temperature, in the magnetized phase, rotational symmetry is broken, because the macroscopic magnetization spontaneously selects one spatial direction. More generally, by knowing the symmetries that the degrees of freedom in a condensed-matter system obey, we can already figure out which phases and phase transitions there could be.

iv) Lorentz symmetry (or its generalization, Poincaré symmetry, which adds translations (in space and in time)) determine much of the properties of elementary particles and their interactions and much of the mathematical structure of QFT.
For instance, the fact that we characterize elementary particles by their mass and their spin is a direct consequence (as we will work out) follows from considering the Poincaré group.

v) ...

This motivates us to dive into the *mathematics of symmetries*, because this appears to be the language in which large parts of nature can be described.

2.1 Symmetries are described by groups

It turns out that there exists a mathematical structure that is exactly adapted to formalizing symmetries, and that is a group.

Definition of a group:

A group G is a set of elements $G_i \in G$, together with a "multiplication" $\cdot$, such that

$$G_i \cdot G_j = G_k, \qquad G_k \in G \quad \forall\, G_i, G_j \in G. \tag{17}$$

This means that we can combine two elements of the group by the multiplication and we obtain another element of the group. The multiplication law satisfies

- associativity

$$(G_i \cdot G_j) \cdot G_k = G_i \cdot (G_j \cdot G_k) \quad \forall\, G_i, G_j, G_k \in G. \tag{18}$$

- $\exists$ identity element E, s.t.

$$G_i \cdot E \in G \qquad \forall\, G_i \text{ and } E \in G. \tag{19}$$

- inverse element

$$\forall\, G_i \in G\, \exists\, G_i^{-1} \in G, \text{ s.t. } G_i \cdot G_i^{-1} = E. \tag{20}$$

Note that the identity element is unique, as is the inverse for each element.
Let's parse this definition and the intuition behind the various requirements in physics language, using rotations as an example and thinking of a spherically symmetric system:

- Two rotations can be performed consecutively, yielding a third rotation (about a different axis). This is the multiplication law which allows us to combine group elements into new group elements.
- When three rotations are performed, either the 1st and 2nd or 2nd and 3rd can be combined, such that the consecutive execution of the three of them is equal in any of the two combinations. (Note that we must not reverse the *order* of the three rotations, because the group is not commutative.)
- There is an identity element, namely rotation by 0° (or no rotation).
- For each rotation, we can reverse the sense of rotation to rotate back, such that the combination of rotation and inverse rotation yields no rotation.

You may already know that rotations can be represented by matrices, such that, e.g., the identity is the unit matrix and the inverse element is the inverse matrix.

We will encounter two *mathematically distinct* sets of groups that encode symmetries in QFT:

1) discrete groups (with a finite set of elements), for instance reflections about a plane (has three elements: the reflection, its inverse, and the identity).
2) continuous groups, which are Lie groups. The rotation group is an example. It is continuous, because it has infinitely many group elements (rotations by different angles) and "neighboring" rotations only differ infinitesimally.

We will also encounter three *physically distinct* types of groups[2]

[2] There is a theorem, the Coleman-Mandula theorem, that says that, under some assumptions, there are no symmetry groups that mix spacetime symmetry transformations with internal symmetry transformations. The realization that, by violating the assumptions, one can get around this theorem, and is then required to introduce so-called "super-partners" led to the development of supersymmetry, which we will not treat in this course, but which is a very interesting mathematical developments worth understanding. In nature, supersymmetry is realized in some low-dimensional settings in condensed-matter theory, but does not appear to be realized in particle physics. It is, however, instrumental in one approach to quantum gravity, namely in string theory.

a) spacetime symmetry groups, which can either be continuous (like the Lorentz group, $SO(1,3)$), or discrete (like time-reversal symmetry which maps the time t to $-t$)

b) internal symmetry groups, where “internal” here means that the symmetry does not act on space and time (like, e.g. the Lorentz group), but only on the field. These come in two different versions:

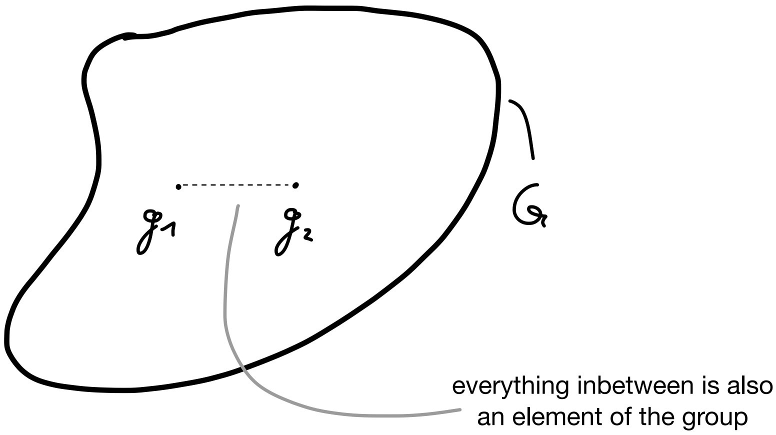
   i) global internal symmetries (like the $\mathbb{Z}_2$-symmetry $\phi \to -\phi$ that we imposed on scalar field theory to ensure that there is no ϕ^3 term in the scalar potential, or the $SO(3)$ symmetry that is imposed on the scalar field in the Heisenberg model that describes phase transitions in certain materials).
   Global means that the symmetry transformation is the same for the field at all spacetime points.

   ii) local internal symmetries (like the $U(1)$ gauge symmetry of electromagnetism).
   Local here means that the symmetry transformation can be different at different spacetime points (even if it doesn’t act on the spacetime itself).

Some of these notions may seem a little abstract at the moment. They will become clearer as we develop our understanding of group theory and come up with examples.
The most relevant groups for us will be *Lie Groups*.

2.2 Lie groups

These are groups in which the group elements form not just a set, but a differentiable manifold (which is a collection of points such that each point has an open neighborhood that is equivalent to $\mathbb{R}^n$ and which can be covered by coordinate charts that overlap partially).
This means that the group is *continuous*, such that you can always find a group element infinitesimally close to any given element. Intuitively, we can see directly that the group of rotations should be such a continuous group, because we can always rotate by an arbitrarily small amount and thus find rotations which are only infinitesimally different from each other.



Examples:

- $U(1)$ is the group of all unitary 1×1 matrices, i.e.,

$$G = e^{i\alpha}, \qquad \alpha \in \mathbb{C}. \tag{21}$$

  The corresponding manifold is the circle (of radius 1) in the complex plane.
  As a global symmetry, the phase α of the transformation does not depend on the spacetime point. As a local symmetry, α is upgraded to a function $\alpha(x^\mu)$. We will explore the consequences of this soon. In fact, this group determines the properties of photons and their interactions with charged particles.

- $SU(2)$ is the group of 2×2 unitary matrices with determinant 1. The corresponding manifold is the 3-sphere, S^3.
  To see this, we write

$$U^\dagger U = \mathbb{1} \implies U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \qquad \text{with } |a|^2 + |b|^2 = 1 \text{ for } a, b \in \mathbb{C} \tag{22}$$

  (Check:

$$U^\dagger = \begin{pmatrix} a^* & -b \\ b^* & a \end{pmatrix} \quad \text{and} \quad U^\dagger U = \begin{pmatrix} |a|^2 + |b|^2 & \cancel{a^* b - b a^*} \\ \cancel{b^* a - a b^*} & |a|^2 + |b|^2 \end{pmatrix} = \mathbb{1}) \tag{23}$$

  Now we write both complex numbers in terms of real and imaginary part,

$$a = x + iy, \qquad b = z + it \tag{24}$$

  $\implies |a|^2 + |b|^2 = |x|^2 + |y|^2 + |z|^2 + |t|^2 = 1$ parametrizes the group manifold $SU(2)$, where $x, y, z, t \in \mathbb{R}$.

This is the equation defining a unit 3-dimensional sphere embedded in 4-dimensional space, i.e., S^3.
SU(2) is the symmetry group determining the properties of the weak gauge bosons (W^+, W^- and Z) and their interactions with the fermions in the Standard Model.

With a Lie group comes a Lie algebra $\mathfrak{g}$, $\mathrm{Lie}(G) = \mathfrak{g}$. Knowing about the Lie algebra is useful, because all properties of the Lie group follow from knowing the so-called *generators* of the Lie algebra and their commutation relations.
A *Lie algebra* is a vector space $\mathfrak{g}$ with a bilinear, antisymmetric map:

$$\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, \qquad (a,b) \mapsto [a,b] = -[b,a] \tag{25}$$

(that we suggestively write in the same notation that we use for the commutator) that satisfies the Jacobi identity

$$[a,[b,c]] + [b,[c,a]] + [c,[a,b]] = 0. \tag{26}$$

We will only need matrix groups and matrix algebras. For matrix Lie groups, the relation between group and algebra is given through the *exponential map*: For $a_i \in \mathfrak{g}$,

$$G_i = \exp(a_i), \tag{27}$$

(defined through its Taylor series) is a group element. Each group element (in the so-called identity component of G) can be written in such a way. For $0 \in \mathfrak{g}$, $\mathbb{1} = \exp(0) \in G$.
We can find a basis in $\mathfrak{g}$ and these elements of the Lie algebra are called the *generators*. Having this basis of generators, we can construct every group element through the exponential map.
Example: Rotation group $SO(3)$
$SO(3)$ is the group of special orthogonal $3x3$ matrices, i.e., matrices which are orthogonal, so $\mathrm{Rot}\,\mathrm{Rot}^\top = \mathbb{1}$, where $\mathrm{Rot}^\top$ denotes the transposed matrix, and special, i.e., their determinant is $+1$. They describe rotations, because we can check that the requirement that a rotation leaves the length of a vector invariant requires $\mathrm{Rot}\,\mathrm{Rot}^\top = \mathbb{1}$. To check this, consider a spatial vector, with components x^i. Under a rotation, it is mapped to

$$x^i \to x^{i'} = \mathrm{Rot}^i{}_k\, x^k. \tag{28}$$

We require that its length stays invariant, so that

$$x^i\, x^j\, \delta_{ij} = x^{i'} x^{j'} \delta_{ij} = Rot^i{}_k\, \mathrm{Rot}^j{}_l\, x^k\, x^l \delta_{ij}. \tag{29}$$

Thus, $\mathbb{1} = \mathrm{Rot}^i{}_k\, \mathrm{Rot}^j{}_l\, \delta_{ij} = \mathrm{Rot}^T\,\mathrm{Rot}$. This is in particular realized by matrices of the form

$$\mathrm{Rot}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \tag{30}$$

and analogously for rotations about the y and the z-axis.
Claim: $\mathrm{Lie}(G) = \{\text{antisymmetric}\, 3 \times 3 \text{ matrices}\}$

Mini-Exercise 2.1. Check that $\mathrm{Rot}\,\mathrm{Rot}^\top = \mathbb{1}$, as required for $\mathrm{Rot} \in SO(3)$, is realized by $\mathrm{Rot} = \exp(T)$, if $T_{ij} = -T_{ji}$.

Let's see how we can reconstruct the group elements, i.e., the Rot matrices, from the Lie algebra generators. An antisymmetric 3×3 matrix with real components (so that Rot is real), has 3 independent components, so we have three basis elements

$$T_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{31}$$

and analogously for the other two *generators* of the Lie algebra. Now we can write a rotation about the x-axis as

$$\mathrm{Rot}_x = \exp\left(\theta\, T_x\right) = \mathbb{1} + \theta \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} + \mathcal{O}(\theta^2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\theta \\ 0 & \theta & 1 \end{pmatrix} + \mathcal{O}(\theta^2), \tag{32}$$

which is clearly the infinitesimal version of the rotation matrix given above.

What will be crucial in our construction of QFT is the notion of *representations* of groups and algebras. For instance, we will construct the spin-0, spin-1/2 and spin-1 representations of the Lorentz group to describe the Higgs field, the electron and the photon in the Standard Model, or various excitations in condensed-matter systems.

Intuitively, a representation is a set of objects which satisfy the same multiplication rules as the abstract group elements, i.e., they are often matrices, for which the multiplication satisfies the combination rules that the group elements satisfy.

More formally, a representation R of a group is a map $G \xrightarrow{R} \mathrm{GL(V)}$ (where GL(V) are the general linear transformations on a vector space), such that $R(\mathbb{1}) = \mathbb{1}$ and $R(gh) = R(g)R(h)$. (In other words, R is a group homomorphism from G to GL(V).) Loosely speaking, we find matrices which represent the symmetry operators.

Examples: representations of the rotation group $SO(3)$

- trivial representation: on scalar quantities $R(\mathrm{Rot}) = \mathbb{1}$, no rotation.
- vector representation: on a vector, $V = \mathbb{R}^3$, $R(\mathrm{Rot}) = \mathrm{Rot}$. This is the so-called *fundamental representation*, in which the rotation matrices take the form that defines the group, namely $3x3$ orthogonal matrices with unit determinant.
- tensor representation: on a tensor, $V = \mathbb{R}^3 \times \mathbb{R}^3$, $R(\mathrm{Rot}) = \mathrm{Rot} \otimes \mathrm{Rot}$, because $T_{ij} \mapsto R_i{}^k R_j{}^l T_{kl}$.

Similarly, Lie algebras have representations and from a representation of a Lie algebra, we can always construct the associated representation of the group (by using the exponential map). Thus, we will sometimes be a bit sloppy and switch back and forth between algebra and group.

Let's make all of this more concrete by looking at the Lorentz and the Poincaré groups as our examples. These are our most important examples, because these encode the fundamental symmetries of spacetime, on which we are constructing our quantum field theory.

2.3 Lorentz transformations and the Lorentz group

Lorentz transformations, abstractly denoted by Λ, act on 4-vectors that denote the spacetime location of an event [3], i.e.,

$$x^\mu = \begin{pmatrix} t \\ \vec{x} \end{pmatrix} \tag{33}$$

as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{34}$$

where the defining equation for a Lorentz transformation is

$$\Lambda^\mu{}_\rho \eta_{\mu\nu} \Lambda^\nu{}_\sigma = \eta_{\sigma\rho}. \tag{35}$$

This equation says that the Minkowski metric is left invariant under Lorentz transformations, which implies that scalar products built with this metric are invariant under Lorentz transformations.
Because η is the Minkowski metric, the Lorentz group is $SO(3,1)$, and contains boosts and spatial rotations, instead of being $SO(4)$, the group of rotations of 4-dimensional space (which Eq. (35) would define for $\eta_{\mu\nu} \to \delta_{\mu\nu}$).
From Eq. (35), we have that $\Lambda_\nu{}^\mu = \eta_{\mu\kappa}\eta_{\lambda\nu}\Lambda^\lambda{}_\kappa$ is the inverse Lorentz transformation. This is easiest to see by writing Eq. (35) in matrix notation, where it reads

$$\Lambda^T \eta \Lambda = \eta, \tag{36}$$

where the first Λ is transposed, in order for the index contraction in Eq. (35) to match index contraction for matrix multiplication. From Eq. (36), we then have that

$$\Lambda^{-1} = \eta^{-1}\Lambda^T \eta, \tag{37}$$

which, in index notation, becomes

$$\left(\Lambda^{-1}\right)^\mu{}_\nu = \eta^{\mu\kappa}\left(\Lambda^T\right)_\kappa{}^\lambda \eta_{\lambda\nu} = \eta^{\mu\kappa}\Lambda^\lambda{}_\kappa \eta_{\lambda\nu} = \Lambda_\nu{}^\mu. \tag{38}$$

When acting on 4-vectors, Λ are in their fundamental representation, which you can think of as the representation that is used to define the group. How is the associated Lie algebra $\mathfrak{so}(3,1)$ characterized?
We use that we can expand the exponential map to first order in the Lie algebra elements, if we consider an infinitesimal transformation. For the fundamental representation

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu + \mathcal{O}(\omega^2) \tag{39}$$

for an infinitesimal transformation. Eq. (35) then implies a property of the ω's:

[3] Note that we use units in which $c = 1$.

Mini-Exercise 2.2. What holds for $\omega^\mu{}_\nu$, such that (35) holds?

It holds that

$$\omega_{\mu\nu} = -\omega_{\nu\mu}, \tag{40}$$

i.e., $\omega_{\mu\nu}$ is an antisymmetric 4×4 matrix and therefore has 6 independent components that can be nonzero. Depending on which components we choose to be nonzero, we obtain a different group element of the Lorentz group.

Let us consider an example: We choose $\omega^{12} = -\omega^{21} = \theta$ and set all other components of ω to zero. Note that we have to be careful with the upper and lower indices on ω, so there will be an $\eta_{\mu\nu}$ that will make an appearance below. We obtain that

$$\begin{aligned}\Lambda^\mu{}_\nu &= \delta^\mu_\nu + \omega^{\mu\rho}\eta_{\rho\nu} \\ &= \mathbb{1} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\theta & 0 \\ 0 & \theta & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.\end{aligned} \tag{41}$$

We observe that this generates nothing but a (infinitesimal) rotation of the four-vector x^μ about the z-axis by an angle θ. We also note that the contraction $\omega^{\mu\rho}\eta_{\rho\nu}$ essentially flips the sign, i.e., $\omega^1{}_2 = -\omega^{12}$ and similarly $\omega^2{}_1 = -\omega^{21}$.

Similarly, if we choose $\omega^{01} = -\omega^{10} = \theta$, we obtain

$$\Lambda^\mu{}_\nu = \mathbb{1} + \begin{pmatrix} 0 & \theta & 0 & 0 \\ \theta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{42}$$

which we can recognize as an infinitesimal boost along the x-axis, with θ being the rapidity, $\tanh\theta = v/c$. In this case, we have used that $\omega^0{}_1 = \omega^{01}\eta_{11} = -\omega^{01}$ and $\omega^1{}_0 = \omega^{10}\eta_{00} = -\omega^{10}$.

These examples help us to see that the six entries in $\omega^{\mu\nu}$ which can be nonzero, select, which among the six possible "basis" transformations (3 rotations along the 3 spatial axis, and 3 boosts along these axis), can be performed and by which amount the physical system is rotated and/or boosted. If we choose more than one component of $\omega^{\mu\nu}$ to be non-zero, we get the corresponding combination of these "basis" transformations.

For a general representation $U(\Lambda)$ of the Lorentz transformation Λ, we have that

$$U(\mathbb{1} + \omega) = \mathbb{1} + \frac{i}{2}\omega_{\mu\nu}M^{\mu\nu} + \mathcal{O}(\omega^2). \tag{43}$$

In this expression, the $\omega^{\mu\nu}$ still selects, which transformation is performed and determines the "amount" of the transformation, but the "basis transformations" are now encoded in the $M^{\mu\nu}$. The $M^{\mu\nu}$ are called the <u>*generators*</u> of the Lorentz group, and there are six of them, representing the three independent rotations and three independent boosts. We have that

$$J^i = \frac{1}{2}\epsilon^i{}_{jk}M^{jk} \qquad \text{generates rotations}$$
$$K^i = M^{i0} \qquad \text{generates boosts,}$$

where ϵ_{ijk} is the Levi-Civita symbol, which is fully antisymmetric under permutations of its indices and $\epsilon_{123} = 1$. Thus, M^{12}, M^{13} and M^{23} generate rotations (along the z-, y- and x-axis, respectively) and M^{01}, M^{02} and M^{03} generate boosts along the x-axis, y-axis and z-axis, respectively.

A Lorentz transformation can act on many different objects, not just on four-vectors. In particular, we will later in the course encounter *spinors*, which are objects that have spinor indices. These are indices, i.e., a spinor is a collection of functions, but they are *not* spacetime indices. Therefore, to have a Lorentz transformation act on a spinor, the $M^{\mu\nu}$ need to carry the appropriate indices, i.e., *each of the six $M^{\mu\nu}$'s, such as M^{01}, M^{12} etc., must be a* matrix *with indices in the space that it acts on.*

This is somewhat abstract at this moment, so in order to make it less abstract, we consider the case in which the Lorentz transformation acts on a four-vector. We already know that we can write this in the form of Eq. (39), but now we want to understand how to write it in the form Eq. (43), in which the generators appear explicitly. In fact, for the fundamental representation of the Lorentz group, we have that

$$(M_{\mu\nu})_{\kappa\lambda} = -i\left(\eta_{\mu\kappa}\eta_{\nu\lambda} - \eta_{\nu\kappa}\eta_{\mu\lambda}\right). \tag{44}$$

By plugging this into Eq. (43), we get back Eq. (39).

While it seems unnecessarily complicated to introduce the M's for the action on 4-vectors, the main point about Eq. (43) is that it is general; it describes the action of a Lorentz transformation on any object.

The *defining property* of the generators of the Lorentz group is that they satisfy a commutation relation. The abstract definition of the Lie algebra of the Lorentz group is through this commutation relation:

The Lie algebra of the Lorentz group $SO(3,1)$, is defined by the commutator relation of its generators, which is

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(\eta^{\mu\rho}\, M^{\nu\sigma} - \eta^{\nu\rho}\, M^{\mu\sigma}) - i(\eta^{\mu\sigma}\, M^{\nu\rho} - \eta^{\nu\sigma}\, M^{\mu\rho}). \tag{45}$$

You will derive this commutation relation in the exercises. You can think of the Lorentz group as being defined by this commutation relation. When we talk about different elementary particles and different fields, they all arise from thinking about different representations of the Lorentz group, i.e., many properties of elementary particles follow from this commutation relation above. At this stage, this is still a rather abstract notion, but over the course of this course, we will see the commutation relation Eq. (45) "unfold its power".

2.4 Poincaré group and why we classify particles by their mass and spin

We classify elementary particles by their mass and spin, plus quantum numbers associated to internal symmetries. For instance, we describe the electron as a particles with rest-mass 511 keV and spin-1/2 (and electric charge -1). Why do we do so? Is it just a conventional choice and

we could be using some completely different characteristics? The answer is no. There is a deep mathematical reason and it has to do with the structure of the *Poincaré-group* - a generalization of the Lorentz group - and its so-called *Casimir-operators.*
The Poincaré group is an extension of the Lorentz group which, in addition to boosts and rotations, contains translations, under which $x^\mu \mapsto x^\mu + a^\mu$. A transformation by an element of the Poincaré group can be written as $x^\mu \mapsto \Lambda^\mu{}_\nu x^\nu + a^\mu$.
This is the full symmetry that 3+1-dimensional Minkowski spacetime enjoys.
An infinitesimal translation in a general representation can be written as

$$U(a) = \mathbb{1} + ia_\mu P^\mu, \tag{46}$$

where P^μ is the generator of translations. By Noethers theorem, P^μ will be identified as the 4-momentum in the corresponding representation. Its commutation relations with the other generators of the Poincaré group are

$$[P^\mu, M^{\rho\sigma}] = i(\eta^{\mu\sigma} P^\rho - \eta^{\mu\rho} P^\sigma) \tag{47}$$

$$[P^\mu, P^\nu] = 0. \tag{48}$$

Now let us consider some state of n particles, which transforms under actions of the Poincaré group. Under such transformations, its properties, such as its 4-momentum, change.
However, the Poincaré group has two *Casimir invariants.* These are (in the simplest case) quadratic combinations of generators, which commute with all other generators. Therefore, their eigenvalues are unchanged under the action of group elements and they provide *invariant* characterizations of particles.
$P^2 = P_\mu P^\mu$ is the first Casimir invariant and $W^2 = W_\mu W^\mu$, with $W_\mu = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma} M^{\nu\rho} P^\sigma$ the Pauli-Lubanski-pseudovector, is the second.

Mini-Exercise 2.3. Show that P^2 commutes with all generators of the Poincaré group.

P^2 acting on a state with some 4-momentum yields the eigenvalue m^2, i.e., because P^2 is a Casimir operator of the Poincaré group, we label elementary particles by their rest mass.
But what is the physical meaning of W^2?

$$W^2 = W_\mu W^\mu = \frac{1}{4}\epsilon_{\mu\nu\rho\sigma} M^{\nu\rho} P^\sigma \epsilon^\mu{}_{\chi\lambda\tau} M^{\chi\lambda} P^\tau. \tag{49}$$

Let's consider this in a massive particles rest frame (massless particles are a separate case and we will get to them later).
Then $P \to (m, \vec{0})$ and $W^0 = 0$. This holds, because ϵ is totally antisymmetric and because the only non-zero component of P is P^0.

$$W^i = -\frac{1}{2}\epsilon^i{}_{\mu\nu 0} M^{\mu\nu} P^0, \tag{50}$$

here μ, ν must be spatial indices, but $\neq i$. Therefore, $\epsilon^i{}_{jk0} = \epsilon^i{}_{jk}$, the 3d Levi-Civita symbol. Thus,

$$W^i = -J^i P^0 = -mJ^i \implies W^2 = m^2 \vec{J} \cdot \vec{J}. \tag{51}$$

Now we need to interpret which angular momentum it is that shows up here. Which angular

momentum does an elementary particle have? *Intrinsic* angular momentum, i.e., spin. You might remember from QM, that the eigenvalues of $\vec{J}^2$ are $s(s+1)$, with s the spin.
$\Longrightarrow$ Because W^2 is the 2nd Casimir operator of the Poincaré group, we label massive elementary particles by their spin.

We have thus come to our first concrete result from our more abstract consideration of group theory:
We have learned that there is a reason why we label elementary particles by mass and spin. This is not an arbitrary choice, but a direct consequence of the fundamental symmetry-structure of Minkowski spacetime and the properties of the underlying Poincaré group.

Next, we may wonder, what spin values[4] are allowed? Can we have elementary particles with spin 0? spin 1/2? Spin 1? What about non-half-integers? Is there a particle with spin 2/3? or spin M? To figure this out, we will classify the representations of the Lorentz group. This will determine what type of fields we will focus on for the rest of the course.[5] Generally, for a field with a general Lorentz index A (could be a 4-vector index, or two 4-vector indices, such that the field is a tensor, but we'll also encounter spinor indices, which label the components of a spinor, but are *not* spacetime indices), $\phi_a(x)$, we have

$$\phi'_a(x) = L_a{}^b(\Lambda)\phi_b(\Lambda^{-1}x). \tag{52}$$

The matrices $L_a{}^b(\Lambda)$ form a *representation* of the Lorentz group, i.e.,

$$L_a{}^b(1+\omega) = \delta_a{}^b + \frac{i}{2}\omega_{\mu\nu}(M^{\mu\nu})_a{}^b \tag{53}$$

where $(M^{\mu\nu})_a{}^b$ are representation matrices of the $\mathfrak{so}(3,1)$ Lie algebra, so that

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(\eta^{\mu\rho}\, M^{\nu\sigma} - \eta^{\nu\rho}\, M^{\mu\sigma}) - i(\eta^{\mu\sigma}\, M^{\nu\rho} - \eta^{\nu\sigma}\, M^{\mu\rho}). \tag{54}$$

To understand which spins elementary particles can have, we must find all possible (finite-dimensional) matrices $M^{\mu\nu}_{ab}$ that obey these commutation relations, in order to finite the possible fields that we can write down. This sounds like a challenging problem, but it turns out that we are lucky if we know something about the representation of the Lie algebra $SU(2)$ [6].
From QM, we know that $[J_i, J_j] = i\epsilon_{ijk}J_k$, which is the $SU(2)$ Lie algebra, is satisfied by sets of 3 hermitian matrices of size $(2j+1)\times(2j+1)$, where the eigenvalues of J_3 are $-j, -j+1, \ldots, +j$.

[4] All in units of $\hbar$, which we set to 1.

[5] In the current discussion, we are switching back and forth between considering particles and fields. In this, we are already using a result that we will see a little later in the course, namely that elementary particles show up as excitations of fields. Therefore, it is to some extent equivalent to talk about particles or about the associated fields, because the properties of the particles follow from the properties of the fields. However, let us highlight that there is a difference when it comes to representations of the Poincaré group: fields transform in the finite-dimensional representations of the group, i.e., they are constructed from a finite set of components. In contrast, particles transform in the infinite-dimensional representation of the Poincaré group. Physically, this is, loosely speaking because if you have a particle with some four-momentum p^μ, then there are infinitely many other four-momenta $p^{\mu'}$ that are related to p^μ by a boost. The choice of an infinite-dimensional representation is also necessary, because no finite-dimensional representation is unitary, and we would like to have probabilities (or scalar products of a state with itself) to be preserved under Poincaré transformations. Therefore, the representation that a field transforms in is not the same one as the particles that it gives rise to transform in. However, for our purposes at the present, we do not yet need to know this, as we will now simply focus on the representations that the fields can transform in.

[6] Note that the Lie algebras for $SO(3)$ and $SU(2)$ are identical. For the groups, there are some subtle differences, which need not directly concern us.

(If you would like a “refresher” on this, a good place to read up on it is, e.g., Sakurai “Modern Quantum Mechanics”.)

Our luck lies in the fact that upon introducing

$$N_i = \frac{1}{2}(J_i - iK_i) \qquad (\text{remember} : J_i = \frac{1}{2}\epsilon_{ijk}M_{jk} \text{ and } K_i = M_{i0}) \tag{55}$$

and

$$M_i = \frac{1}{2}(J_i + iK_i) \tag{56}$$

(Note: J_i, K_i are hermitian; N_i is not; in fact $M_i = N_i^\dagger$.) we find that

$$[N_i, N_j] = i\epsilon_{ijk}N_k, \qquad [M_i, M_j] = i\epsilon_{ijk}M_k, \qquad [N_i, M_j] = 0. \tag{57}$$

The Lie algebra of $SO(3,1)$ is nothing but two separate $SU(2)$ Lie algebras!
Thus, we can build the representations of the $SO(3,1)$ Lie algebra from representations of the $SU(2)$ Lie algebra!
$\Longrightarrow$ Each irreducible (i.e., not give by a product of two smaller representations) representation of the $SO(3,1)$ Lie algebra is specified by *two* integers or half-integers n' and n, which are the eigenvalues of M_3 and N_3.
We *label* these representations by n and n' or by the number of components in each representation, $(2n+1)$ and $(2n'+1)$.
To understand the corresponding spin of the field (and the particles that are the excitations of the field), we go back to the Pauli-Lubanski pseudovector and the associated Casimir operator, in the rest-frame, $W^2 = m^2\vec{J}\cdot\vec{J}$ and also use that $J_i = M_i + N_i$. Thus,

(n, n')	$\big(2n+1,\, 2n'+1\big)$	spin	name of the field
$(0,0)$	$(1,1)$	0	scalar (singlet)
$\left(\frac{1}{2}, 0\right)$	$(2,1)$	$\frac{1}{2}$	left-handed spinor
$\left(0, \frac{1}{2}\right)$	$(1,2)$	$\frac{1}{2}$	right-handed spinor
$\left(\frac{1}{2}, \frac{1}{2}\right)$	$(2,2)$	1	vector (this has $2\cdot 2 = 4$ components,
			which is the right number for a 4-vector)

Therefore, we now have a clear idea which fields we are going to consider. Rather than guessing that maybe there could be elementary particles with spin 2/3 (or other non half-integer values) out there, and somehow trying to come up with ideas for what the corresponding fields could be, we already know that such fields/particles do not exist and we do not need to spend our time trying to find a description for them, because our considerations, based on symmetries, tell us that such an effort is futile.
To sum up, by considering the fundamental symmetry of Minkowski spacetime, that a theory of fields and associated particles living on that spacetime has to satisfy, we have developed a comprehensive list of possible fields that can exist. Thus, rather than proceeding by trial-and-error, we have found a systematic structure that the rest of this course (and Quantum Field Theory) will follow. This structure is very restrictive and only allows us to consider fields which are associated to integer or half-integer spins. It is therefore not an accident that all elementary particles have integer or half-integer spin; there are no other options for them, based on the underlying symmetry

group, the Poincaré group.
We will work our way through the spin 0, 1/2 and 1 cases in the course, because, as it turns out, they are all part of the Standard Model of particle physics.
Higher spins (3/2, 2) do not correspond to detected elementary particles, although spin 3/2 plays a role in supergravity, where a spin 3/2 particle is the superpartner of the graviton. The graviton, which is the expected quantum of the gravitational field, has spin 2.

2.5 Noether's theorem

Symmetries not only help us to understand the building blocks of our theory (i.e., which fields there may be and how we characterize particles), they also imply conserved quantities and thus determine dynamical processes. The link between symmetries and conserved quantities is at the heart of Noether's theorem, just as in classical mechanics.

Noether's theorem in QFT states that: Every continuous symmetry of the action implies a conserved current density and a conserved charge.

This is similar to Noether's theorem in classical mechanics with the key difference being the conservation of the current. To derive the theorem, we will consider a scalar field; the theorem generalizes to non-zero spin fields, such as the gauge field and spinor fields.

As an example of a continuous spacetime symmetry, consider a translation $x \to x' = x + d$. How does $\phi \to \phi'$ look like? Note that we will take the *active point of view*, where we are assuming that the physical field configuration changes (in contrast to the *passive point of view*, where the coordinates change). It should hold that the transformed field at the transformed point is equal to the untransformed field at the original point, because, if we are shifting the field, but then also shift all points, the system remains unchanged. Thus

$$\phi'(x') = \phi(x), \tag{58}$$

which is shown in Fig. 1.
Thus, $\phi'(x)$ is defined by applying the inverse transformation to the argument, i.e.,

$$\phi'(x) = \phi(x - d). \tag{59}$$

When we generalize to a Lorentz transformation $x' = \Lambda x$, we have the same behavior: the scalar field is evaluated at a point that corresponds to the inverse of the transformation.

$$\phi'(x) = \phi(\Lambda^{-1}x). \tag{60}$$

We can also consider *internal* symmetries, e.g., for a complex scalar field $\phi(x)$ taking values in $\mathbb{C}$ instead of in $\mathbb{R}$, we can write an action that has a $U(1)$ symmetry:

$$S_{\mathrm{U(1)\,complex\,scalar}} = \int d^4x \left(\frac{1}{2}\partial_\mu\phi^*\partial^\mu\phi - \frac{1}{2}m^2\phi^*\phi\right), \tag{61}$$

which is invariant under $\phi(x) \to e^{i\alpha}\phi(x)$, and, accordingly $\phi^*(x) \to e^{-i\alpha}\phi^*(x)$. The infinitesimal

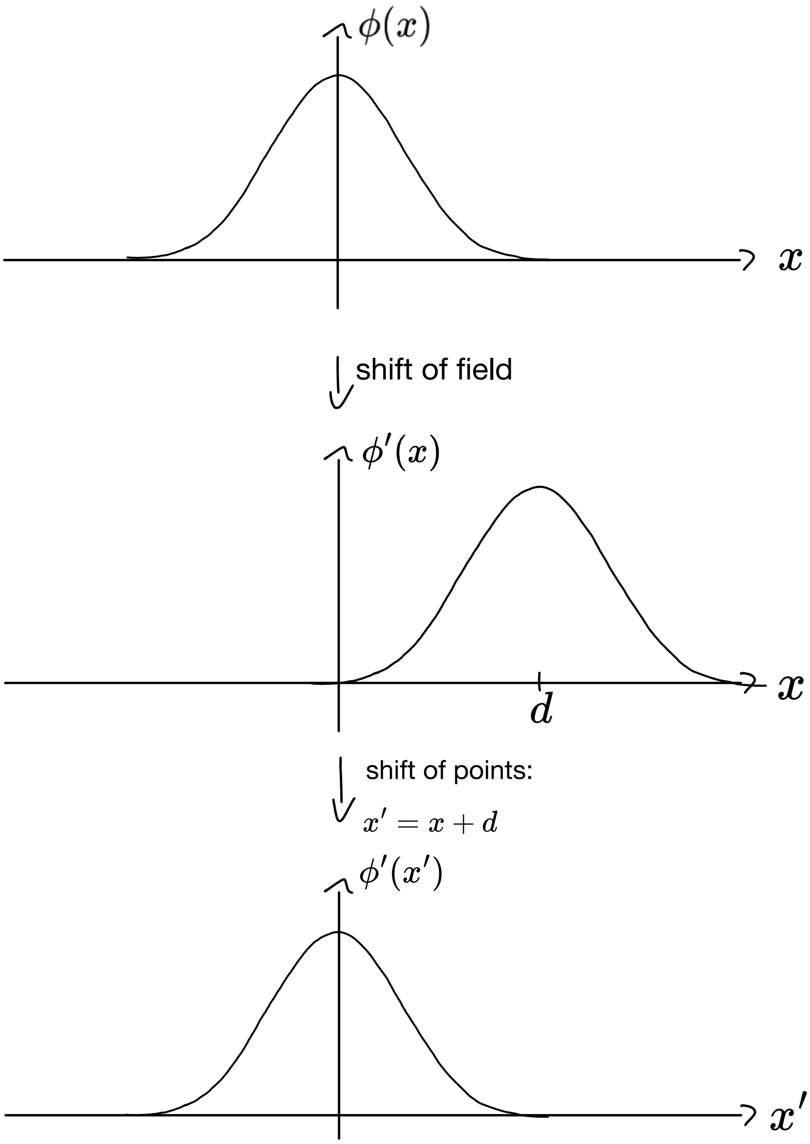


Figure 1: We show a field configuration $\phi(x)$ in the upper panel. In the central panel, we have shifted the field by a distance d (active transformation) and in the lower panel, we have then additionally shifted the coordinates by the same distance d, so that $x' = x + d$.

version of this transformation is

$$\phi'(x) = \phi(x)\,(1 + i\alpha + ...)\,. \tag{62}$$

We will consider this example in much more detail later in the lecture.
To derive Noether's theorem, we assume some continuous symmetry, but we do not need to specify whether it is a spacetime symmetry or an internal symmetry. Noether's theorem holds for both. Because we are assuming a continuous symmetry, there is an infinitesimal version of this transformation of the field

$$\phi(x) \to \phi'(x) = \phi(x) + \varepsilon\chi(x). \tag{63}$$

(For a discrete symmetry, there are only finite transformations, e.g. a $\mathbb{Z}_2$-symmetry under which

$\phi(x) \to -\phi(x)$ has no infinitesimal version. This is why all that follows holds for continuous, but not for discrete symmetries.)

For instance, for an infinitesimal translation, we can write the right-hand side in terms of a Taylor expansion

$$\phi'(x) = \phi(x) + \frac{\partial \phi}{\partial x^\mu} d^\mu + ..., \tag{64}$$
$$= \phi(x) + \epsilon_\mu \chi^\mu(x), \tag{65}$$

where we consider d^μ to be an infinitesimal shift and we defined $\partial_\mu \phi d^\mu = \epsilon_\mu \chi^\mu(x)$. When we perform a translation in a single direction in spacetime, this reduces back to the form $\epsilon_\mu \chi^\mu \to \epsilon\, \chi$. We denote the difference between the transformed and the untransformed field

$$\delta_\varepsilon \phi \coloneqq \phi' - \phi. \tag{66}$$

Under this change in the field, the Lagrangian changes as follows:

$$\delta_\varepsilon \mathcal{L} = \mathcal{L}' - \mathcal{L} = \mathcal{L}(\phi', \partial\phi') - \mathcal{L}(\phi, \partial\phi) \tag{67}$$
$$= \frac{\partial \mathcal{L}}{\partial \phi} \delta_\varepsilon \phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta_\varepsilon \partial_\mu \phi, \tag{68}$$

where $\delta_\varepsilon \partial_\mu \phi = \partial_\mu \phi' - \partial_\mu \phi$. (Note that we're slightly abusing naming conventions, as advertised, because this is the Lagrangian *density*, but we are referring to it as the Lagrangian. This is very common practise in QFT.)
Because we assume that the transformation corresponds to a symmetry of the action, the action must stay invariant under it. Thus, the Lagrangian may at most change by a total derivative, so we can write

$$\delta_\varepsilon \mathcal{L} = \varepsilon \partial_\mu F^\mu(\phi, \partial\phi, \partial^2\phi, x), \tag{69}$$

where, depending on the symmetry F^μ may actually be zero, so that even the Lagrangian is invariant under the symmetry.
We know that $\delta_\varepsilon \mathcal{L} \sim \varepsilon$, because $\delta_\varepsilon \mathcal{L} \to 0$ for $\varepsilon \to 0$. In principle, $F^\mu = F^\mu(\phi, \partial\phi, \partial^2\phi, x)$ can have dependencies on x and on $\partial^2\phi$, etc., even if $\mathcal{L}$ does not.
Now we want to derive the conserved current. $\partial_\mu F^\mu$ is a good starting point, because it already has the required form for a conservation law, $\partial_\mu j^\mu = 0$.

$$\varepsilon \partial_\mu F^\mu = \delta_\varepsilon \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \phi} \delta_\varepsilon \phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta_\varepsilon \partial_\mu \phi. \tag{70}$$

In the next step we use the equations of motion,

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = 0, \tag{71}$$

to rewrite the 1st term into a form that also has a partial derivative in front, as needed to derive a conservation law. Note that this will mean that everything that follows only applies for field configurations which satisfy the equations of motion. (In QFT, these are often called "on-shell" configurations. In a few weeks, when we talk about the path integral quantization, we will explicitly see the difference to the "off-shell" configurations.) We obtain

$$
\begin{aligned}
\Longrightarrow \ \varepsilon \partial_\mu F^\mu &= \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right) \delta_\varepsilon \phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta_\varepsilon \partial_\mu \phi && (72)\\
&= \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta_\varepsilon \phi\right). && (73)
\end{aligned}
$$

Thus,

$$\partial_\mu \Big(\underbrace{F^\mu - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \chi}_{j^\mu} \Big) = 0. \tag{74}$$

j^μ is a conserved current.

Example: Energy-momentum tensor and its conservation

In classical mechanics, the symmetry-transformation underlying energy-momentum conservation is a space-time-translation:

$$x^\mu \to x'^\mu = x^\mu - \epsilon^\mu. \tag{75}$$

(These are really 4 symmetries packaged into one.)
The resulting transformation of the field is, as we wrote above,

$$
\begin{aligned}
\phi'(x) &= \phi(x + \varepsilon) && (76)\\
\Longrightarrow \ \delta_\varepsilon \phi &= \phi(x+\varepsilon) - \phi(x) = \varepsilon^\nu \underbrace{\partial_\nu \phi(x)}_{\chi_\nu} && (77)
\end{aligned}
$$

Eq. (76) means that the new field at x is the same as the old one at $x+\varepsilon$, because the shift is by $-\varepsilon$, and we are again using the active view on transformations. In Eq. (77), χ has an index, because there are *4* symmetries. Eq. (77) is to 1st order in ε, because we can Taylor expand $\phi(x+\varepsilon)$. Thus, if we focus on the dependence of $\mathcal{L}$ on x (through its dependence on ϕ),

$$
\begin{aligned}
\mathcal{L}'(x) &= \mathcal{L}(x+\varepsilon) && (78)\\
\Longrightarrow \ \delta_\varepsilon \mathcal{L} &= \mathcal{L}(x+\varepsilon) - \mathcal{L}(x) && (79)\\
&= \varepsilon^\mu \partial_\mu \mathcal{L}(x) \qquad \text{(to 1st order in } \varepsilon) && (80)\\
&= \varepsilon^\nu \partial_\mu (\underbrace{\delta^\mu{}_\nu \mathcal{L}}_{\coloneqq F^\mu{}_\nu\,.}) && (81)
\end{aligned}
$$

Now we can use the general expression we derived before to get the conserved currents. Because we are looking at 4 symmetries at the same time, we will have 4 conserved currents, each of which is a 4-vector. In Eq. (81), you can think of the index μ as the index that belongs to a conserved current (which is a four-vector) and the index ν as the one that labels the four distinct currents that there are for the four distinct translations. Which translation is performed, is selected by the non-zero components of ϵ^ν.

To “package” the four conserved currents into one expression, we write

$$j^\mu{}_\nu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\chi_\nu - F^\mu{}_\nu \tag{82}$$

$$= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi - \delta^\mu{}_\nu\mathcal{L}. \tag{83}$$

This conserved *tensor* is usually written as

$$T^{\mu\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L}. \tag{84}$$

It is conserved, $\partial_\mu T^{\mu\nu} = 0$, and called the energy-momentum (or stress-energy) tensor.
Side-note: This is an entry-point into General Relativity, because, if we promote $\eta^{\mu\nu} \to g^{\mu\nu}$, then $T^{\mu\nu}$ acts as a source for spacetime curvature in the Einstein equations. The physical meaning behind that is that any form of energy or momentum sources spacetime curvature.

From the conservation of the current, we can also derive the conservation of a charge:

$$Q(t) = \int d^3x\, j^0(t,\vec{x}). \tag{85}$$

It holds that

$$\dot{Q} = \frac{\mathrm{d}}{\mathrm{d}t}Q(t) = 0, \tag{86}$$

if we assume that all fields and their derivatives vanish at $|x| \to \infty$, i.e., we only consider nonzero field configurations away from spatial infinity. This is reasonable to describe all realistic physical situations that we are interested in (e.g., particle physics experiments at CERN, phonons in the Bose-Einstein-Condensates of our experimental friends in Neuenheimer Feld, or superconducting Cooper-pairs in superconductors in various labs, all of which are described by QFT.)
We can show $\dot{Q} = 0$ as follows:

$$\dot{Q} = \frac{\mathrm{d}}{\mathrm{d}t}\int d^3x\, j^0(t,\vec{x}) = \int d^3x \big(\partial_0 j^0(t,\vec{x})\big) \tag{87}$$

$$= \int d^3x\, \partial_i j^i(t,\vec{x}) \qquad \text{(by conservation of the current)} \tag{88}$$

$$= \int dxdydz(\partial_x j^x + \partial_y j^y + \partial_z j^z) \tag{89}$$

$$= \int dydz\, j^x\Big|_{x\to\pm\infty} + \int dxdz\, j^y\Big|_{y\to\pm\infty} + \int dxdy\, j^z\Big|_{z\to\pm\infty} \tag{90}$$

$$= 0, \tag{91}$$

if fields and derivatives vanish at $|x| \to \infty$, so that j vanishes there.
Let us highlight that the conservation of a *current* is stronger than the conservation of the charge, because it implies that the charge is conserved *locally*, i.e., changes of the charge in a (finite) volume in time must be accounted for by a current flowing though the surface of the volume. To see this, write:

$$\frac{\mathrm{d}Q_V}{\mathrm{d}t} = -\int_V d^3x\, \vec{\nabla}\cdot\vec{j} = -\int_{A=\partial V}\vec{j}\cdot d\vec{S}. \tag{92}$$

Q_V is the charge in a volume V. In the last step we used Gauss’ law.

Example: the conserved charges following from the conservation of the energy-momentum tensor are:

$$\int d^3x\, T^{00} = \underbrace{\int d^3x \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \dot{\phi} - \mathcal{L} \right)}_{\substack{\text{We recognize this} \\ \text{as the Hamiltonian!}}} = H =: P^0 \tag{93}$$

The other conserved charges are the spatial momenta, so $P^\nu = \int d^3x\, T^{0\nu}$ is conserved.

Noether's theorem also applies to continuous internal symmetries. We'll consider an example later in the course.

Mini-Exercise 2.4. Take

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2. \tag{94}$$

What is P^i?

We note that the conserved quantities in turn are the *generators* of the associated symmetry. This closes our considerations of symmetries. We have learned that symmetries are encoded in groups. Continuous symmetries of interest in physics are Lie groups, for which each symmetry transformation can be generated by the generators of the Lie algebra. In turn, Noether's theorem tells us that each symmetry leads to a conserved quantity. This conserved quantity is the generator that generates this symmetry.
If we did not know about the Lie group associated to a symmetry, we could therefore learn about it from the action of the symmetry and the resulting conserved quantities.

A Solutions to Mini-exercises

Chapter 1

Solution A.1.1.

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - C(\varphi-\gamma) - \frac{1}{2}m^2(\varphi-\gamma)^2 \\ &= \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \underline{C\varphi} + C\gamma - \frac{1}{2}m^2\varphi^2 + \underline{m^2\varphi\gamma} - \frac{1}{2}m^2\gamma^2\end{aligned}$$

Define $\gamma = \frac{C}{m^2}$:

$$\to \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi + \frac{C^2}{m^2} - \frac{1}{2}m^2\varphi^2 - \frac{1}{2}m^2\frac{C^2}{m^2}.$$

Chapter 2

Solution A.2.1.

$$\mathrm{R}\,\mathrm{R}^\top = \exp(T)\exp(-T) = \mathbb{1},$$

because

$$\mathrm{R}^\top = (\exp(T))^\top = \exp T^\top = \exp(-T).$$

Solution A.2.2.

$$\begin{aligned}&\big(\delta^\mu{}_\rho + \omega^\mu{}_\rho\big)\eta_{\mu\nu}(\delta^\nu{}_\sigma + \omega^\nu{}_\sigma) \overset{!}{=} \eta_{\sigma\rho} \\ &\eta_{\rho\sigma} + \omega_{\sigma\rho} + \omega_{\rho\sigma} + \mathcal{O}(\omega^2) = \eta_{\sigma\rho} \implies \omega_{\rho\sigma} = -\omega_{\sigma\rho}.\end{aligned}$$

Solution A.2.3.

$$\begin{aligned}[P^2, M_{\mu\nu}] &= [P^\rho P_\rho, M_{\mu\nu}] \\ &= P^\rho[P_\rho, M_{\mu\nu}] + [P^\rho, M_{\mu\nu}]P_\rho \\ &= P^\rho(i\eta_{\rho\nu}P_\nu - i\eta_{\rho\mu}P_\nu) + (i\eta_{\rho\nu}P_\mu - i\eta_{\rho\mu}P_\nu)P^\rho \\ &= i(P_\nu P_\mu - P_\mu P_\nu + P_\mu P_\nu - P_\nu P_\mu) \\ &= 0.\end{aligned}$$