Quantum Field Theory I

Lecture in SS 2023 at the KFU Graz

Axel Maas

Contents

Chapter 1

Introduction

Quantum-field theory is the (extremely successful) combination of quantum physics and special relativity. As a field theory, it replaces the concept of particles with that of extended fields, like continuum mechanics replaces point particles with matter density fields. This turns out to be a necessary step. However, such a step is not (mainly) motivated by quantum physics, but indeed already by classical electrodynamics.

When considering classical electrodynamics, it becomes apparent that the concept of (charged) point particles is problematic. The first glimpse of that is already visible when attempting to calculate the electric energy density of a point particle, which diverges. It becomes very apparent when attempting to understand the back reaction of the electromagnetic field created by an accelerating, massive particle on the particle. While some effective description have been derived over time, all of them need to deal with complicated singularities. On the other hand, when considering, e. g., magnetohydrodynamics, which just has matter fields in terms of fluids, these problems become very much reduced. This already suggests that a description of matter in terms of fields is likely more appropriate. As will be seen, the particle concept will then arise again in a similar way as classical mechanics emerges from quantum mechanics by means of the Ehrenfest relations, by a suitable averaging. However, it will require first a good understanding of the underlying theory before it is possible to return to this question.

Thus, the main aim here is to formulate a quantum special relativistic field theory. Extending this to also include general relativity leads to the, not yet fully obtained, quantum gravity, which is subject of another lecture. In contrast, the following is, while not yet mathematical watertight, at least much better under control. For this purposes the path integral formulation, which will be introduced in chapter 2, is best suited, especially with hindsight to gauge theories in the lecture "Quantum Field Theory II". How this relates to the more familiar operator, or canonical, formalism, will therefore be relegated to chapter

10. The non-relativistic limit, and thus how quantum mechanics emerges from quantum field theory, will also be there discussed in chapter 10.3.

Just as with quantum mechanics, perturbation theory is both a mainstay of applying quantum field theory as well as very important to define terminology. It will therefore be treated twice, once for lowest order calculations in chapter 3 and once for higher orders in chapter 8. In particular, this will give a first idea of why the usage of field and particle is often interchangeable, even if not fully accurate.

Interestingly enough, the spin of particles (or, more precisely, fields) plays a much more important role than in quantum mechanics, and turns out to be intimately tied to special relativity. For scalar particles this is actually relatively trivial, and can be mostly ignored. They will therefore serve as the first objects already in chapter 2 and 3. Particles of different spin will then be introduced in chapter 4.

Finally, there are many interesting subjects, which cannot be treated perturbatively. At the forefront of them are bound states and quantum phase diagrams. While the methods necessary to deal with them in any non-trivial way substantial exceeds this lecture, and are covered in lectures of their own, a first glimpse will be provided in chapter 9.

As with any topic of such relevance there are numerous textbooks and overview articles. While anyone should make their own choice, the following books were very helpful in the preparation of this lecture, in order of increasing complexity:

- Peskin & Schröder, “Quantum Field theory” (Cambridge)
- Das, “Lectures on quantum field theory” (World Scientific)
- Dyson, “Advanced quantum physics” (World Scientific)
- Böhm et al. “Gauge theories of the strong and electroweak interaction” (Teubner)
- Weinberg, “Quantum Theory of Fields” (Cambridge)
- DeWitt, “The global approach to quantum field theory” (Oxford)
- Haag, “Local quantum physics” (Springer)

Interested students may find a first access to the philosophical implications of quantum field theory (and physics in general) in

- Rickles, “The philosophy of physics” (Cambridge)
- Friebe at al., “The philosophy of quantum physics” (Springer)
- Butterfield & Earman (editors), “Philosophy of Physics” (Elsevier)

- H. Lyre, "Lokale Symmetrien und Wirklichkeit" (in German) (Mentis)
- Berghofer et al., "Gauge Symmetries, Symmetry Breaking, and Gauge-Invariant Approaches" (Cambridge)

There are, however, very few textbook focused on non-gauge quantum field theories, due to overwhelming relevance of gauge theories.

Chapter 2

The path integral and scalar particles

2.1 Classical field theories

As has been emphasized in the introduction, already classical physics suggests that fundamental physics needs to be done in terms of fields. In principle, already quantum mechanics is a field theory, as the wave-function itself is already a field. However, position was different than time, and especially an operator, except when going into a position-space representation. On the other hand, special relativity (and also general relativity) strongly suggests to keep space and time on equal footing.

How these two aspects can be unified is actually not unique, even though attempting to treat time by an operator turns out to be much more involved. Eventually, it is experiment, which decides, as it will decide how suitable theories look like, and how they are quantized. In the corresponding theoretical developments, this implies that certain postulates need to be made. These will be clearly marked in the following. However, while there are possibilities to motivate them, and eventually derive them from a full quantum gravity theory of matter, eventually they cannot be reasoned for[1].

Here, it is therefore time to start with postulating the classical structure of the field theories to be discussed, and later quantized. While later on more complicated entities will arise, at the moment only the simplest case, spinless, uncharged particles, will be considered. Extending the postulate to cover other cases will happen throughout the lecture, and especially in section 2.5 and chapter 4, and later in the lecture "Quantum Field Theory II".

The basic entity is therefore a real-valued field, $\phi(x)$, which is parametrized by a four-vector x. Thus, the field is a map from a four-vector in a Minkowski space-time to the

[1]Of course, also in quantum gravity new postulates appear. Whether a postulate-free theory is even possible is an unresolved question.

real numbers. At this moment no probability interpretation will be attached to it, and in fact it is not the same as a wave-function. Furthermore, under a Poincaré transformation of its argument, the field remains unchanged, i. e.

$$\begin{aligned} x_\mu &\rightarrow x'_\mu = \Lambda_{\mu\nu} x^\nu + a_\mu \\ \phi(x) &\rightarrow \phi(\Lambda x + a_\mu), \end{aligned}$$

where Λ is a Lorentz transformation and a_μ a translation vector. Of course, in a non-relativistic setting the four vectors would be replaced by time and a three-vector separately, and Poincaré symmetry by Galileo symmetry. That changes many of the things in the following only marginally, and can be understood from the non-relativistic limit to be discussed in section 10.3. Thus, here the relativistic notion will be used. Conversely, some aspects like spin cannot be explained in a non-relativistic setting, but only postulated. But they emerge naturally in a relativistic framework.

To describe classically the dynamics of such a field requires either a Hamiltonian or a Lagrangian density. Since in a path-integral, relativistic treatment the Lagrangian turns out to be the more useful one, it will be concentrated on for now. The role of the Hamiltonian will be discussed again in chapter 10. The classical theory of a free field is then given by the Lagrangian density

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi, \tag{2.1}$$

which describes a freely vibrating relativistic field. These are in turn determined by the corresponding equations of motions, the Lagrange equations. It should be noted that as initial conditions it is sufficient to know the field and its derivatives at all points in space at a fixed time.

The Lagrangian itself is of second order in time and space, and manifestly invariant under Poincaré symmetry. It is thus different from quantum mechanics, where the time derivative is only of first order. Again, this is an artifact of the non-relativistic limit, see section 10.3.

It is natural to ask, whether other derivatives can appear. Less derivatives will lead to interactions, as described in section 2.2. More derivatives will imply also more time derivative, if special relativity should be maintained. As a consequence, more initial data would be needed. Especially, it is possible to have a different evolution of a field and its derivatives fully fixed at a given time by changing the conditions at some later time. This appears to be acausal, and is certainly in contradiction to existing experiments. While it is an interesting question if such theories can be given a meaning at the quantum level, this question is so open that here only theories up to second order in the derivatives will be considered.

2.2 The linear sigma model

While (2.1) is a complete theory, it is only describing a free field, and thus at most plane waves. To have other phenomena, it requires interactions, either with another field, or with itself. Relegating additional fields to later, this implies that the interacting field theory should be given by

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi, \partial_\mu\phi),$$

with the understanding that only up to order two derivatives should be admitted. Ultimately experiment will fix the potential. But some quite general rules have been found for the potential from the actually observed physics. These are that the potential, also in more general theories, is a polynomial of the fields up to mass dimension d, where d is the number of space-time dimension. This requires to define the mass dimension.

In the following, the action

$$S = \int d^dx\mathcal{L}$$

will become an important object. The action is in natural units, i. e. $k_B = \hbar = c = 1$, which will be used throughout, dimensionless. Because the integration measure has dimension[2] Length^d =Energy^{-d}, and derivatives have unit Length^{-1}, this implies that the scalar field has mass dimension Length$^{-\frac{d-2}{2}}$=Energy$^{\frac{d-2}{2}}$. Hence, a polynomial of order n has mass dimension $(n(d-2))/2$, and thus 4 in 4 dimensions, to which this will be mostly restricted. While both features are an empirical result, they have profound consequences to be addressed in sections 3.1 and 8.2.

The most general Lagrangian consistent with these rules is thus

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2 - \frac{\zeta}{6}\phi^3 - \frac{\lambda}{24}\phi^4, \tag{2.2}$$

where the numeric prefactors have been chosen for later convenience. This theory is known as the linear σ-model. Note that while m^2 is written like a positive quantity, it will be allowed to be negative as well. However, because the Lagrangian and the action are real, all three constants need to be real as well. As will be seen, the quantity m will be connected to the mass of the particle, while ζ and λ will be related to the interactions. The latter are therefore called coupling constants. It should be noted that classically $\lambda > 0$ is required for the theory to have a stable least-energy configuration, and thus describing a stable system. If $\lambda = 0$, this requires that $m^2 \geq 0$ and $\zeta = 0$ to have a stable system.

The linear σ-model is a well-defined theory, but also already classically no longer exactly solvable. It is, however, the simplest interacting theory, which plays a larger role as a

[2]The units employed are usually eV and fm, which can be converted using $\hbar c = 197$ MeV·fm. But only one of them is needed in any given expression in the chosen natural units.

quantum field theory. It will therefore serve for large parts of this lecture as the standard example, especially in the case that $\zeta = 0$.

2.3 Path integral quantization

In this section, the basic formulation of the path integral will be constructed, together with the necessary tools to treat real-valued, scalar fields. The application of perturbation theory and the derivation of Feynman rules will be discussed in chapter 3.

2.3.1 Heuristic introduction

The path integral formulation is as axiomatic as is canonical quantization, it cannot be derived. However, it is possible to motivate it.

This is done most easily by defining it as a limiting procedure, using a so-called lattice regularization. As the name implies, it works by replacing space-time with a finite lattice of discrete points. This technique is useful also for doing calculations in quantum field theories, and is explored in more detail in a dedicated lecture. For the moment, it is, however, just a tool to define a limiting process. The approach is therefore quite similar to what is done by defining the path integral in quantum mechanics by creating it from a product of propagators. It is actually possible to define, even mathematically rigorous in the non-interacting case, the path integral directly in the continuum. However, this requires more general ways of integrating, so-called Ito integration, as quantum fields are usually non-continuously differentiable functions, which cannot be treated by Riemann integration.

In fact, it is best to start with the quantum mechanics version, but in a different than usual way. The heuristic reasoning is then as follows. Take a quantum-mechanical particle which moves in time T from a point a of origin to a point b of measurement. This is not yet making any statement about the path the particle followed. In fact, in quantum mechanics, due to the superposition principle, a-priori no path is preferred. Therefore, the transition amplitude U for this process must be expressible as

$$U(a, b, T) = \sum_{\text{All paths}} e^{i \cdot \text{Phase}} \tag{2.3}$$

which are weighted by a generic phase associated with the path. Since all paths are equal from the quantum mechanical point of view, this phase must be real. Thus it remains only to determine this phase. Based on the correspondence principle, in the classical limit the classical path must be most important. Thus, to reduce interference effects, the phase

should be minimal for the classical path. A function which implements this is the classical action S, determined as

$$S = \int_{\mathcal{C}_a^b} dt L,$$

where the integral is over the classical path $\mathcal{C}_a^b$ from a to b. If the classical path is now replaced with some arbitrary path C_a^b, again connecting the points a and b, the action can be considered to be a functional of the path C_a^b and the classical Lagrange function L. Of course, it is always possible to add a constant to the action without altering the result. Rewriting the sum as a functional integral over all paths, this yields already the definition of the functional integral

$$U(a,b,T) = \sum_{C_a^b} e^{iS\left[C_a^b\right]} \equiv \int \mathcal{D}C_a^b e^{iS\left[C_a^b\right]}.$$

This defines the quantum mechanical path integral in a symbolic way.

It then remains to give this functional integral a mathematically concise meaning, such that it becomes a mathematical description of how to determine this transition amplitude. It is here where the lattice comes into play. However, for arbitrary interacting theories, there are still conceptual and practical problems, so that the following remains often an unproven procedure.

The starting point was the transition amplitude. In quantum mechanics, this amplitude is given by

$$U(a,b,T=t_N-t_0) = \left\langle a,t_N \left| e^{-iHT} \right| b,t_0 \right\rangle.$$

In the next step, insert at intermediate times a sum, or integral in cases of a continuous spectrum, over all states

$$U(a,b,T) = \sum_i \left\langle a,t_N \left| e^{-iH(t_N-t_1)} \right| i,t_1 \right\rangle \left\langle i,t_1 \left| e^{-iH(t_1-t_0)} \right| b,t_0 \right\rangle.$$

By this, the transition amplitude is expressed by a sum over all possible intermediate states, already quite in the spirit of (2.3). To fully embrace the idea, divide the time interval into N steps of size $\epsilon = T/N$, where N is large and will later be send to infinity. That is actually already a lattice in time. This yields

$$\begin{aligned} U(a,b,T) &= \sum_j \sum_{i_j} \left\langle a,t_N \left| e^{-iH\epsilon} \right| i_{N-1},t_{N-1} \right\rangle ... \left\langle i_1,t_1 \left| e^{-iH\epsilon} \right| b,t_0 \right\rangle \\ &= \int \Pi_i dq_i \left\langle q_a,t_N \left| e^{-iH\epsilon} \right| q_{N-1},t_{N-1} \right\rangle ... \left\langle q_1,t_1 \left| e^{-iH\epsilon} \right| q_b,t_0 \right\rangle, \end{aligned} \tag{2.4}$$

where in the second line the result was rewritten in terms of a set of continuous eigenstates of the (generalized) position operator Q_i. These are therefore $N-1$ integrals. In this way,

the time is given a lattice structure. Space is still a continuum, which extends over all space.

If, as is the case for all systems of interest in the following, the Hamiltonian separates as

$$H = \frac{1}{2}P_i^2 + V(Q),$$

where the P_i and Q_i are the M canonically conjugated momenta, then for ϵ arbitrarily small the Baker-Campbell-Hausdorff formula

$$\exp F \exp G = \exp\left(F + G + \frac{1}{2}[F,G] + \frac{1}{12}\left([[F,G],G] + [F,[F,G]]\right) + ...\right).$$

yields

$$e^{-iH\epsilon} \approx e^{-\frac{i\epsilon}{2}P_i^2} e^{-i\epsilon V},$$

i. e. for infinitesimally small time steps the exponentials can be separated. Assuming the states to be eigenstates of the position operator and furthermore inserting a complete set of (also continuous) momentum eigenstates allows to rewrite the transition matrix elements as ordinary functions

$$\left\langle q_{i+1}, t_{i+1} \left| e^{-iH\epsilon} \right| q_i, t_i \right\rangle = e^{-\epsilon V(q_i)} \int \Pi_j \frac{dp_j^i}{2\pi} \Pi_k e^{-i\epsilon\left(\frac{p_k^{i2}}{2} - ip_k \frac{q_k^{i+1} - q_k^i}{\epsilon}\right)}, \tag{2.5}$$

where products run over the number of independent coordinates M. The infinitesimal step (2.5) is also known as the transfer matrix, which transfers the system from one time to another. In fact, even if the Hamilton operator is not known, but only the transfer matrix, it is possible to construct the full transition amplitude, as this only requires to create a product over all transfer matrices.

Defining

$$\mathcal{D}p\mathcal{D}q = \Pi_i^N \Pi_j^M \frac{dp_j^i dq_j^i}{2\pi}, \tag{2.6}$$

and thus in total $2NM$ integration measures yields the first formulation of the path integral

$$U(a,b,T) = \int \mathcal{D}p\mathcal{D}q e^{-\epsilon p_k \frac{q_j^{i+1} - q_j^i}{\epsilon}} e^{-i\epsilon H(p_j^i, q^j)}$$

Defining

$$\frac{q_j^{i+1} - q_j^i}{\epsilon} = d_t q_j^i$$

and performing the Gaussian integrals over the momenta yields

$$\begin{aligned} U(a,b,T) &= \int \mathcal{D}q e^{i\sum^N \epsilon L(q_j^i, d_t q_j^i)} \overset{N\to\infty}{=} \int \mathcal{D}q e^{iS} \\ \mathcal{D}q &= \Pi_i^N \Pi_j^M \frac{dq_j^i}{\sqrt{2\pi\epsilon}}, \end{aligned}$$

where L is the Lagrange function of the system, thus arriving at the original idea (2.3).

Considering the result in detail, it is important to note one important feature. The definition requires to chose any straight line between every point at every time. Thus, in general paths will contribute which are not differentiable. This is a very important insight: Quantum physics differs from classical physics not only by including all possible paths, but also by including not only differentiable paths. This is in stark contrast to Hamilton's principle of classical mechanics.

Passing now to a field theory, the transition is the same as in classical mechanics: The paths are replaced by the fields, the Lagrange function by the Lagrangian density, and the action is an integral over space-time. Of particular importance is then the partition function

$$Z = \int \mathcal{D}\phi e^{i\int d^dx\mathcal{L}(\phi,\partial_\mu\phi)}, \tag{2.7}$$

where the integral is over all possible field configurations, i. e. the set of all possible values of the fields at every space-time point, including any non-differentiable ones[3]. Since any field configuration includes the time-dependence, the path-integral can be considered as an integral over all possible field configurations, and thus histories of the universe described by the Lagrangian $\mathcal{L}$, from the infinite past to the infinite future. Thus, the path integral makes the absence of locality in quantum physics quite manifest. The partition function (2.7) is essentially the transition function from the vacuum to the vacuum. It is important to note that in the whole setup the field variables are no longer operators, like in canonical quantization, but ordinary functions.

It is also important to note that the integration is over all possible field configurations. Thus, in contrast to the beginning and end of the paths in the quantum-mechanical case, there are no boundary conditions imposed. Of course, if desired, additional boundary conditions can be imposed, as can be any kind of underlying space-time manifold. But for any set of boundary conditions and/or any change of manifold a different theory will result. Neither of this will be done in this lecture.

While the vacuum-to-vacuum transition amplitude is a very useful quantity, what is really important are the expectation values of the correlation functions, i. e. expectation values of products, or other functions, of more than one field. These can be determined in a very similar way as before to be

$$\langle T(\phi(x_1)...\phi(x_n))\rangle = \int \mathcal{D}\phi\phi(x_1)...\phi(x_n)e^{i\int d^dx\mathcal{L}(\phi,\partial_\mu\phi)}. \tag{2.8}$$

[3]In fact, it can be shown that those are the dominating one. Making sense out of this expression in the continuum is highly non-trivial and requires to pass from Riemann integrals to different definitions of integrals, but this is not the subject of this lecture.

Here, T means time-ordering, i. e. that the time arguments x_i^0 are ordered in descendant order, i. e.

$$T(\phi(t_1)\phi(t_2)) = \theta(t_1 - t_2)\phi(t_1)\phi(t_2) \pm \theta(t_2 - t_1)\phi(t_2)\phi(t_1), \tag{2.9}$$

where the minus sign appears if the ϕ anticommute. Note that this time-ordering does not appear inside the path integral, but is created automatically. Inside the path integral, the ordering is arbitrary. In a path-integral context, this is less important, but will be found to be relevant during canonical quantization in chapter 10, when the fields are interpreted as operators, rather than as functions. It also is important if fields anticommute, what will happen when introducing fermions in chapter 4.

It is here implicitly assumed that $Z = 1$, i. e. that the measure of the path integral is normalized such that the expectation value of unity is one, $\langle 1 \rangle = 1$. Otherwise

$$\langle T\phi(x_1)...\phi(x_n)\rangle = \frac{\int \mathcal{D}\phi\phi(x_1)...\phi(x_n)e^{iS[\phi]}}{\int \mathcal{D}\phi e^{iS[\phi]}}, \tag{2.10}$$

holds.

There are two important remarks. This can be seen by noting that the fields are evaluated at fixed times, and therefore evaluate to functions of the positions in their respective transfer matrix elements (2.5). Thus, any expectation value is a path integral over the fields as ordinary functions weighted by the phase.

2.3.2 Functional analysis

So far, any explicit calculation with the path integral would require to go back to the limiting prescription (2.4). This would be quite awkward. Fortunately, this is usually not necessary, and it is often possible to work with the path integral in much the same way as with ordinary integrals.

How to do so falls into the mathematical purview of functional analysis and distribution theory to provide any level of rigor. For most applications in physics, little of this rigor is necessary. Thus, here results will be collected, which are useful to work in an operational way with the path integral, but they will not be proven. Rather, the corresponding mathematical literature and/or lectures can provide this, where desired and necessary. Nonetheless, it is advised that, with all simplicity the following seems to bring, one should always be wary that many issues run much deeper.

The starting point before defining functional integration is the definition of a functional derivative. Essentially, this is an extension of variational calculus in classical mechanics.

The basic ingredient for a functional derivative δ with respect to a function (field) $\phi(x)$ are the definitions

$$\begin{aligned}\frac{\delta 1}{\delta\phi(x)} &= 0\\ \frac{\delta\phi(y)}{\delta\phi(x)} &= \delta(x-y)\\ \frac{\delta}{\delta\phi(x)}(\alpha(y)\beta(z)) &= \frac{\delta\alpha(y)}{\delta\phi(x)}\beta(z)+\alpha(x)\frac{\delta\beta(z)}{\delta\phi(x)},\end{aligned}$$

in analogy to conventional derivatives. The last identity is known as the Leibnitz rule.

Consequently, a power series of a functional is defined as

$$F[\phi]=\sum_{n=0}^{\infty}\int dx_1...dx_n\frac{1}{n!}T(x_1,...,x_n)\phi(x_1)...\phi(x_n),$$

where the coefficients of an ordinary power series are now replaced by coefficient functions T. In particular, they can be obtained as

$$T(x_1,...,x_n)=\left.\frac{\delta^n}{\delta\phi(x_1)...\delta\phi(x_n)}F[\phi]\right|_{\phi=0}.$$

This defines the most important concepts for differentiation. If not stated otherwise, it will be assumed in the following that any functional can be written as functional Taylor series. Just as with ordinary functions, this is not guaranteed.

Concerning the functional integrals, they are as usually defined to be the inverse operation to functional derivatives. Therefore, integration proceeds as usual. In most practical cases, the relevant functional are either polynomials or can be expanded in a power series, and then functional integrals are straight-forward generalizations of the usual integrals. In particular

$$\begin{aligned}\int\mathcal{D}\phi &= \phi(x)\\ \int\mathcal{D}\phi\phi &= \frac{1}{2}\phi(x)^2,\end{aligned}$$

where the first expression implies that $\delta\int$ equals not to one, but equals a δ-function.

Of particular importance are Gaussian integrals, i. e. the generalization of

$$\int_{-\infty}^{\infty}\frac{dx}{\sqrt{\pi}}e^{-ax^2}=\frac{1}{\sqrt{a}}. \tag{2.11}$$

The result can be either obtained from the power series expansion or directly gleaned from the finite-dimensional generalization of Gaussian integrals, which is given by

$$\int_{-\infty}^{\infty}\frac{dx_1}{\sqrt{\pi}}...\int_{-\infty}^{\infty}\frac{dx_n}{\sqrt{\pi}}e^{-x^TAx}=\frac{1}{\sqrt{\det A}},$$

with an arbitrary matrix A, though for a finite result the square-root of its determinant must be invertible, i. e., no zero eigenvalues may be present.

The functional generalization is then

$$\int \mathcal{D}\phi e^{-\int dxdy\phi(x)A(x,y)\phi(y)} = \frac{1}{\sqrt{\det A(x,y)}},$$

where A may now be operator-valued. Especially derivative operators may appear in this context. The determinant of such an operator can be evaluated by the expression

$$\det A = \exp \operatorname{tr} \ln(A), \tag{2.12}$$

just like for matrices, which is of great practical relevance. Herein, both the logarithm and the exponential of an operator are defined by the usual power series of these operations[4]. Alternatively, $\det A$ can be expressed in terms of the solutions of the eigenvalue equation

$$\int dy A(x,y)\phi(y) = \lambda\phi(x),$$

where the eigenvalues λ can form a (complex) continuum, and finite, or even infinite, degeneracies are possible. The determinant is then given as the product of all eigenvalues.

An important property is the definition that a functional integral is translationally invariant. Thus, for an arbitrary functional F and an arbitrary function η and constant α

$$\int \mathcal{D}\phi F\left[\phi + \alpha\eta\right] \overset{\phi\to\phi-\alpha\eta}{=} \int \mathcal{D}\phi F\left[\phi\right] \tag{2.13}$$

holds by definition.

From these properties follows the validity of the substitution rule as

$$\int \mathcal{D}\phi F\left[\phi\right] = \int \mathcal{D}\psi \det\left(\frac{\delta\phi}{\delta\psi}\right) F\left[\phi\left[\psi\right]\right], \tag{2.14}$$

where the Jacobi determinant $\det(\delta\phi/\delta\psi)$ appears. In case of a linear transformation

$$\phi(x) = \int dy \eta(x,y)\psi(y), \tag{2.15}$$

the determinant is just $\det \eta(x,y)$ of the infinite-dimensional matrix $\eta(x,y)$ with the indices x and y.

[4]This implies that problems may arise, as this is not always justified. In fact, there exist operators for which even the trace is not well defined. Even though this rarely plays a role in the following, caution is in general advised.

2.3.3 Sources

Combining all of the above allows for a different way of calculating correlations functions than by direct evaluation of (2.10). For this, introduce[5] a so-called (external) source $j(x)$, and replace

$$iS[\phi] \to iS[\phi] + \int d^d x \phi(x) j(x).$$

Thus, the partition function will become a functional of $j(x)$. Then a more elegant way to express correlation functions is by

$$\begin{aligned}\langle T\phi(x_1)...\phi(x_n)\rangle &= \frac{1}{Z[0]} \int \mathcal{D}\phi e^{iS[\phi]+\int d^d x\phi(x)j(x)}\ \phi(x_1)...\phi(x_n)\Big|_{j=0} \\ &= \frac{1}{Z[0]} \int \mathcal{D}\phi \frac{\delta^n}{\delta j(x_1)...\delta j(x_n)} e^{iS[\phi]+\int d^d x\phi(x)j(x)}\Big|_{j=0} = \frac{1}{Z[0]}\ \frac{\delta^n}{\delta j(x_1)...\delta j(x_n)} Z[j]\Big|_{j=0}.\end{aligned}$$

Note that it was assumed that functional derivation and integration commute, and that often the normalization $Z[0] = 1$ is chosen.

Furthermore, this permits to reconstruct the original path-integral, or generating functional, as

$$Z[j] = \sum_{n=0}^{\infty} \int d^d x_1 ... d^d x_n \langle T\phi(x_1)...\phi(x_n)\rangle j(x_1)...j(x_n), \tag{2.16}$$

which can be proven by comparing both expressions in an expansion term-by-term. This reconstruction theorem can be readily generalized to theories with more than one field.

2.4 Free particles and generating functionals

As a useful first example, consider the Lagrangian of a free particle,

$$\mathcal{L} = \frac{1}{2}\phi(-\partial^2 - m^2)\phi$$

for a real, scalar, non-interacting particle. Note that a partial integration has been performed, where the boundary term vanishes. It is here implicitly assumed that the fluctuations at infinity are sufficiently independent to average any boundary term out. If there is an explicit boundary, or correlations across arbitrary distances should be described, this assumption may need revisiting. As the action is even in ϕ, this implies that any expectation value of an odd-power monomial of the field will vanish.

[5] There are subtleties involved, if the source breaks any symmetry explicitly. Then the limit of vanishing source may be different from the situation at zero source. This will not play a role in this lecture, but does sometimes in applications.

The two-point correlation function, the so-called propagator, is then the simplest, non-trivial and non-vanishing correlation function. For the scalar follows

$$\langle\phi(z)\phi(y)\rangle = \int \mathcal{D}\phi\phi(z)\phi(y)e^{iS+\int d^dxj\phi}\Big|_{j=0} = \frac{\delta^2}{\delta j(z)\delta j(y)}\int \mathcal{D}\phi e^{iS+\int d^dxj\phi}\Big|_{j=0}.$$

This can be either directly integrated, or by completing the square. The propagator is hence the inverse of the square root of the determinant of the Klein-Gordon operator

$$\Delta = \partial^2 + m^2. \tag{2.17}$$

Rather then to directly determine the determinant of this operator, it is better to switch to momentum space in the action, yielding

$$S = \int d^dp\phi(p)(p^2 - m^2)\phi(-p).$$

At first sight, the path integral is then not well defined, due to its highly oscillatory behavior. Just as with ordinary Fourier integration, this can be remedied by replacing m^2 by $m^2 - i\epsilon$, and eventually taking ϵ to zero. Using then (2.12) yields

$$\det(p^2 - m^2 + i\epsilon)^{-\frac{1}{2}}e^{\frac{-i}{2}\int d^dpj(p)^2(p^2-m^2+i\epsilon)} = e^{-\frac{i}{2}\int d^dp\ln(\int d^dpp^2-m^2+i\epsilon)-\frac{i}{2}\int d^dpj(p)^2(p^2-m^2+i\epsilon)}$$

This is used to define the free energy W as

$$e^{W[j]} = Z[j].$$

Because of (2.16) this implies that

$$\langle T\phi(q)\phi(-q)\rangle = \frac{1}{Z[0]}\frac{\delta^2 Z[j]}{\delta j(q)\delta j(-q)}\Big|_{j=0} = \frac{e^{W[j]}}{e^{W[0]}}\frac{\delta^2 W[j]}{\delta j(q)\delta j(-q)}\Big|_{j=0} = \frac{\delta^2 W[j]}{\delta j(q)\delta j(-q)}\Big|_{j=0}$$

Hence, the derivatives of the free energy are the correlation functions. In the present case

$$\langle T\phi(q)\phi(-q)\rangle = -i(q^2 - m^2 + i\epsilon).$$

As will become apparent later, it is also useful to define the so-called quantum effective action as the Legendre transformed of the free energy[6]. To this end, define the conjugate variable to the sources as the so-called classical field Φ,

$$\Phi = \frac{\delta W[j]}{\delta j}$$

[6]That a Legendre transform exists is highly non-trivial. However, it can be shown that it is possible for any well-defined quantum field theory. From this follows that the free energy and the quantum effective action are convex functionals of their arguments.

and then[7]

$$\Gamma[\Phi] = \int d^d p j(p)\Phi(p) - W[j(\Phi)].$$

The derivatives of the quantum effective action with respect to the classical fields are called the vertex functions. Especially, because of the existence of an inverse Legendre transformation

$$j = \frac{\delta\Gamma(\Phi)}{\delta\Phi} \tag{2.18}$$

it follows that

$$\delta^d(p-q) = \frac{\delta\Phi(p)}{\delta\Phi(q)} = \int d^d k \frac{\delta\Phi(p)}{\delta j(k)}\frac{\delta j(k)}{\delta\Phi(q)} = \int d^d k \frac{\delta^2 W}{\delta j(p)\delta j(k)}\frac{\delta^2\Gamma}{\delta\Phi(k)\delta\Phi(q)} \tag{2.19}$$

and therefore necessarily

$$\frac{\delta^2 W}{\delta j(p)\delta j(k)}\delta(p-k) = \left(\frac{\delta^2\Gamma}{\delta\Phi(k)\delta\Phi(q)}\right)^{-1}\delta(q-k). \tag{2.20}$$

It hence follows that

$$\frac{\delta^2\Gamma}{\delta\Phi(-q)\delta\Phi(q)} = \frac{i}{q^2 - m^2 + i\epsilon} = D(q) = \Gamma_2^{-1}(q)$$

is the so-called propagator.

These quantities will take on important meanings in the following, starting with chapter 3.

2.5 Internal symmetries and Ward identities

2.5.1 Symmetries

Symmetries play an important role in physics, and thus also in quantum field theory. A symmetry is defined to be an operation on the variables of the theory such that any observable quantity remains unchanged. Though their theory is more multilayered then often perceived.

Symmetries thus act in general as

$$\phi(x) \rightarrow \phi(x') + \delta(x') \tag{2.21}$$

$$x \rightarrow x' = x + \epsilon(x), \tag{2.22}$$

[7]As a Legendre transformation, it is necessary that Γ is a convex function of the classical fields. However, the existence of the Legendre transformation is not obvious, as this requires the free energy to be also a strictly monotone function of the sources, and thus $W[0]$ to be an absolute minimum. This follows from the fact that the total derivative of a path integral, and thus also with respect to the sources, vanishes, and the Hessian is positive.

in which the (vectors) δ and ϵ are taken to be infinitesimal. Finite transformations are build up from infinitesimal transformations[8]. Symmetries which change the coordinates are called external symmetries. If $\epsilon(x)$ is linear in x, this will be in this lecture exclusively Poincaré transformations. Their role will be analyzed in detail in chapter 4, but has already been briefly used in section 2.1. If $\epsilon(x)$ is non-linear in x, this will generically become a theory involving general relativity. This subject is addressed in the lecture on "Advanced General Relativity and Quantum Gravity".

Symmetries with $\epsilon = 0$ are called internal symmetries. If $\delta(x)$ is only linearly dependent on the fields, with space-time independent coefficients, it is a so-called global symmetry, or sometimes also called rigid symmetry. Otherwise it is called a local symmetry or gauge symmetry. The latter type of symmetries is subject of the lecture "Quantum Field Theory II". Symmetries, for which both ϵ and λ are simultaneously non-zero are supersymmetries, and are discussed in the lectures on "Beyond the Standard Model physics".

For a transformations like (2.21-2.22) to be a symmetry it is necessary that it leaves the generating functional (2.7) invariant. The reason is that the partition function can be associated with relative differences in free energies (thus the introduction of W) or pressure, and is hence observable. This is different from classical physics, where a symmetry needs to leave only the action invariant. Very often, the Lagrangian alone is already invariant. However, not always.

As an example of symmetries, consider again the free scalar field (2.1). However, duplicate the number of degrees of freedom, by replacing the single field ϕ by a two-dimensional vector ϕ_i, giving the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i.$$

Acting with a space-time-independent orthogonal matrix O on the two-dimensional, real vector $\vec{\phi}$, yields

$$(O\vec{\phi}) = O_{ij}\phi_j \approx \phi_i + \omega_{ij}\phi_j + \mathcal{O}(\omega^2) \tag{2.23}$$

where the infinitesimal form in the second step makes use of the fact the two-dimensional orthogonal groups form a Lie group, and thus can be expanded in this way. This will leave the action invariant, since O is space time-independent, and thus commutes with the derivatives,

$$\frac{1}{2}\partial_\mu O_{ij}\phi_j\partial^\mu O_{ik}\phi_k = (O^T O)_{jk}\frac{1}{2}\partial_\mu\phi_j\partial^\mu\phi_k = \frac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i,$$

and where it was used in the last step that for orthogonal matrices $O^T O = 1$. Hence, the Lagrangian and the action are invariant.

[8]There are some subtleties involved, which will not play a role in this lecture. This is taken up again in the lecture "Quantum Field Theory II".

However, for the partition function also the measure $\mathcal{D}\phi_i$ needs to be invariant. Since (2.23) can be regarded as a variable transformation, this requires that the Jacobi determinant is unity. As the transformation is linear, the multi-component generalization $\eta(x,y)_{ij}$ in (2.15) is $\delta(x-y)O$ and $\det O$ is ± 1. Thus, the path integral could change sign. However, because of the normalization of the path integral by $Z[0]$, this will not change any correlation function, and thus all observables and the generating functional remain unchanged. If the determinant would have spoiled this, it is called an anomaly, a symmetry broken by the quantization. This indeed happens, as will be explored in the lecture "Quantum Field Theory II", and indeed occurs and has observable consequences, as discussed in the lecture "The Standard Model".

A useful consequence is the existence of so-called Ward-Takashi identities, which relate correlation functions. To understand it, this requires a few more preliminary remarks. If the transformations (2.21-2.22) form a group, which is almost always the case, the group is called the symmetry group of the symmetry. Then, the field form an orbit of the group, and any element of the orbit is physically equivalent. Any element is thus a representative of the orbit. Since the path integral is defined to integrate over all field configurations, it necessarily integrates over the orbits as well.

But this implies that for any symmetry group for which the integral over the group orbit g with representation $T(\alpha_i)$

$$0 = \int dg\phi^g = \int \Pi_i d\alpha_i T(\alpha_i)_{jk}\phi_k \tag{2.24}$$

holds, where the α_i are parameters which allow to access all group elements, can have only non-vanishing correlation functions if the combination of fields involved do create an invariant under the symmetry. This applies, e. g., to all Lie groups. Conversely, for

$$\langle T a_{i_1\dots i_n}\phi_{i_1}(x_1)\dots\phi_{i_n}(x_n)\rangle \neq 0 \tag{2.25}$$

to hold, a must be an invariant tensor. To see this, split the integration over the field into a representative ϕ^R and the integral over the group orbit g. For a non-invariant correlation function this implies

$$\langle T\phi_i\dots\rangle = \int \mathcal{D}\phi^R dg\phi_i\dots e^{iS} = \int \mathcal{D}\phi^R \Pi_j d\alpha_j T(\alpha_j)_{ik}\phi_k\dots e^{iS} = 0. \tag{2.26}$$

Hence, only invariants can be non-zero.

However, when including a source term, the source is external, and thus does not transform under the symmetry. Therefore, a source term breaks the symmetry necessarily explicitly. Especially, in presence of a source term even non-invariant correlation functions

remain non-zero. When taking the limit of the source to zero, they may again vanish or not. In the latter case the system is non-analytic in the source, a phenomenon know as spontaneous symmetry breaking. This will play an important role in phenomenology.

However, there is another way to consider the structure. Because an invariant tensor will not depend on the fields, it appears valid to write

$$\langle T a_{i_1\dots i_n}\phi_{i_1}(x_1)\dots\phi_{i_n}(x_n)\rangle = a_{i_1\dots i_n}\langle T\phi_{i_1}(x_1)\dots\phi_{i_n}(x_n)\rangle. \tag{2.27}$$

But, by virtue of (2.26), this seems to contradict (2.25).

What happens can be understood in terms of the Wigner-Eckart theorem. Any tensor under a given symmetry can be decomposed in a basis of invariant tensors $\{b^a\}$ and scalar functions s^a, and thus necessarily

$$\phi_i\dots\phi_n = \sum_a b^a_{i\dots n}s^a(\phi^2). \tag{2.28}$$

Multiplying thus with an invariant tensor a, this will project out the corresponding sum of the s^a, based on how a decomposes into the b^a. Then, this becomes transparent, as this separates the ϕ^R integration and the α integration in (2.26) as

$$\int \mathcal{D}\phi^R \Pi_j d\alpha_j T(\alpha_j)_{ik}\phi_k\dots e^{iS} = \sum_a \left(\int \Pi_j d\alpha_j T(\alpha_j)_{ik} b^a_{i\dots}\right)\left(\int \mathcal{D}\phi^R s^a((\phi^R)^2)e^{iS}\right)$$

Thus, the first factor vanishes, while the second does not. Now it becomes evident, why (2.27) is actually not valid. While the invariant tensor does not depend on the fields, it does depend on the parametrization of the orbits and the invariant tensors used to write down the expressions. By pulling it out, this information is lost, and the result vanishes. Essentially, what happens is that the orbit integration also includes all possible changes of basis, and the a would be needed to be changed accordingly. Thus, (2.27) it is not a valid operation.

Even though therefore only scalar quantities can be non-zero, it is often very convenient to act as if also correlation functions can be written as (2.28),

$$\Gamma_{i_1\dots i_n}(x_1,\dots,x_n) \equiv \langle T\phi_{i_1}(x_1)\dots\phi_{i_n}(x_n)\rangle = \sum_a b^a_{i_1\dots i_n}s^a(x_1,\dots,x_n). \tag{2.29}$$

However, this should be regarded as a short-hand notation for

$$\Gamma_{i_1\dots i_n}(x_1,\dots,x_n) = \sum_a b^a_{i_1\dots i_n}\langle T b^a_{j_1\dots j_n}\phi_{j_1}(x_1)\dots\phi_{j_n}(x_n)\rangle.$$

While so far only scalar particles have been considered, later on this will change. As a consequence, the fields will also have Lorentz indices, and transform non-trivially under Lorentz transformations. Then, correlation functions for this external symmetry need to be treated likewise.

2.5.2 Ward-Takahashi identities

Not withstanding these subtleties, symmetries can be used to derive powerful relations between correlation functions. Let a theory be invariant under the change

$$\phi \to \phi' = \phi + \delta\phi = \phi + \epsilon f(\phi, x), \tag{2.30}$$

with ϵ infinitesimal. Then the generating functional $Z[j]$ should not change, i. e., δZ should be zero. This can be analyzed using the general behavior of a quantity F under the variation δ,

$$\delta F(\phi) = \frac{\delta F}{\delta\phi}\delta\phi = \frac{\delta F}{\delta\phi}\epsilon f.$$

In case of Z, it acts on two components in the path integral. One on the action itself, which yields

$$\frac{1}{\epsilon}\delta\left(e^{iS+i\int d^dxj\phi}\right) = i\left(\frac{\delta S}{\delta\phi} + j\right)\epsilon f e^{iS+i\int d^dxj\phi},$$

to first order in ϵ. The second is the measure. As mentioned before, the shift (2.30) is a variable transformation, which generates a Jacobian determinant. This Jacobian determinant can also be expanded in ϵ, yielding

$$\det\frac{\delta\phi'}{\delta\phi} = \det\left(1 + \frac{\delta\epsilon f}{\delta\phi}\right) = 1 + \epsilon\frac{\delta f}{\delta\phi} + \mathcal{O}(\epsilon^2).$$

Together, this yields the variation

$$0 = \delta Z = \int \mathcal{D}\phi\epsilon\left(\frac{\delta f}{\delta\phi} + i\left(\frac{\delta S}{\delta\phi} + j\right)f + \mathcal{O}(\epsilon)^2\right)e^{iS+i\int d^dxj\phi}. \tag{2.31}$$

Note that if the term from the Jacobian is non-vanishing as $\epsilon \to 0$, this yields a so-called anomaly, and potentially breaks the symmetry. If this is the case, the original assumption that $\delta Z = 0$ is no longer valid, and the following does not work[9].

Differentiating this expression once with respect to the source and setting the sources afterwards to zero yields an expression connecting different correlation functions. E. g., performing a single derivative will yield

$$\left\langle T\phi(y)\frac{\delta f(\phi, x)}{\delta\phi(x)}\right\rangle + i\left\langle\phi(y)\frac{\delta S}{\delta\phi(x)}f\right\rangle + \langle Tf\rangle = 0.$$

In general, there will not only be one field involved, but many fields, numerated by a field index i. In this case, expression (2.31) takes the form

$$0 = \int \mathcal{D}\phi_i\left(\frac{\delta f_k}{\delta\phi_k} + i\left(\frac{\delta S}{\delta\phi_k} + j_k\right)f_k\right)e_k^{iS+i\int d^dxj_k\phi},$$

[9] In fact, this yields so-called broken or modified Ward-Takahasi identities, which, however, is beyond the scope of this lecture.

i. e., it becomes a sum over all fields. Deriving this expression in total n times for any sequence of field types i_l yields the set of all Ward-Takahashi identities, sometimes also called rigid identities given that global symmetries are sometimes called rigid symmetries,

$$\left\langle T\Pi_{l=1}^{n}\phi_{i_l}(x_l)\frac{\delta f_k}{\delta\phi_k(y)}\right\rangle + i\left\langle T\Pi_{l=1}^{n}\phi_{i_l}(x_l)\frac{\delta S}{\delta\phi_k(y)}f_k\right\rangle + \sum_{m=1}^{n}\left\langle \Pi_{l=1}^{m-1}\phi_{i_l}(x_l)f_{i_m}\Pi_{r=m+1}^{n}\phi_{i_r}(x_r)\right\rangle = 0. \tag{2.32}$$

These relate different correlation functions with each other especially also those with a different number of fields. The Ward-Takashi identities are an expression of the redundancy introduced by the global symmetry. Similar to the Wigner-Eckart theorem, they allow to eliminate the redundancy by relating different correlation functions.

To obtain practical cases requires to insert an action with a certain invariance. Take again as an example the action for the σ-model (or ϕ^4 model) with positive mass squared,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^\dagger\partial^\mu\phi + \frac{1}{2}m^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2.$$

The transformation function is then $f_i = \mp i\phi_i$, where $i = 1$ refers to ϕ and $i = 2$ refers to $\phi^\dagger$. The derivative of f actually vanishes in this case, since the Jacobian matrix under a linear shift of the fields is zero, by the definition of translational invariance of the path integral (2.13).

Furthermore, the action is invariant under the global symmetry transformation. This implies

$$\frac{\partial S[\phi_i + \epsilon f_i]}{\partial\epsilon} = 0 = \int d^dx\frac{\delta S}{\delta\phi_i}\frac{\partial(\phi_i + \epsilon f_i)}{\partial\epsilon} = \int d^dx\frac{\delta S}{\delta\phi_i}f_i,$$

and thus also the second term in (2.31) vanishes. Hence, only the third term remains, which can be conveniently written as

$$0 = \delta\left\langle T\Pi_{l=1}^{n}\phi_{i_l}\right\rangle, \tag{2.33}$$

which are called Ward identities in this context. E. g., at level $n = 2$, this identity implies

$$\langle T(\delta\phi(x))\phi(y)^\dagger\rangle + \langle T\phi(x)\delta\phi(y)^\dagger\rangle = \langle\phi(x)\phi(y)^\dagger\rangle - \langle\phi(x)\phi(y)^\dagger\rangle = 0,$$

which seems rather trivial. However, when rewriting the theory in terms of a scalar σ field and a pseudoscalar χ field, i. e. $\phi = \sigma + i\chi$, this implies

$$\langle T\delta\sigma\chi\rangle + \langle T\sigma\delta\chi\rangle = \langle\chi\chi\rangle - \langle\sigma\sigma\rangle = 0,$$

which implies that the propagators of both fields are identical. At tree-level, this is immediately visible, but gives a constraint for the results beyond tree-level. Of course, this is a rather simple result, and much more interesting ones are obtained at higher order and/or for more complicated theories.

Chapter 3

Perturbation theory

To actually determine correlation functions, it appears necessary to evaluate the path integral. Except for non-interacting theories, a few 2-dimensional field theories, and a few theories with extremely strong symmetries, this is so far not exactly possible. Thus, approximation methods are required. As so often, there are a wide range of such methods available. However, the mainstay, as in quantum mechanics, is very often perturbation theory. Also, many concepts from perturbation theory are essential to go beyond perturbation theory eventually. Thus, here perturbation theory will be treated first. In chapter 9 approximations transcending the bounds of perturbation theory will be introduced.

3.1 Perturbative expansion and Feynman rules

Correlation functions can be calculated using the path integral as

$$\langle T\phi_1...\phi_n\rangle = \left.\frac{\int \mathcal{D}\phi\phi_1...\phi_n e^{iS[\phi,J]}}{\int \mathcal{D}\phi e^{iS[\phi,J]}}\right|_{J=0}. \tag{3.1}$$

However, this is so far only a tautology, as this gives no constructive way of calculating actually the correlation functions. As in quantum mechanics, in many cases the correlation functions are quantitatively well described by small perturbations from a known, usually exactly soluble, case. To calculate the differences, perturbation theory is again useful. However, its formulation is slightly different, due to manifest Lorentz covariance. Especially, time-independent perturbation theory is usually not useful, as space-time are linked by Lorentz boosts.

The starting point will be expression (3.1). Perturbation theory will boil down essentially to expanding the exponential in the fields, giving an infinite series of quasi-Gaussian integrals. The result is that the transition matrix elements are determined by a sum over

correlation functions in a theory with quadratic action. Such an expansion of the field is most often an expansion around free fields, and thus assumes that the field amplitudes are small and thus interaction terms small compared to the kinetic terms. It is thus a saddle-point approximation around the free system, and can thus also be considered as an expansion in quantum corrections. It will later be seen that this is manifest as it is also a power series in $\hbar$.

To do this, split the Lagrangian into a quadratic part $\mathcal{L}_2$ and a remainder part $\mathcal{L}_I$, which includes all the interactions. This yields for the generating functional

$$\begin{aligned} Z\left[J\right] &= \int \mathcal{D}\phi e^{i\int d^dx\mathcal{L}_I} e^{i\int d^dx(\mathcal{L}_2+J\phi)} \\ &= e^{i\int d^dx\mathcal{L}_I\left[\frac{\delta}{i\delta J}\right]} \int \mathcal{D}\phi e^{i\int d^dx(\mathcal{L}_2+J\phi)}. \end{aligned}$$

This is a rewriting of the expression[1] with the operator-valued exponential defined in terms of a power series. The argument of $\mathcal{L}_I$ is just indicating that all appearances of the field have been replaced by the derivative with respect to the source. To see the equivalence, take as an example a theory with cubic interaction term

$$\mathcal{L}_I = \frac{\lambda}{3!}\phi^3$$

and expand the exponential

$$\begin{aligned} & e^{i\int d^dx\mathcal{L}_I\left[\frac{\delta}{i\delta J}\right]} \int \mathcal{D}\phi e^{i\int d^dx(\mathcal{L}_2+J\phi)} \\ = & \int \mathcal{D}\phi \sum_n \frac{1}{n!}\left(\frac{\lambda}{3!}i\int d^dy \frac{\delta^3}{i\delta J(y)^3}\right)^n e^{i\int d^dx(\mathcal{L}_2+J\phi)} \\ = & \int \mathcal{D}\phi \left(1 + i\int d^dy \frac{\lambda}{3!}\frac{\delta^2}{i\delta J(y)^2}\frac{\delta i\int d^dxJ\phi}{i\delta J(y)} + ...\right) e^{i\int d^dx(\mathcal{L}_2+J\phi)} \\ = & \int \mathcal{D}\phi \left(1 + i\int d^dy \frac{\lambda}{3!}\frac{\delta^2}{i\delta J(y)^2}\int d^dx\phi\delta(x-y) + ...\right) e^{i\int d^dx(\mathcal{L}_2+J\phi)} \\ = & \int \mathcal{D}\phi \left(1 + i\int d^dy \frac{\lambda}{3!}\phi\frac{\delta^2}{i\delta J(y)^2} + ...\right) e^{i\int d^dx(\mathcal{L}_2+J\phi)} \\ = & \int \mathcal{D}\phi \left(1 + i\int d^dy \frac{\lambda}{3!}\phi^3 + ...\right) e^{i\int d^dx(\mathcal{L}_2+J\phi)} \\ = & \int \mathcal{D}\phi \sum_n \frac{1}{n!}\left(\frac{\lambda}{3!}i\int d^dy\phi^3(y)\right)^n e^{i\int d^dx(\mathcal{L}_2+J\phi)} \\ = & \int \mathcal{D}\phi e^{i\int d^dx\mathcal{L}_I} e^{i\int d^dx(\mathcal{L}_2+J\phi)}. \end{aligned}$$

[1] Which is actually only approximately valid, as will be discussed in chapter 9.

Such manipulations are very helpful in general.

To proceed it is necessary to perform the remaining shifted Gaussian integral. Following section 2.4, this yields

$$\begin{aligned}\int D\phi e^{i\int d^dx(\phi(x)(\Omega-i\epsilon/2)\phi(x)+J(x)\phi(x))} &= \int D\phi e^{i\int d^dx(\phi(x)\Omega\phi(x))} e^{-\frac{i}{2}\int d^dxd^dyJ(x)\Delta(x-y)J(y)} \\ &= Z_2\left[0\right] e^{-\frac{i}{2}\int d^dxd^dyJ(x)\Delta(x-y)J(y)}. \end{aligned} \tag{3.2}$$

There are a number of points to take into account. Ω is just the quadratic part of the Lagrangian, e. g., for a free scalar field it is just $(-\partial^2 - M^2)/2$. The addition of the term $i\epsilon$ is actually needed to make the integral convergent, and has to be carried through all calculations. Next, the so-called Feynman propagator Δ is defined such that

$$(2\Omega - i\epsilon)\Delta(x-y) = i\delta^d(x-y),$$

i. e. it is the Green's function of differential operator 2Ω. That it depends only on the difference $x - y$ comes from the assumption of translational invariance, which applies for unbroken Poincare symmetry. To solve this differential equations requires boundary conditions. In perturbation theory, it is assumed that the fields fall off sufficiently fast towards (spatial) infinity. That assumption can be lifted in other methods.

For a scalar particle of mass M and thus $\Omega = (-\partial^2 - M^2)/2$ the Feynman propagator can be calculated using Fourier transformation,

$$\begin{aligned}(-\partial^2 - M^2 + i\epsilon)\int d^dp e^{ip(x-y)}\Delta(p) &= i\int d^dp e^{ip(x-y)} \\ \int d^dp e^{ip(x-y)}(p^2 - M^2 + i\epsilon)\Delta(p) &= i\int d^dp e^{ip(x-y)}\end{aligned} \tag{3.3}$$

to yield

$$\Delta(p) = \frac{i}{p^2 - M^2 + i\epsilon}, \tag{3.4}$$

which is more useful for a calculation than the rather involved position space expression, which can be obtained by reverting the Fourier transformation. E. g., in four dimensions it is given by

$$\Delta(x-y) = \frac{\delta((x-y)^2)}{4\pi} + \frac{i}{4\pi^2}\frac{m}{\sqrt{2(x-y)^2 + i\epsilon}}K_1\left(m\sqrt{2(x-y)^2 + i\epsilon}\right), \tag{3.5}$$

where K_1 is the modified Bessel function of the first kind. However, there is one interesting information to be gleaned from here. The function decays exponential over space-like

distances, just as with a quantum-mechanical wave-function in classically forbidden regions. It thus manifests the property that special relativity does not transfer information at space-like distances: The correlation of fields is quickly diminishing over space-like distance. On the contrary, for time-like distances it oscillates, and thus allows for interference of the fields, just like wave functions. Finally, there is a curious divergence for coinciding space-time points. This will play an essential role for the limitations of a given quantum-field theory, especially in section 8.2.

Finally, the factor $Z_2[0]$ in front of the integral containing the Feynman propagator is just the factor $1/a$ in the conventional integral (2.11), conveniently rewritten as an exponential. This factor will cancel partly the denominator in (3.1) when taking the limit $J \to 0$ at the end of the calculation.

This is then sufficient to write down a perturbative expression for an arbitrary correlation function. Take, for example, the linear-σ model of section 2.2 with

$$\mathcal{L}_I = -\frac{\lambda}{4!}\phi^4.$$

The perturbative expression up to linear order in $\lambda/4!$ for a process involving two particles in the initial and final state, which will be needed for ((in)elastic) scattering later is then

$$\begin{aligned}
&\langle T\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle \\
= &\left.\frac{\int \mathcal{D}\phi\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)e^{iS[\phi,J]}}{\int \mathcal{D}\phi e^{iS[\phi,J]}}\right|_{J=0} \\
= &\left.\frac{1}{Z[0]}\frac{\delta^4}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)}\int \mathcal{D}\phi e^{iS[\phi,J]}\right|_{J=0} \\
= &\left.\frac{Z_2[0]}{Z[0]}\frac{\delta^4}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)}e^{i\int d^dx\mathcal{L}_I\left[\frac{\delta}{i\delta J}\right]}e^{-\frac{i}{2}\int d^dxd^dyJ(x)\Delta(x-y)J(y)}\right|_{J=0}.
\end{aligned}$$

The next step is to expand both exponentials, the first in a formal power series in $\mathcal{L}_I$, and the second one in the conventional exponential series,

$$\begin{aligned}
= &\frac{Z_2[0]}{Z[0]}\frac{\delta^4}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)}\left(\sum_n \frac{1}{n!}\left(\frac{\lambda}{4!}i\int d^dy\frac{\delta^4}{i\delta J(y)^4}\right)^n\right)\times \\
&\times\left(\sum_m \frac{1}{m!}\left(-\frac{i}{2}\int d^dxd^dyJ(x)\Delta(x-y)J(y)\right)\right)\Bigg|_{J=0}.
\end{aligned}$$

Both are polynomial in the sources. The expansion of the exponential of the interaction Lagrangian yields terms with zero, four, eight,... derivatives with respect to the sources. The second term produces terms with zero, two, four,... powers of the sources. Since the sources are set to zero at the end, only terms without sources will remain. Reinstantiating

factors of $\hbar$, it is immediately visible that every power of λ comes with a power of $\hbar$. This justifies the interpretation that it is also simultaneously an expansion in quantum fluctuations.

Thus, to order zero in the interaction Lagrangian only the term with four sources will survive the external derivative. To first order in the interaction Lagrangian only the term with eight powers of the sources will survive. To this order in the expansion, the expression takes therefore the form

$$\begin{aligned} = & \frac{Z_2[0]}{Z[0]} \frac{\delta^4}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)} \left(\frac{1}{2!} \left(-\frac{i}{2} \int d^d x d^d y J(x) \Delta(x-y) J(y) \right)^2 \right. \\ & \left. -\frac{i\lambda}{4!} \int d^d z \frac{1}{4!} \frac{\delta^4}{\delta J(z)^4} \left(-\frac{i}{2} \int d^d x d^d y J(z) \Delta(x-y) J(y) \right)^4 + \mathcal{O}(\lambda^2) \right). \end{aligned}$$

In principle, taking the derivatives is straight-forward. However, e. g., the first term is given by the expression

$$\frac{\delta^4}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)} \int d^d y_1 d^d y_2 d^d y_3 d^d y_4 J(y_1)\Delta(y_1-y_2)J(y_2)J(y_3)\Delta(y_3-y_4)J(y_4).$$

The first derivative, with respect to $J(x_4)$ could act equally well on all four sources under the integral. It will therefore provide four terms. Correspondingly, the second derivative can act on three different terms, making this 12 terms, and so on, giving in total 24 terms, with all possible combinations, or partitions, of the four arguments.

To illustrate the process, two steps for a particular combination will be investigated. The first derivative acts as

$$\begin{aligned} & \frac{\delta}{\delta J(x_4)} \int d^d x d^d y J(x) \Delta(x-y) J(y) \\ = & \int d^d x d^d y \delta^d(x-x_4) \Delta(x-y) J(y) + ... = \int dy \Delta(x_4-y) J(y) + ..., \end{aligned} \tag{3.6}$$

where the points indicate further contributions. For the action of the next derivative, there are two possibilities. Either it acts on the same factor of the product of the integrals, or on a different one. Take first the possibility of the same factor. If it is a distinct factor, this just provides the same action. If it is the same factor, this immediately yields

$$\frac{\delta}{\delta J(x_3)} \int d^d y \Delta(x_4-y) J(y) = \int d^d y \Delta(x_4-y) \delta^d(y-x_3) = \Delta(x_4-x_3)$$

In total, this yields for the term proportional to $\lambda^0 = 1$

$$A = -\sum_{P_{ijkl}} \Delta(x_i-x_j)\Delta(x_l-x_k), \tag{3.7}$$

where P_{ijkl} indicates that the sum is over all 4! possible permutations of the index set $\{ijkl\}$.

The situation becomes somewhat more complicated for the terms proportional to λ, since now multiple derivatives with respect to the same source $J(x)$ appears. Again, a single such derivative acts like (3.6). A difference occurs when the second derivative occurs. This can either act again on another factor, but it could also act on the same factor. The first case just produces another factor of type (3.6). The second situation is different, and yields

$$\frac{\delta}{\delta J(z)} \int d^d y \Delta(z-y) J(y) = \int d^d y \Delta(z-y) \delta^d(y-z) = \Delta(z-z) \tag{3.8}$$

which appears to look like $\Delta(0)$. However, this not quite the case, as will be visible later. In particular, the expression $\Delta(0)$ cannot be easily interpreted, as (3.5) shows. Furthermore, an integral over z still appears. It is therefore useful to keep explicitly terms of $\Delta(z-z)$ in the following for now.

After a slight change in notation, there will then be 8! possibilities for the order λ contribution. However, many of them turn out to be identical, yielding in total three further contributions

$$\lambda B = -i\lambda \int d^d x \Delta(x-x_1)\Delta(x-x_2)\Delta(x-x_3)\Delta(x-x_4) \tag{3.9}$$

$$\lambda C = -\frac{i\lambda}{2} \sum_{P_{ijkl}} \Delta(x_i - x_j) \int d^d x \Delta(x-x)\Delta(x-x_k)\Delta(x-x_l) \tag{3.10}$$

$$\lambda AD = -\frac{i\lambda}{8} \int d^d x \Delta(x-x)\Delta(x-x) \sum_{P_{ijkl}} \Delta(x_i-x_j)\Delta(x_k-x_l). \tag{3.11}$$

These four terms have simple interpretations, if each factor of Δ is considered to be a particle[2] propagating along the connecting line of $x-y$. Then, the first term (3.7) corresponds to the interference pattern of identical particles when they are observed at two different initial and final positions: Since the particles are identical, any combination is possible, including that one particle vanishes and the other one appears. This can be visualized by using a line to symbolize a factor of Δ, and draw all possible combinations between the four points.

Similar interpretations hold for the three remaining terms (3.9-3.11). The expression (3.9) contains for each factor of Δ a common point. This can be taken to be just a meeting of all four particles at a common vertex point x. Since there appears a pre-factor of λ, it

[2]It is formally the correlation of a free field with itself over the distance $(x-y)^2$. This can, however, be interpreted as a particle, as will be seen in section 8.5. Therefore, this standard jargon will be kept in the following.

can be said that the four particles couple with a strength λ, thus also the name coupling constant for λ. Such an interaction vertex could be denoted by a dot.

The third term (3.10) can be seen as one particle just propagating, while the second particle has an interesting behavior: It emits at an intermediate point a particle, and reabsorbs it then. Such a emitted and reabsorbed particle is called a virtual particle and contributes to a cloud of virtual emission and absorption processes, which becomes more common and more relevant at higher orders. Pictorially, this corresponds to a loop in the propagation, which again harbors an interaction vertex.

The last contribution is different, as when drawing lines there appears an additional graph, which is disconnected from the initial and final positions, and has the form of the number eight. Such a disconnected diagram is also called a disconnected contribution or a vacuum contribution, as it is not connected to any external input, and is thus a property of the vacuum alone.

In general, the expression (3.7-3.11) are very cumbersome to deal with in position space. It is therefore more useful to perform a Fourier transformation, and perform the calculations in momentum space. In particular, this removes many of the cumbersome sums over partitions. How to switch to momentum space will be discussed in more detail after taking care of the remaining factor $Z_2[0]/Z[0]$.

Since the current calculation is a perturbative calculation, it is adequate to also expand $Z_2[0]/Z[0]$ in λ. This can be most directly done again using the formula (3.2). Thus, the factor Z_2 cancels immediately, and the remaining expansion terms are, up to combinatorial factors, very similar as before. Its inverse is thus given, to order λ, by

$$\frac{Z[0]}{Z_2[0]} = 1 + \frac{i\lambda}{2^3}\int d^dx \Delta(x-x)\Delta(x-x) + \mathcal{O}(\lambda^2) = 1 + \lambda D.$$

This term is easily identified as the prefactor appearing in (3.11). To order λ, this yields

$$\begin{aligned} & \langle T\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle = \frac{A+\lambda(B+C+AD)}{1+\lambda D} + \mathcal{O}(\lambda^2) \\ = \;& (A+\lambda(B+C+AD))(1-\lambda D) + \mathcal{O}(\lambda^2) = A+\lambda(B+C+AD) - \lambda AD + \mathcal{O}(\lambda^2) \\ = \;& A+\lambda(B+C) + \mathcal{O}(\lambda^2). \end{aligned} \tag{3.12}$$

Thus, to order λ, the term with a disconnected contribution is canceled. It turns out that this is a generic result, and that all diagrams with disconnected contribution in a perturbative expansion always cancel, and a general proof can be constructed in a very similar way to this evaluation in ϕ^4 theory up to leading order, essentially as an inductive proof. However, this is beyond the scope of this lecture.

As stated, the explicit expression in position space turns out to be very awkward to use in actual calculations, and their evaluation in momentum space is preferable. This

can be done using the expression for the Feynman propagator in momentum space, (3.4). The total sum then becomes

$$
\begin{aligned}
&(2\pi)^d \int \frac{d^dp_1}{(2\pi)^{\frac{d}{2}}}\frac{d^dp_2}{(2\pi)^{\frac{d}{2}}}\frac{d^dp_3}{(2\pi)^{\frac{d}{2}}}\frac{d^dp_4}{(2\pi)^{\frac{d}{2}}} e^{-i(p_1x_1+p_2x_2+p_3x_3+p_4x_4)} \times \\
&\times \left(\sum_{P_{ijkl}} (2\pi)^d \delta^d(p_k+p_l)\delta^d(p_i+p_j)\frac{i}{p_i^2-m^2}\frac{i}{p_k^2-m^2} \right. \\
&-i\lambda\delta^d(p_1+p_2+p_3+p_4)\frac{i}{p_1^2-m^2}\frac{i}{p_2^2-m^2}\frac{i}{p_3^2-m^2}\frac{i}{p^2-m^2} \\
&\left. -\frac{(2\pi)^d\lambda}{2}\sum_{P_{Ijkl}} \delta^d(p_i+p_j)\delta^d(p_k+p_l)\frac{i}{p_i^2-m^2}\frac{i}{p_k^2-m^2}\frac{i}{p_l^2-m^2}\int\frac{d^dq}{(2\pi)^{\frac{d}{2}}}\frac{i}{q^2-m^2} \right).
\end{aligned}
$$

Note that the $i\epsilon$ contributions have not been written explicitly in the propagators, but left implicit. This is the standard conventions for such a representation of a perturbative expression. Of course, if the result is desired in momentum space rather than position space, which is normally the case, the Fourier transformation can be dropped.

The result already shows a number of regularities, which can be generalized to the so-called Feynman rules, which permit to directly translate from a graphical representation to the mathematical expression in perturbation theory. These can be derived rather generally, though this becomes rather cumbersome. Here, these will be stated simply without proof:

- Select the type and number of all external lines
- Determine the order (in all coupling constants, i. e., in all vertices) to which the process should be evaluated
- Draw all possible diagrams connecting in all possible ways the external lines with up to order vertices, and add them
- For each line, write a propagator of this particle type
- For each vertex, write the interaction vertex, i. e. essentially $\delta^n\mathcal{L}_I/\delta\phi^n$, for each
- Impose the conservation of all quantities, including momentum, conserved by a giving vertex at each vertex. This can be most directly done by following each input conserved quantity through the whole diagram until its final result
- Integrate over all undetermined momenta, i. e., each momentum running through a loop

- Lines, which are attached to the outside of a diagram, receive a further propagator of the corresponding type

Two things can further facilitate the result. On the one hand, any diagram will be zero, if any conservation law is not respected by the transition from initial to final state. However, in such cases also the perturbative vertices vanish identically. Secondly, there are many diagrams, which are identical up to reordering, as in the previous example. They can be collected, and yield as pre-factors so-called symmetry factors, which are essentially combinatorial factors. They can be obtained by counting all possible ways of how a diagram can be drawn by connecting the external lines.

It can be immediately shown that the previous results can be obtained from these rules, as an explicit example of the more general Feynman rules. In general, the necessary propagators and vertices can be derived from the classical action as

$$i\Gamma^{\phi_1^{a_1}\ldots\phi_n^{a_n}}(x_1,...,x_n) = \left.\frac{i\delta^n S}{\delta\phi_1^{a_1}(x_1)...\delta\phi_n^{a_n}}\right|_{\phi_i=0}, \tag{3.13}$$

which afterwards can be transformed to momentum space by a Fourier transformation. The conservation of momenta and other quantum numbers will then appear as suitable δ-functions, which are usually factored out and manually imposed afterwards. For the propagators, it is still necessary to invert the results, see section 2.4.

Perturbative calculations can be further simplified by passing to connected, amputated diagrams.

The so-called connected diagrams are diagrams in which all lines are connected with each other. In the previous case, the result can be symbolically written as

$$\Delta\Delta + \Delta\Delta' + \Pi, \tag{3.14}$$

where Δ is a propagator, Δ' is a propagator with a loop attached, and Π is the graph where all four lines are connected. This is called the full correlation function. Of course, Δ and Δ' can also be determined from the two-point function $\langle T\phi\phi\rangle$, to the same order, and therefore contain no new information. The only new contribution for the four-point function at this order of perturbation theory is Π. It would therefore be useful, if it is possible to only calculate this contribution, instead of the whole one. Indeed, it can be shown that for a correlation function with n external legs

$$\begin{aligned} G(x_1,...,x_n) &= G_c(x_1,...,x_n) + \sum G_c(x_i,...,x_j)G_c(x_j,...,x_k) \\ &\quad + \sum G_c(x_i,...,x_j)G_c(x_k,...,x_l)G_c(x_m,...,x_n) + ... \end{aligned} \tag{3.15}$$

where the sums are over all possible ways to split the index set $\{x_i\}$ in two, three,... subsets. Furthermore, every connected correlation function G_c, i. e. anyone which cannot

be written as a product of two or more other connected correlation functions, is a series in the coupling constant. Thus, in the present case,

$$\begin{aligned} G(x_1,x_2,x_3,x_4) &= G_c(x_1,x_2,x_3,x_4) + \sum_{P_{ijkl}} G_c(x_i,x_j)G_c(x_k,x_l) \\ G_c(x,y) &= \Delta(x-y) + \Delta'(x-y), \end{aligned}$$

where again the Δ is the propagator, and Δ' is the propagator to order λ, which includes the attached loop, and G_c is the only diagram with all points connected. Contributions proportional to Δ'^2 have to be dropped, as they are of higher order in the perturbative expansion. This relation can be inverted to obtain the connected correlation functions from the other, but it is more interesting to calculate just the connected, and then calculate the complete one by the formula (3.15).

Finally, all external lines have the propagators attached to them, they are called non-amputated. Removing this yields the amputated correlation functions Γ, which can immediately yield again the non-amputated one. Thus, it is sufficient to calculate the amputated ones. In the same way, explicit momentum conserving factors can always be reinstantiated.

Thus, the calculation of the four-point function boils finally down to the calculation of the amputated, connected two-point function to order λ, and the amputated, connected four-point function of order λ. These are just given by

$$\begin{aligned} \Gamma_c(p,q) &= -i\lambda \int \frac{d^d r}{(2\pi)^4} \frac{i}{r^2 - m^2} \\ \Gamma_c(p,q,k,l) &= -i\lambda, \end{aligned} \tag{3.16}$$

rather simple expressions indeed. Only missing is the explicit form of the contribution from Δ', i. e. Γ_c. However, it requires renormalization, and will be discussed in section 8.2. These are also called proper or vertex correlation functions. In fact, for the four-point vertex function to leading order in λ, the entire result is given by (3.16).

There is a further possibility to reduce the effort of perturbative calculations, though these do not reduce it further for the present example. It is rather simple to imagine situations, were it is possible to cut a single internal line to obtain two separate graphs. Such graphs are called one-particle reducible. It can be shown that it is sufficient two know all graphs, which cannot be separated in such a way, so-called one-particle irreducible graphs (1PI), to obtain all relevant results, and to reconstruct also the one-particle reducible ones. The generic connection can again be illustrated. Take two graphs which are 1PI, say graphs $A(p,q)$ and $B(k,l)$. They can be joined to a one-particle reducible graph by

$$A(p,q)\Delta(q)B(q,l),$$

i. e., by the insertion of a propagator. This can be repeated as necessary.

Thus, the final addition to the Feynman rules is

- Identify in all the diagrams the connected, amputated 1PI graphs. Calculate these, and the result can be obtained by just multiplying and adding the results together such as to obtain the original graphs

Note that the construction can be extended further, to so-called nPI graphs. However, their recombination is in general no longer possible by multiplications, but usually involves integration over intermediate momenta. This is beyond the scope of this lecture.

From this construction it follows that there are two distinct classes of perturbative calculations. One is the class of so-called tree-level calculations, in which no loops appear. Since graphs without loops are always one-particle reducible, they can always be cut so long as only to consist out of vertex and propagator expressions. On the other hand, this implies that a tree-level calculation can always be written as just a multiplication of propagators and vertices, without any integration. These contributions turn out to be furthermore the classical contribution, i. e., whatever remains when taking the limit of $\hbar \to 0$. Nonetheless, even tree-level calculations, in particular for many external particles, can become very cumbersome, and both a technical as well as a logistical problem.

The second type of diagrams are all graphs with loops. Since they vanish in the classical limit, this implies that these are the quantum, or also radiative, corrections to a process. The integrals make an evaluation much more complicated. Furthermore, the integrals are usually not finite, leading to the necessity of the renormalization process to be discussed in section 8.2.

3.2 Cross sections and the LSZ construction

Matrix elements are not yet observables. Just like in quantum mechanics it is necessary to compute, e. g. transition rates or spatial probability distributions, so it is necessary to pass to observables in quantum field theory. And just as in quantum mechanics many possible observables can be constructed. However, the most common case is that $n > 0$, usually $n = 1$ or $n = 2$, particles are considered, which start at a space-like distance, and the question is posed, what the probability is that they react and form $m > 0$ particles, which are observed again at space-like separations. The usage of space-like separations is here necessary, because otherwise the particles would still interfere with each other.

This can be formally stated as cluster decomposition. The question is, whether

$$\langle O(x)O(y)\rangle = \langle O(x)\rangle\langle O(y)\rangle \tag{3.17}$$

if $|(x-y)^2| \to \infty$ and space-like. It is in general very complicated to proof cluster decomposition for any given (interacting) theory. In almost all cases additional assumptions are required. In the following, it will simply assumed to hold. This is fairly well established, at least by circumstantial evidence, for all conventional quantum field theories[3]. It is certainly what is observed in experiments[4].

The initial particles are called in states and the outgoing particles out states. The probability to observe this is then given by the Born rule,

$$P_{i\to o} = |\langle P_1(x_1)...P_n(x_n)|Q_1(y_1)...Q_m(y_m)\rangle|^2.$$

In general, this is a usual matrix element squared, which could be calculated using, e. g., perturbation theory. There is, however, an issue with the objects P_i and Q_j. As was visible, the objects in quantum field theories are fields. And while the first idea is to just use the elementary fields for those, it is still quantum mechanics. And thus every combination of fields with the same quantum numbers mix. The correct mixture would thus appear to be necessary to be determined from the details of the initial state and the final state, e. g. as coherent states. This appears to make the calculations extremely dependent on the experiment. The decisive insight is that the x_i and the y_i are separately pairwise (formally) infinitely space-like, even though the x_i and y_i are time-like. Using this information, the Lehmann-Symanzik-Zimmermann (LSZ) construction yields that it actually does not matter.

This works as follows. Start with an initial state created by operators Q_i on a common space-like hypersurface. This state is then propagated by a time-development operator S, called the S-matrix. Finally, the Born rule is applied for the probability to measure a final state created by a set of operators P_i, which are again located on common space-like hypersurface. The time-like distance, i. e. the eigenzeit between space-like hypersurfaces, between the initial state and final state is taken to become infinite. Thus, the desired probability is

$$\begin{aligned} P_{i\to f} &= |\langle i|S|f\rangle|^2 \\ |i\rangle &= P_1...P_n|0\rangle \\ |f\rangle &= Q_1...Q_m|0\rangle. \end{aligned} \tag{3.18}$$

This requires two more statements. The first is the ordering of the operators. Because of cluster decomposition (3.17) and that all operators act at space-like distances, the ordering

[3]There are subtleties for gauge theory to be addressed in the quantum field theory II lecture.

[4]Note that in the case of entanglement in non-relativistic quantum mechanics it appears different. However, e. g. in the Bell equation, the particles originate from a common source, to which both are time-like, and thus causally connected.

will not matter, if the operators commute at space-like distances. When moving to the path-integral evaluation of the desired matrix elements, this will not matter. But it will matter when performing canonical quantization in chapter 10. The other is the vacuum state $|0\rangle$. It will be assumed to be unique. In the presence of symmetries, this is no longer true. After all, it would be possible to construct states, which are invariant. And what ensures that such states are not the one's of lowest energy, and thus should be considered the vacuum. This is highly non-trivial, and to some extent on the theory. This will therefore be glimpsed over here, and be taken up again in the quantum field theory II as well as phenomenological lectures.

It is useful to split go to momentum space, and split off the possibility that nothing happens explicitly as

$$\begin{aligned}\langle i|S|f\rangle &= \delta_{if} + i(2\pi)^4\delta(P_i - Q_f)\mathcal{M}_{i\to f} = \delta_{if} + i(2\pi)^4\delta(P_i - Q_f)\langle f|T|i\rangle,\\ P_i &= \sum p_i\\ Q_f &= \sum q_i\end{aligned}$$

where it was assumed that $\langle i|f\rangle$ is normalized to unity. The quantity $\mathcal{M}_{\rangle\to\{}$ is called often the matrix element, though it is so exactly only if the initial state and final state differ. Overall momentum conservation is explicitly taken care of also by pulling out a δ function from the matrix element. In turns, this gives an implicit definition of the T-matrix. If $n = 1$ and $m > 1$, this is called a decay. If $n \geq 2$, this is called scattering, except when $m = 1$, when it is called fusion. If $\langle i|f\rangle \neq 0$, this is called elastic scattering, otherwise it is called inelastic scattering.

Especially, this allows to determine inelastic scattering cross sections and decay widths. In four dimensions, for a $1 \to m$ particle decay, this yields the differential decay width as

$$d\Gamma = \frac{(2\pi)^4}{2E}|\mathcal{M}_{i\to f}(p,\{q_i\})|^2\delta\left(p - \sum_i q_i\right)\Pi_i\frac{d^{d-1}q_i}{(2\pi)^3E_i},$$

i. e. the number of particles emitted with a specified momentum and direction. The partial decay width is then obtained by integrating over the possible phase space, and the total width by summing all partial decay widths. Likewise, the most common case of $2 \to m$ scattering, given some influx of particles, yields the differential cross section width in four dimensions as

$$d\sigma = \frac{(2\pi)^3}{4\sqrt{(p_1p_2)^2 - m_1^2m_2^2}}|\mathcal{M}_{i\to f}(p_1,p_2,\{q_i\})|^2\delta\left(p_1 + p_2 - \sum_i q_i\right)Pi_i\frac{d^{d-1}q_i}{(2\pi)^3E_i}.$$

Integrating over the final-state momenta, the so-called phase space, yields the inclusive cross-sections. Summing afterwards over all possible final states the total cross section or exclusive cross section. Any partial summation is called a partially inclusive cross section.

Particularly useful special cases are the center-of-mass expressions for a two-particle final state, i. e. $m = 2$, which implies $\vec{q}_1 = -\vec{q}_2 = \vec{q}$ and thus integration about one solid angle can be performed,

$$\frac{d\Gamma}{d\Omega} = \frac{1}{32\pi^2 m}\frac{|\vec{q}|}{m}|\mathcal{M}_{i\to f}(p,q)|^2\theta(m-m_1-m_2) \tag{3.19}$$

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s}\frac{|\vec{q}|}{|\vec{p}|}|\mathcal{M}_{i\to f}(p_1,p_2,q)|^2, \tag{3.20}$$

where $s = E^2$ is the center-of-mass energy. Note that only the integrated decay width and cross section are frame-independent.

The aim is thus to calculate T. While the previous procedure of section 3.1 allows to perform the calculation for a fixed set of operators P_i and Q_i, as noticed, this can be simplified. For this, it is first to understand more general propagators.

A propagator is determined by the proper two-point function, created from two fields and fixed quantum numbers,

$$\langle O_1(p)O_2(-p)\rangle = R(p^2) \tag{3.21}$$

where the structure of the result comes from Poincaré symmetry alone. Furthermore, for any kind of particle interpretation, $m^2 = p^2$ determines the satisfaction of the on-shell condition, i. e. that the four-momentum is indeed associated with the mass. Otherwise, the relativistic energy-momentum relation is violated. This off-shell behavior can occur in quantum physics in intermediate steps, but not for an physical particle. While in general the propagator can be an arbitrary complicated function of p^2, this cannot be the case if the object described is space-like infinitely separated from every other particle. Then, due to (3.17), there can be no interactions, and the described particle is free. Thus, in such an asymptotic case[5]

$$R(p^2)|_{p^2=m^2} = \frac{1}{p^2-m^2} \tag{3.22}$$

needs to hold. Strictly speaking, these equalities only hold for $p^2 = m^2$, in which case this quantity diverges. This is true for any operators with appropriate quantum numbers in (3.21), and thus the choice of operators does not matter. This property is automatically fulfilled by the perturbative propagator, which therefore describes also such an asymptotic state. Hence, within the scope of perturbation theory, elementary fields automatic correspond to a particle concept, a feature known as particle-field duality.

[5]If there exists multiple stable particles with the same conserved quantum numbers, this needs to be replaced with a sum over all these states.

Now

$$\left.\frac{\delta^n Z[j]}{Z[0]\delta j_1...\delta j_n}\right|_{j_i=0} = \left.\frac{\delta^n W[j]}{\delta j_1...\delta j_n}\right|_{j_i=0} = \left.\frac{\Phi_1}{\delta j_1}...\frac{\delta\Phi_n}{\delta j_n}\frac{\delta^n W}{\delta\Phi_1...\delta\Phi_n}\right|_{j_i=0} \tag{3.23}$$

$$= D_1...D_n \left.\frac{\delta^n\left(\int d^dx\Phi(x)j(x)-\Gamma\right)}{\delta\Phi_1...\delta\Phi_n}\right|_{j_i=0} = -D_1...D_n \left.\frac{\delta^n\Gamma}{\delta\Phi_1...\delta\Phi_n}\right|_{\Phi_i=0}, \tag{3.24}$$

where the indices also denote the arguments of the propagators. Thus, any correlation function can be written in terms of a vertex function and a product of propagators. This implies that every correlation function has singularities if the particles corresponding to the external fields are on-shell. These are removed when moving to the vertex function, which are thus the truncated ones. It should also be noted that perturbatively, because of (3.12), these are, in fact, the connected diagrams. Note that these operators do not impose any condition on the space-time positions of the fields, and especially, this results does not make use of the asymptotics.

Now, finally, an interaction can be given in terms of a conditional probability. Start with an initial state, which has space-like separated particles. They propagate to some points in causal contact, interact, and then need to propagate to the final states. The probability factors are given by the asymptotic propagators, the correlation function given the interaction, and then propagation to the asymptotics. In all cases, this requires the connected correlation function, due to the choice of normalization. This yields

$$\begin{aligned}\langle i|S|f\rangle &= (R^i_1...R^i_n)^{-\frac{1}{2}}\langle P_1...P_nQ_1...Q_m\rangle(R^f_1...R^f_n)^{-\frac{1}{2}}\\ &= -(-i)^{n+m}(Z^i_1...Z^i_nZ^f_1Z^f_m)^{-\frac{1}{2}}\left.\frac{\delta^n\Gamma}{\delta\Phi_1...\delta\Phi_n}\right|_{\Phi_i=0}\end{aligned} \tag{3.25}$$

where

$$i\sqrt{Z} = R^{-1}D\big|_{p^2=m^2} \tag{3.26}$$

is the difference on the pole between the interacting propagator and the asymptotic, non-interacting propagator. Thus, it is the change of normalization due to the interaction of the fields.

This implies that it is sufficient to calculate the vertex function to determine general cross sections. Especially, in perturbation theory, these are by constructions just the connected, amputated vertex functions. In perturbation theory this can furthermore be reduced to the case of 1PI functions.

3.3 Sample tree-level calculation

With the tools at hand, the next step is to perform a few examples. These will all be tree-level calculations for the moment. There is one issue to clarify before starting, and that has to do with the expansion parameter. As noted, the expansion is formally in the interaction term. However, in the example cases so far, this was also an expansion in the coupling, and $\hbar$, and the number of loops. Now, the fixed association of $\hbar$ and loops will remain. However, as soon as a theory has more than one coupling, the situation becomes involved.

Strictly speaking, and expansion is determining a (functional) Taylor series in the coupling constant. Just like with an ordinary Taylor series, when there are two variables, there will appear mixed terms. Statements about convergence and analyticity properties will require to take into account all of these mixed terms, and obtain all terms up to order n, where n can be obtained by any combination of the orders of the couplings.

However, in many practical applications, the order does not reflect the quantitatively relevant terms. It is thus often more useful to select those contribution, for which the total power $g_1^{n_1}...g_m^{n_2}$ of the involved couplings are bigger than a fixed value. As increasing the order will reduce this factor, in as far as a Taylor series is justified at all, this gives a good first estimate. However, the actual terms multiplied by the prefactor can themselves can upset this ordering, even if the series would be ultimately convergent. Thus, selecting a particular order is a-priori a good first step. But it by no means guarantees to get a meaningful, quantitatively relevant result in the end. Only be improving the calculations to higher orders and/or compare to experiment or other methods, the quality of the perturbative result can be ascertained.

3.3.1 Tree-level decay

The first example will be the decay of a particle in four dimensions. To this end, of course, there needs to be a kinematically allowed decay channel. To this end, a modification of the O(2) linear σ-model appears useful, in which there is a three-point interaction between two particles, each having a different mass,

$$\mathcal{L} = \sum_i \left(\frac{1}{2}\partial_\mu \phi_i \partial^\mu \phi_i - \frac{m_i^2}{2}\phi^2 - \frac{\lambda_i}{24}\phi_i^4 \right) + \frac{\zeta}{2}\phi_1\phi_1\phi_2$$

Thus, the two particles have different masses, and the last term allows a transition from particle 2 to particle 1, if $m_2 \geq 2m_1$.

At tree-level, the only relevant contribution is the three-point vertex. It's Feynman

rule is given, by virtue of (3.13), as

$$\Gamma^{\phi_1^2\phi_2} \left.\frac{\delta^3 S}{\delta\phi_1\delta\phi_2\delta\phi_1}\right|_{\phi_1=\phi_2=0} = \zeta.$$

Since there are no internal lines, this already gives the 1PI matrix element necessary for the decay width formula (3.19),

$$\begin{aligned}\frac{d\Gamma_{\phi_2\to\phi_1\phi_1}}{d\Omega} &= \frac{1}{32\pi^2 m}\frac{|\vec{q}|}{m}|\mathcal{M}_{\phi_2\to\phi_1\phi_1}|^2\theta(m-m_1-m_2)\\ &= \frac{1}{32\pi^2 m_2}\frac{\sqrt{\frac{m_2^2}{4}-m_1^2}}{m_2}|\zeta|^2\theta(m_2-2m_1)\\ &= \frac{1}{32\pi^2 m_2}\sqrt{\frac{1}{4}-\frac{m_1^2}{m_2^2}}|\zeta|^2\theta(m_2-2m_1)\end{aligned}$$

In the second step, the rest frame of the decaying particle was chosen, to reexpress the three-momenta of the decay products in terms of the masses. This can be obtained as in classical special relativity from four-momentum conservation,

$$p = q_1 + q_2$$

and thus

$$m_2^2 = p^2 = (q_1^2+q_2^2)^2 = q_1^2+q_2^2+2q_1q_2 = 2m_1^2+2(E_1E_2-\vec{q}_1\vec{q}_2).$$

Using that in the rest frame of the decaying particle $\vec{q}_1 = -\vec{q}_2$, and thus $E_1 = E_2$ and the relation

$$m^2 = E^2 - \vec{q}^2$$

yields finally

$$m_2^2 = 2m_1^2 + 2(E^2+\vec{q}^2) = 4m_1^2 + 4\vec{q}^2,$$

which can then be solved for $\vec{q}^2$.

As the result does not depend on the direction, it is straightforward to integrate it to get the partial decay width,

$$\Gamma_{\phi_2\to\phi_1\phi_1} = \frac{1}{8\pi m_2}\sqrt{\frac{1}{4}-\frac{m_1^2}{m_2^2}}|\zeta|^2\theta(m_2-2m_1).$$

As it should, it has units of energy, given that ζ has units of energy. A few general remarks are in order. The first is that the θ-function, which implements four-momentum conservation, also ensures that the result is real. Moreover, the decay will increase quadratically

with the coupling strength, rather than linear as may have been thought because of its linear appearance in the Lagrangian. This is a general feature. Also, the sign of ζ does not matter. That is, however, an artifact of using the lowest order in perturbation theory. The decay width also depends on the kinematics. It increases the lighter the final state particles are, i. e., there is a so-called phase space effect. However, it does not do so arbitrarily, and there is an upper limit. This is also expected. The maximum of a decay is, if the particle decays immediately. But as a quantum effect, even that will not always happen, and thus the life-time, defined as the inverse of the decay width, remains always finite. In fact, the decays will follow an exponential decay law, $\exp(-\Gamma t)$. Thus, there is a maximum rate with which the initial amount of particles can be depleted in a quantum-mechanical system like this. Finally, the decay width drops at fixed decay products when increasing the mass of the initial particle. Thus, heavier particles decay quicker.

These features are relatively characteristic for decays. However, higher orders and special theories can deviate substantially from these general rules. But there are still good qualitative first estimates, if nothing else is known yet.

3.3.2 Tree-level scattering

For an example of scattering in four dimensions, the one-particle linear σ model, (2.2), will be used. To avoid the problems appearing with loops for now, this will be done to order (ζ^2, λ^1), which could be expected to be a suitable approximation if $\zeta/m \gg \lambda$, where m is needed to create a dimensionless quantity. The simplest process would be a $2 \to 2$ scattering. However, here enters already a quantum-mechanical problem: The two particles are the same, and thus indistinguishable. This will be needed to be taken into account.

The two relevant Feynman rules are the propagator and the the three-point vertex, which is again just ζ, and the four-point vertex, λ. At this order, there are four distinct diagrams. The initial four-momenta will be labeled p_1 and p_2, and the final four-momenta q_1 and q_2, and

$$p_1 + p_2 = q_1 + q_2$$

due to four-momentum conservation. It is convenient to choose the center-of-mass frame in the following, implying

$$\vec{p_1} + \vec{p_2} = \vec{q_1} + \vec{q_2} = \vec{0},$$

which will be used in the following.

The simplest diagram is the one including the four-point vertex, as this already ap-

peared in section 3.1. Here, the 1PI matrix element is just given by

$$\mathcal{M}^{\lambda}_{\phi\phi\to\phi\phi} = \lambda. \tag{3.27}$$

In the other cases always an internal propagator appears. However, there will be no loops attached to it, as the propagator always emerges from a three-point vertex. Such diagrams are hence of order $\zeta^2\lambda$, and thus do not contribute at this order. It remains a tree-level process.

The first diagram is obtained from the two initial particles annihilating each other into a single particle, which then propagates, and finally decays into the two final state particles. There is only one way to draw this diagram. Its matrix element is

$$\mathcal{M}^{s}_{\phi\phi\to\phi\phi} = \frac{i\zeta^2}{k^2 - m^2} \tag{3.28}$$

with the internal four-momentum k being

$$k = -p_1 - p_2 = -q_1 - q_2$$

due to 4-momentum conservation, and the convention to have all momenta going into the vertex, and thus $k + p_1 + p_2 = k + q_1 + q_2 = 0$.

In this situation, it is useful to introduce the so-called Mandelstam variables, which are relativistic invariants,

$$\begin{aligned}
s &= (p_1 + p_2)^2 = (q_1 + q_2)^2 = 2m_i^2 + 2\vec{p}_1\vec{p}_2 = 2m_f^2 + 2\vec{q}_1\vec{q}_2 = 4E^2 \\
t &= (p_1 - q_1)^2 = (p_2 - q_2)^2 = m_i^2 + m_f^2 - 2E^2 + 2|\vec{p}_1||\vec{q}_1|\cos\theta \\
&= m_i^2 + m_f^2 - 2E^2 + 2|\vec{p}_2||\vec{q}_2|\cos\theta \\
u &= (p_1 - q_2)^2 = (p_2 - q_1)^2 = m_i^2 + m_f^2 - 2E^2 - 2|\vec{p}_1||\vec{q}_2|\cos\theta \\
&= m_i^2 + m_f^2 - 2E^2 - 2|\vec{p}_2||\vec{q}_1|\cos\theta \\
s + t + u &= 2m_i^2 + 2m_f^2
\end{aligned}$$

where the masses of initial state particles and final state particles are separated as m_i and m_f for later use, but here $m_i = m_f = m$. The angle θ is the scattering angle between the incoming momenta and outgoing momenta. While their definitions appears relatively special, it turns out that these combinations appear very often in practical calculations. Especially, since (3.28) can be written as

$$\mathcal{M}^{s}_{\phi\phi\to\phi\phi} = \frac{i\zeta^2}{s - m^2}$$

it has become practice to call quantities after their dependency on Mandelstam variables. Thus, this matrix element is called an s-channel process, as at tree-level the process only

depends on s. It describes a process, were the particles are annihilated into a single one, and then split again into two. This already implies that this matrix element only depends on the center-of-mass energy, but not on the scattering angle. Moreover, (3.27) does not even depend on any Mandelstam variables, and thus neither on the energy of the process nor on the scattering angle. It remains to determine any possible symmetry factors. None appear in the s-channel, as there is only one way of drawing the diagram.

The remaining diagram corresponds to an exchange of the particle. At first it appears as if there would be also a symmetry factor. But here the different connections create a situation with exchanged external momenta. This is not the same diagram,. but rather a so-called exchange diagram, originating from the fact that particles cannot be distinguished. Its matrix element is

$$\mathcal{M}^s_{\phi\phi\to\phi\phi} = \frac{i\zeta^2}{t-m^2} + \frac{i\zeta^2}{u-m^2},$$

and is thus called t-channel, and the last term stems from the exchange diagram.

The total expression for the matrix element is thus

$$\mathcal{M}_{\phi\phi\to\phi\phi}\lambda + i\zeta^2\left(\frac{1}{t-m^2}+\frac{1}{u-m^2}+\frac{1}{s-m^2}\right).$$

This yields the differential cross section in the center-of-mass frame

$$\begin{aligned}
\frac{d\sigma}{d\Omega} &= \frac{1}{64\pi^2 s}\left(\lambda^2 + \zeta^2\frac{(5m^4+(s+t)^2-st-4m^2(s+t))^2}{(m^2-s)^2(m^2-t)^2(s+t-3m^2)}\right)\\
&= \frac{1}{64\pi^2 s}\left(\lambda^2 + \zeta^4\frac{(24m^2(m^2-s)+7s^2+(s-4m^2)^2\cos(2\theta))^2}{(m^2-s)^2(8m^4-s^2+2(s^2-4m^2)^2\cos(2\theta))^2}\right)\\
&\overset{m\to 0}{=} \frac{1}{64\pi^2 s}\left(\lambda^2 + \zeta^4\frac{(7+\cos(2\theta))^2}{4s^2\sin^4(\theta)}\right).
\end{aligned}$$

There are a few general remarks in order. First, if $\zeta = 0$, the result does not depends on the scattering angle, but is isotropic. This is changed by the ζ-dependent part, which shows a characteristic 2θ-dependence. Such a dependence can be used to characterize the particles involved in the scattering. Second, the interference term between the genuine three-particle interaction and the exchange and annihilation diagram. That is an artifact of the three-level situation, and the relative factor of i between both diagrams due to the appearing propagators. This would change beyond tree-level. Third, in the limit of $m \to 0$, or equivalently $s \to \infty$, there appears the typical Rutherford singularity of $\sin^{-4}\theta$, which is, however, tamed when there is a finite mass. Fourth, there is a singularity at $s = m^2$, due to the exchanged particle. This is a very typical situation, and one of the possibilities to look for, then different, particles in the collision of particles. If the particle would be

unstable, this singularity becomes tamed by the width. Fifth, the total cross section is diverging for $m \to 0$. Just like in the Rutherford case, this is because a genuine bare point-particle is assumed. Effects at NLO and/or any internal structure of the involved particles will tame this singularity. Especially, if the particles are massive, the total cross section becomes

$$\begin{aligned}\sigma &= \frac{1}{64\pi s}\left(\lambda + \frac{2\zeta^4}{m^2(3m^2-s)\sqrt{(s-3m^2)(ms-2m^3)}}\times\right.\\ &\quad\times\left(16m^8 - 28m^6 s + s^4 - m^2 s\left(6s^2 + m\sqrt{(s-3m^2)(s-2m^2)^2}\right.\right.\\ &\quad\left.\left.+m^4\left(17s^2 + 3m\sqrt{(s-3m^2)(s-2m^2)}\right)\right)\right).\end{aligned}$$

Sixth, and finally, both the total cross section and the partial cross section decay like $1/s$. Thus, the probability for elastic scattering decreases with increasing momenta. This is actually generic, as will be discussed in section 8.5,

To obtain an inelastic process, it would be necessary to have either more particles in the final state or to introduce the possibility to produce different particles in the final state. In such cases, it may happen that an inelastic scattering process can only take place if the initial state particles have enough energy. This is a threshold, and will be explored in more detail in section 9.4.

Chapter 4

Particles and the Poincare group

So far, only scalar particles have been considered. However, already in non-relativistic quantum physics other kinds of particles have been encountered, most notably electrons. They have an additional feature, called spin, which, by-and-large, works similarly as an angular momentum, and thus rotation symmetry. The concept of rotation symmetry is here, of course, replaced by Lorentz transformations. It will be seen that this has further consequences.

4.1 Representations of the Poincare group and spin

Since Poincare symmetry, consisting out of translation symmetry and Lorentz symmetry, is a symmetry, it implies that fields can carry representations of the symmetry. This is very alike the case of internal symmetries in section 2.5. So far, the scalar field had been in the trivial representation. Thus, the natural question is, whether there are fields ψ, which transform non-trivially under the Poincare group, i. e.

$$\psi' = \Lambda_{ij}\psi_j \tag{4.1}$$

where Λ is a linear operator implementing a Lorentz transformation.

The Poincare group is a semi-direct product of the Abelian non-compact translation group with generators P_μ, and the pseudogroup SO(1,3). There are also time reversal and parity, which yield then O(1,3). These will be addressed in section 6 It is convenient to decompose in the following the generators of spatial rotations J and boosts K as

$$\begin{aligned} A &= \frac{1}{2}(J + iK) \\ B &= \frac{1}{2}(J - iK). \end{aligned}$$

The Lorentz algebra becomes then a semi-direct product of two SU(2) algebras

$$\begin{aligned}[A_i, A_j] &= \epsilon_{ijk} A_k \\ [B_i, B_j] &= \epsilon_{ijk} B_k \\ [A_i, B_j] &= 0. \end{aligned} \tag{4.2}$$

In dimensions different from four, the situation changes. Especially, the group SO(1,N) decomposes into the N-dimensional rotation group SO(N), but the decomposition in SU(2) groups is in general not possible. Especially, in 1+1 dimensions SO(1,1)$\sim$U(1) becomes Abelian[1]. However, the group SO(1,N)/SO(N) will always create a semi-direct product of the SO(N). Especially, in three dimensions, SO(1,2) is itself only an SU(1,1) group, with one rotations and two boosts. In dimensions larger than four there are more independent rotations than boosts.

To characterize a representation requires as usual invariant Casimirs. For the involved group, there is one Casimir each.

For the Abelian translation group, this Casimir is obtained as

$$m^2 = P_\mu P^\mu,$$

which is invariant under the Lorentz group. Suggestively, it is already called m, as it will be later on be used to identify the mass. Note, however, that m can have either sign or be zero. It is purely empirical, which of those Casimirs are encountered in nature, and it turns out to be only objects having $m \geq 0$. They will be therefore concentrated on in the following. As an Abelian group, all irreducible representation are one-dimensional. The parameter m classifies thus the little group of the orbits, where $m > 0$, $m = 0$, and $m < 0$ are the strata for the little groups SO(3), SO(2), and SO(1,2). The second already indicates that particles with $m = 0$ cannot have a rest frame.

Both SU(2) groups can be characterized by an independent half-integer or integer quantum number j. Since the group is non-Abelian, its representations are not necessarily one-dimensional, but rather $2j + 1$ dimensional. In fact, this is the spin of the particle. However, in quantum mechanics spin behaved like an angular momentum. But angular momentum is no longer well-defined in a relativistic setting, as it can be changed by a boost, due to the semi-direct nature. Thus, a generalization is necessary.

Thus, a Poincare-invariant operator is needed to represent spin. It is obtained using the Pauli-Lubanski vector

$$W_\mu = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} P^\nu M^{\rho\sigma}. \tag{4.3}$$

[1]Note that this implies that the Casimir will be continous. Given the following, this implies that spin in 1+1 dimensions is not quantized, but continous. To distinguish such particles, they are called anyons.

The antisymmetric operator M is defined as

$$\begin{aligned} M^{0i} = -M^{i0} &= K_i \\ M^{ij} &= \epsilon_{ijk} J_k, \end{aligned}$$

and is therefore the generator of a Lorentz transformation with anti-symmetric parameter matrix ω, as $\omega_{\alpha\beta} M^{\alpha\beta}$.

The Pauli-Lubanski vector is, due to the Levi-Civita tensor, orthogonal to the momentum vector, and thus linearly independent. To show that its square W^2 is actually a Casimir operator requires to show that it commutes with both the momentum operator and the generator of rotation. Since

$$[P_\lambda, W_\mu] = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} P_\nu [P_\lambda, M^{\rho\sigma}] = \frac{i}{2}\epsilon_{\mu\nu\rho\sigma} P_\nu (\delta^\lambda_\rho P^\sigma - \delta^\sigma_\lambda P^\rho) = 0,$$

already the momentum operator and the Pauli-Lubanski vector commute, so does the square of any of the two vectors with the other vector. Since W^2 is a scalar, it also commutes with the generator of rotations, which can be shown explicitly in the same way. Together with the linear independence, this is sufficient that W^2 is an independent Casimir.

To show that this vector is indeed associated with the usual spin consider its commutator

$$[W_\mu, W_\nu] = i\epsilon_{\mu\nu\rho\sigma} P^\rho W^\sigma,$$

which can again be obtained by direct evaluation. In the rest frame only those commutators remain with $\mu \neq 0$ and $\nu \neq 0$, yielding

$$[W_i, W_j] = im\epsilon_{ijk} W_k \tag{4.4}$$

which is, up to a normalization, just the spin algebra. This especially implies that its eigenvalues behave, up to a factor of m, like the ones of a spin, and indeed the eigenvalues of W^2 are thus spin eigenvalues. A more explicit proof can be obtained by using the explicit form of the Pauli-Lubanski vector (4.3), if desired. This immediately raises the question of what happens for $m = 0$. However, in this case, no rest frame exists, and thus the full expression is necessary. This will be treated in more detail later. Note that (orbital) angular momentum is defined, like in quantum mechanics, based on the relative momenta and positions, but transforms like an integer representation.

Thus, any field can be represented by a triple of quantum numbers (m^2, j_1, j_2), and will form a $2(j_1 + j_2) + 2$-dimensional multiplet. But because the group is non-compact, finite-dimensional representations will not be unitary, and thus need to be functions. This

is realized by choosing the objects to be fields, functions of space-time. Hence, a field is full characterized as[2] $\psi^{j_1 j_2 m^2}_{m_1 m_2}$. The scalar field considered so far had representation $(m^2 > 0, 0, 0)$.

Thus, a Lorentz transformation Λ will act on a field $\psi^{j_1 j_2 m^2}_{m_1 m_2}$ as

$$\Lambda \psi^{j_1 j_2 m^2}_{m_1 m_2} = \lambda^{n_1}_{m_1} \lambda^{n_2}_{m_2} \psi^{j_1 j_2 m^2}_{n_1 n_2}(\lambda p), \tag{4.5}$$

where the λ are the corresponding non-unitary representations, and the action on the momentum p is the conventional one of special relativity. The transformation in momentum space operates likewise. In this case, $\Lambda^\dagger = \Lambda^{-1}$. Correspondingly, m is the eigenvalue of the (lowest-order) Casimir of this representation with respect to the orbits, while the spin is the Casimir for the non-unitary finite-dimensional representations.

It is now required to say a few words about the $m = 0$ case for any $j_i \neq 0$. While for $m > 0$ the fields can be considered in a rest frame, and thus by virtue of (4.4) can be associated with a rest frame, this is no longer possible for the massless case. Thus, in the massive case the momentum vector becomes $p = (m, \vec{0})$, and thus the little group is SO(3), as noted above. hence, they carry a representation of the little group, and thus of SO(3). If $m = 0$, there is no rest frame, and thus at most the momentum vector can be $p = (p, 0, 0, p)$, which has little group SO(2)$\sim$U(1). The projection of the Pauli-Lubanski vector on the momentum vanishes

$$P_\mu W^\mu = \frac{1}{2} P^\mu P^\nu \epsilon_{\mu\nu\rho\sigma} M^{\rho\sigma} = 0,$$

and hence W_μ can only be light-like parallel or anti-parallel to the momentum vector, as $W_0 = P_3 \epsilon_{30jk} M^{jk} \neq 0$ and $W_i = 0$ for $i = 1$ or $i = 2$. Thus, for a zero-mass particle with non-zero spin, the Pauli-Lubanski vector can only be projected in either direction. Normalize to the momentum, it can thus have only $W_0 = \pm c$, rather than the full range of m_i values. This value is called the helicity, and takes the role of spin for massless particles.

It thus remains to understand the features of the representation. Here, the cases with either $j_i = 1/2$ will be considered. Cases with higher spins will follow in sections 7.1 and 7.7.

[2]Sometimes, the two SU(2) groups are distinguished by giving one of them an overdot. This notation of dotted and undotted indices will not be used here, as will be for the most of this lecture not be important to distinguish both.

4.2 Weyl spinors

In the following, the case of four dimensions will be considered only. However, different dimensionalities can be constructed analogously[3]. Consider the case with either $j_i = 1/2$ and the other zero. In that case, the representation in the rest frame needs to be a spinor representation with $m_i = \pm 1/2$. It should be noted that these representations are complex, and thus the fields are necessarily complex. Thus, the field needs to transform in such a spinor representation, and a consequently called (Weyl) spinors. It is worthwhile to construct their transformation properties explicitly.

An infinitesimal Lorentz transformation acts like

$$\begin{aligned} p_0 = E \to E' &= E - \eta_i p_i \\ \vec{p} \to \vec{p}' &= \vec{p} - \vec{\epsilon} \times \vec{p} - \vec{\eta} E, \end{aligned}$$

with parameter vectors η for boosts and ϵ for rotations. For the two-dimensional spinor representation spinors(=vectors in the representation space) ψ and χ with $j_1 = 1/2$ or $j_2 = 1/2$, respectively, this implies

$$\begin{aligned} \psi \to \psi' &= \psi + (i\epsilon_i \sigma_i/2 - \eta_i \sigma_i/2)\psi \\ \chi \to \chi' &= \chi + (i\epsilon_i \sigma_i/2 + \eta_i \sigma_i/2)\chi \end{aligned}$$

where the σ_i are the Pauli matrices. They differ only by their reaction to boosts, but are affected by rotation in the same way. It is also visible that only for $|\vec{\epsilon}| = 4\pi$ at $\vec{\eta} = \vec{0}$. Hence, only under spatial rotations by 4π spinors are transformed back, while vectors do so after 2π. While this seems counter-intuitive at first, it should be remember that the intuition comes from our everyday experience being completely determined by objects with integer spin in an (essential) Galilean way. Minkowski space-time is very different, as rotations also change time, and thus 'longer' rotation is needed to be back for spinors.

Defining the non-unitary matrix

$$V = 1 + i\epsilon_i \sigma_i/2 - \eta_i \sigma_i/2$$

the transformation rules simplify to

$$\begin{aligned} \psi' &= V\psi \\ \chi' &= V^{-1\dagger}\chi = (1 + i\epsilon_i \sigma_i/2 - \eta_i \sigma_i/2)^{-1\dagger}\chi \\ &= (1 - i\epsilon_i \sigma_i/2 + \eta_i \sigma_i/2)^{\dagger}\chi = (1 + i\epsilon_i \sigma_i/2 + \eta_i \sigma_i/2)\chi. \end{aligned}$$

[3]It should be noted that four dimensions are special, and their are a few more possibilities in other dimensions, triggered by the non-equivalent spinor representation of even-dimensional and odd-dimensional SO(N) groups. This will not be needed here, but can play a role in theories with extended space-time structure.

Furthermore, because

$$(E - \sigma_i p_i)\psi = \sigma^\mu p_\mu \psi \sim \chi \tag{4.6}$$
$$(E + \sigma_i p_i)\chi = \bar{\sigma}^\mu p_\mu \chi \sim \psi, \tag{4.7}$$

where $\bar{\sigma}_\mu^T = (1, \vec{\sigma})$, and

$$V \sigma^\mu p_\mu = \bar{\sigma^\mu} p_\mu V^{-1\dagger}$$

by explicit calculation, it follows that a multiplication with $\sigma^\mu p_\mu$ exchanges a ψ-type-spinor and a χ-type-spinor, and thus changes the respective transformation properties under Lorentz transformations. The non-unitary operator $\sigma_\mu p^\mu$ thus maps between the two SU(2) subgroups of SO(1,3).

To determine the constant of proportionality, note that p^2 can be rewritten as

$$p^2 \delta_{kl} = p_\mu g^{\mu\nu} \delta_{kl} p_\nu = p_\mu \sigma^\mu_{ki} \bar{\sigma}^\nu_{il} p_\nu = \left(\sigma^\mu p_\mu \bar{\sigma}^\nu p_\mu\right)_{kl}.$$

Since on-shell $p^2 = m^2$ also holds, this implies

$$m^2 \psi_k = p^2 \delta_{kl} \psi_l = \left(\sigma^\mu p_\mu \bar{\sigma}^\nu p_\mu\right)_{kl} \psi_l,$$

and likewise for χ. This implies that every component of ψ and χ fulfill the Klein-Gordan equation using the Klein-Gordan operator (2.17),

$$\Delta \psi_i = (\partial_\mu^2 - m^2)\psi_i = 0.$$

This fixes the constant of proportionality in (4.6-4.7) to be m, up to a sign. Especially, for massless fermions, the right-hand side of (4.6-4.7) is zero. This can be recast into the statement that both states are eigenstates of a helicity operator $\sigma_i \vec{p}_i / \sqrt{\vec{p}_i \vec{p}_i}$, but with opposite values. Thus, such a Weyl fermion has definite helicity of ± 1, and therefore in the massless case the spin has a fixed projection onto the momentum. Because of this, the Weyl fermions are called chiral or handed, and the two representations $(1/2, 0)$ and $(0, 1/2)$ are called right-handed and left-handed. No matter the mass, the two components of the Weyl spinors always correspond to the two magnetic quantum numbers $m = \pm 1/2$.

To obtain a theory of Weyl fermions will require to build scalars from them. For that note

$$\sigma_2 \psi^{*'} = \sigma_2 V^* \psi^* = \sigma_2 (1 - i\epsilon_i \sigma^{i*}/2 - \eta_i \sigma^{i*}/2)\psi^* = (1 + i\epsilon_i \sigma^i/2 + \eta_i \sigma^i/2)\sigma_2 \chi = V^{-1\dagger} \sigma_2 \psi^*, \tag{4.8}$$

where in the last step it was used that σ_2 anti-commutes with the real σ_1 and σ_3 matrices, but commutes with itself and is itself purely imaginary. Hence multiplying the complex-conjugate by σ_2, also maps between the representations, but this time without involving the momentum like in (4.6-4.7). Hence, e. g., the quantity

$$(-i\sigma_2\psi^*)^\dagger\psi = (-i\sigma_2\psi)^T\psi = \psi^T(i\sigma_2)\psi$$

is a scalar, and likewise for the other handedness.

4.3 Dirac spinors

While theories of single Weyl fermions are possible, and indeed appear in nature in the weak interactions, it is also very common that a theory contains equal numbers of left-handed Weyl spinor and right-handed Weyl spinors with equal masses. The electromagnetic interaction and the strong interaction are of this type. In such a situation it becomes convenient to assemble a new object, a so-called Dirac spinor.

By and large, a Dirac spinor is a four-dimensional reducible representation of both Weyl representations. This can be made explicit by assembling a four-dimensional spinor

$$\Psi = \begin{pmatrix} \psi \\ \chi \end{pmatrix}. \tag{4.9}$$

If the masses of both Weyl spinors is the same, (4.6-4.7), together with the Klein-Gordan equation, imply that this Dirac spinor fulfills the Dirac equation

$$\left(i\gamma^\mu\partial_\mu + m\right)\Psi = 0. \tag{4.10}$$

The four γ_μ matrices are given by

$$\gamma_\mu = \begin{pmatrix} 0 & \bar{\sigma}_\mu \\ \sigma_\mu & 0 \end{pmatrix}.$$

They fulfill the Clifford algebra

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu}. \tag{4.11}$$

Note that this is not a Lie algebra. Still, from this algebra many technically useful relations for expressions of γ matrices can be derived, which will be done on a case-by-case basis. As the (4.10) is invariant under unitary transformation, any similarity transformation of the γ_μ works equally well.

To obtain the Weyl spinors again from the Dirac spinor, it is convenient to define the matrix

$$\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.12}$$

This matrix can be used to project out the Weyl spinors as

$$\frac{1}{2}(1+\gamma_5)\Psi = \begin{pmatrix} \psi \\ 0 \end{pmatrix} = \Psi_R \tag{4.13}$$

$$\frac{1}{2}(1-\gamma_5)\Psi = \begin{pmatrix} 0 \\ \chi \end{pmatrix} = \Psi_L. \tag{4.14}$$

It should be noted that this implies that Ψ is an eigenstate of γ_5 with eigenvalue 1.

The Dirac spinor inherits its transformation properties under Lorentz transformations from the Weyl spinors as

$$\Psi' = \Psi + \frac{i}{2}\begin{pmatrix} \epsilon_i\sigma_i - \eta_i\sigma_i & 0 \\ 0 & \epsilon_i\sigma_i + \eta_i\sigma_i \end{pmatrix}\Psi.$$

Because of that a scalar can be constructed more straightforwardly from a Dirac spinor as

$$\Psi^\dagger\gamma_0\Psi = \begin{pmatrix} \psi^\dagger & \chi^\dagger \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \psi \\ \chi \end{pmatrix} = \psi^\dagger\chi + \chi^\dagger\psi.$$

Because this makes the combination $\Psi^\dagger\gamma_0$, and frequent, the definition

$$\overline{\Psi} = \Psi^\dagger\gamma_0$$

is convenient.

Chapter 5

Fermions and the path integral

With the spinor fields available, it seems to be straightforward to construct to a theory from them. However, one more feature is needed. As of so far, an expression like $\psi(x)^T\psi(x)$ may not be Lorentz invariant, but it can be well non-zero. However, fermions are expected to fulfill the Pauli principle, and thus such an expression should vanish. Indeed, as will be seen in section 6.3, this will even be necessary. But so far, there is no possibility to realize this in the context of the classical fields in the path integral formalism. To do so requires to extend the concepts of fields beyond complex numbers.

5.1 Grassmann mathematics

The tool to do so is Grassmann mathematics, which allows for anti-commuting structures, a necessity to realize the Pauli principle.

5.1.1 Grassmann variables

The starting point is to define anti-commuting numbers, α^a, by the property

$$\{\alpha^a, \alpha^b\} = 0 \tag{5.1}$$

where the indices a and b serve to distinguish the numbers. In particular, all these number are nilpotent,

$$(\alpha^a)^2 = 0.$$

Hence, the set $\mathcal{S}$ of independent, so-called, Grassmann numbers with $a = 1, ..., N$ base numbers are

$$\mathcal{S} = \{1, \alpha^a, \alpha^{a_1}\alpha^{a_2}, ..., \alpha^{a_1} \times ... \times \alpha^{a_N}\},$$

where all α_i are different. This set contains therefore only 2^N elements. Of course, each element of $\mathcal{S}$ can be multiplied by ordinary complex numbers c, and can be added. This is very much like the case of ordinary complex numbers or matrices. Such combinations z take the general form

$$z = c_0 + c_a\alpha^a + \frac{1}{2!}c_{ab}\alpha^a\alpha^b + ... + \frac{1}{N!}c_{a_1...a_N}\alpha^{a_1} \times ... \times \alpha^{a_N}. \tag{5.2}$$

Here, the factorials have been included for later simplicity, and the coefficient matrices can be taken to be antisymmetric in all indices, as the product of α^as are antisymmetric. For $N = 2$ the most general Grassmann number is therefore

$$z = c_0 + c_1\alpha^1 + c_2\alpha^2 + c_{12}\alpha^1\alpha^2,$$

where the antisymmetry has already been used. It is also common to split such numbers in their (Grassmann-)odd and (Grassmann-)even part. Since any product of an even number of Grassmann numbers commutes with other Grassmann numbers, this association is adequate. Note that there is no possibility to invert a Grassmann number, but products of an even number of Grassmann numbers are ordinary numbers and can therefore be inverted.

The conjugate of a product of complex Grassmann-numbers, with independent real and imaginary part, is defined as

$$(\alpha^a...\alpha^b)^* = (\alpha^b)^*...(\alpha^a)^* \tag{5.3}$$

Note that the Grassmann algebra (5.1) is different from the Clifford algebra (4.11). They connect to two different aspects.

5.1.2 Grassman analysis

To do analysis, it is necessary to define functions of Grassmann numbers. First, start with holomorphic functions. This is rather simple, due to the nilpotency of Grassmann numbers. Hence, for a function of one Grassmann variable

$$z = b + f$$

only, with b even and f odd, the most general function is

$$F(z) = F(b) + \frac{dF(b)}{db}f.$$

Any higher term in the Taylor series will vanish, since $f^2 = 0$. Since Grassmann numbers have no inverse, all Laurent series in f are equivalent to a Taylor series. For a function of two variables, it is

$$F(z_1, z_2) = f(b_1, b_2) + \frac{\partial F(b_1, b_2)}{\partial b_1} f_1 + \frac{\partial F(b_1, b_2)}{\partial b_2} f_2 + \frac{1}{2} \frac{\partial^2 F(b_1, b_2)}{\partial b_1 \partial b_2} f_1 f_2.$$

There are no other terms, as any other term would have at least a square of the Grassmann variables, which therefore vanishes. Note that the last term is not zero because $F(b_1, b_2) \neq F(b_2, b_1)$ in general, but even if this is the case, it is not a summation.

This can be extended to more general functions, which are no longer analytical in their arguments,

$$F(b, f) = F_0(b) + F_1(b) f \tag{5.4}$$

and correspondingly of more variables

$$F(b_1, b_2, f_1, f_2) = F_0(b_1, b_2) + F_i(b_1, b_2) f_i + F_{12}(b_1, b_2) f_1 f_2.$$

The next step is to differentiate such functions. Note that the function F_{12} has no definite symmetry under the exchange of the indices, though by using an antisymmetric generalization this term can be again written as $F_{ij} f_i f_j$ if F_{ij} is anti-symmetric.

Differentiating with respect to the ordinary variables occurs as with ordinary functions. For the differentiation with respect to Grassmann numbers, it is necessary to define a new differential operator by its action on Grassmann variables. As these can appear at most linear, it is sufficient to define

$$\begin{aligned} \frac{\partial}{\partial f_i} 1 &= 0 \\ \frac{\partial}{\partial f_i} f_j &= \delta_{ij} \end{aligned} \tag{5.5}$$

Since the result should be the same when $f_1 f_2$ is differentiated with respect to f_1 irrespective of whether f_1 and f_2 are exchanged before derivation or not, it is necessary to declare that the derivative anticommutes with Grassmann numbers:

$$\frac{\partial}{\partial f_1} f_2 f_1 = -f_2 \frac{\partial}{\partial f_1} f_1 = -f_2 = \frac{\partial}{\partial f_1}(-f_1 f_2) = \frac{\partial}{\partial f_1} f_2 f_1.$$

Alternatively, it is possible to introduce left and right derivatives. This will not be done here. As a consequence, the generalized product (or Leibnitz) rule reads

$$\frac{\partial}{\partial f_i}(f_j f_k) = \left(\frac{\partial}{\partial f_i} f_j \right) f_k - f_j \frac{\partial}{\partial f_i} f_k.$$

In this context, it is useful to define the Grassmann parity π of a quantity to be zero if the number is Grassmann-even and one if it is Grassmann-odd. Then the Leibnitz rule can be written as

$$\partial_x(ab) = (\partial_x a)b + (-1)^{\pi(a)\pi(\partial_x)} a\partial_x b \tag{5.6}$$

as this will always generate the correct relative sign.

Likewise, the integration needs to be constructed differently. In fact, it is not possible to define integration (and also differentiation) as a limiting process, since it is not possible to divide by infinitesimal Grassmann numbers. Hence it is necessary to define integration. As a motivation for how to define integration the requirement of translational invariance is often used. This requires

$$\begin{aligned} \int df &= 0 \\ \int df\, f &= 1 \end{aligned} \tag{5.7}$$

Translational invariance follows then immediately as

$$\int df_1 F(b, f_1 + f_2) = \int df_1 (h(b) + g(b)(f_1 + f_2)) = \int df_1 (h(b) + g(b) f_1) = \int df_1 F(b, f_1)$$

where the second definition of (5.7) has been used. Note that also the differential anti-commutes with Grassmann numbers. Hence, this integration definition applies for $df\, f$. If there is another reordering of Grassmann variables, it has to be brought into this order. In fact, performing the remainder of the integral using (5.7) yields $g(b)$. It is an interesting consequence that integration and differentiation thus are the same operations for Grassmann variables, as can be seen from the comparison of (5.5) and (5.7).

To describe fermionic matter requires then to replace all fields describing fermions by fields of Grassmann variables. I. e., a Dirac fermion field is described by a spinor $\psi(x)$ with components $\psi_a(x)$ being Grassmann-odd functions of x. This implies that $\psi_a(x)\psi_a(x) = 0$ (where summation may or may not be implied), as such functions at the same space-time points are nilpotent. This implements the Pauli principle, as thus every time two Grassmann quantities coincide at the same space-time point the result automatically vanishes.

If in this way in the Lagrangians the fermion fields are replaced by Grassmann-odd functions, the use of the rules for integration and differentiation can be extended in a straightforward way to the path integral. Especially, the most important relation necessary later on is again the Gaussian integral over Grassmann fields. To illustrate the use of Grassmann functions, this will be calculated in detail. The starting point is the integral

$$\int d\alpha^* d\alpha \exp(\alpha^* A \alpha),$$

with some ordinary number A. The Taylor expansion of this expression is

$$\int d\alpha^* d\alpha \exp(\alpha^* A \alpha) = \int d\alpha^* d\alpha \alpha^* A \alpha,$$

and any terms linear or constant in the Grassmann variables will vanish during the integration, and likewise all higher-order terms will be zero, since $\alpha^2 = \alpha^{*2} = 0$. In the next step, it is necessary to be very careful in the ordering of the integrals, as also the differentials anti-commute with the variables. To act with $d\alpha$ on the variable α requires to anti-commute it with α^* and $d\alpha^*$ first, giving a factor of $(-1)^2$,

$$\int d\alpha d\alpha^* \alpha^* A \alpha = -\int d\alpha^* d\alpha \alpha^* A \alpha = \int d\alpha^* \alpha^* A d\alpha \alpha = \int d\alpha^* \alpha^* A = A \qquad (5.8)$$

which is remarkably different from the normal Gaussian integral (2.11), as it returns A instead of $A^{-1/2}$.

It can be likewise shown, that the generalization to many variables yields $\det A$ instead of $(\det A)^{-1/2}$. Similarly, it can be shown that for the substitution rule the inverse Jacobian appears, instead of the usual Jacobian.

5.2 The free fermion

With the technology at hand, it is now possible to introduce the first theory with fermions. This will be, of course, a free theory. As a free theory, it requires that it is at most quadratic in the fields. It needs also to be a Poincaré-invariant theory. Based on the experience with scalars, it is to be expected that a term without derivative will act like a mass term. Thus, it remains to be seen what can be done in terms of a derivative term.

The first attempt would be to include a ∂^2 term, as with the scalars. However, because the action of a Lorentz transformation on a spinor is not given by a unitary transformation, but by a non-unitary one (4.6). However, the operators $\sigma_\mu p^\mu$, $\bar{\sigma}_\mu p^\mu$ have been found to map between the two SU(2), and also $\sigma_2 \psi^*$ belong to the other SU(2) group. Thus, possible free Lagrangians for the Weyl spinors are

$$\begin{aligned} \mathcal{L}_\psi &= \psi^\dagger i \sigma^\mu \partial_\mu \psi - \frac{1}{2} m \psi^T (-i\sigma_2) \psi \\ \mathcal{L}_\chi &= \chi^\dagger i \bar{\sigma}^\mu \partial_\mu \chi - \frac{1}{2} m \chi^T (-i\sigma_2) \chi. \end{aligned}$$

Using (4.9) yields then immediately the Lagrangian for a Dirac fermion,

$$\mathcal{L}_\Psi = \overline{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi.$$

From this follows that the Dirac equation (4.10) is the Euler-Lagrange equation of $\mathcal{L}_\psi$. There is an interesting corollary from that. The Dirac equation, or the Klein-Gordon equation, yield four real conditions for the Dirac spinor. These conditions merely ensure the energy-momentum relation, and thus that the described particle ultimately respect special relativity. Thus, out of the eight real degrees of freedom for a Dirac spinor, only four are free if the Dirac equation holds. However, this is only required for an asymptotic state, and thus on-shell. Off-shell, the energy-momentum relation does not need to hold, and thus the degrees of freedom doubles. That is very different from the scalar case, where also the Klein-Gordon equation applies. But here no additional degrees of freedom become available off-shell.

It seems a bit odd at first that the derivative is only acting on one of the fields. However, due to the possibility to perform partial integrations, the derivative could be swapped or separated into two parts, acting on each field equally. It is, however, more convenient in practice to use this, slightly asymmetric, notation. Another interesting feature is that, because there is only one derivative, the mass-dimension of the fermion field needs to be $(d-1)/2$, and thus different from the scalar field. This also suggests that the constant term will have the mass, rather than the mass-squared, as constant of proportionality.

Finally, to quantize the theory, the path integral needs to be formulated. Just as with the multi-component scalar field in section 2.5.1, it is in principle possible to integrate separately over all real and imaginary components individually. However, in view of (5.8), it appears to be more suitable to perform a variable shift such that the independent variables are the field and the complex conjugate one. In fact, for the Dirac fermion it is even more convenient to integrate over $\overline{\Psi}$. Thus, the quantized free fermion is obtained from the path integral

$$Z_F = \int \mathcal{D}\overline{\omega}\mathcal{D}\omega e^{i\int d^dx\mathcal{L}_\omega + i\int d^dx(\bar{j}_\omega\omega + \overline{\omega}j_\omega)}$$

where ω can now be any of the different types of fermions, and ω the corresponding complex conjugated one. Because ω and $\overline{\omega}$ are treated as independent variables, each with half the number of degrees of freedom then the original field, they require each an individual source. Furthermore, because the action needs to be an ordinary, scalar number, both sources need to be Grassmann-valued spinors themselves. Because of the general Leibnitz rule 5.6, this implies that derivatives with respect to j_ω and $\bar{j}_\omega$ do not commute with each other, and also not with themselves. Hence, the order of derivation for matrix elements becomes important. Likewise, the time-ordering prescription (2.9) needs to respect the Grassmann nature now,

$$T(\alpha(t_1)\beta(t_2)) = \theta(t_1-t_2)\alpha(t_1)\beta(t_2) \pm (-1)^{\pi(\alpha)\pi(\beta)}\theta(t_2-t_1)\beta(t_2)\alpha(t_1).$$

This makes it more important to track orderings of expressions, as otherwise minus signs may get lost, which especially for relative minus signs can have serious consequences.

Using (3.13), it is straightforward to obtain the propagators in momentum space[1],

$$D^{\omega}_{ij}(p) = \frac{i(\Omega^{\mu}_{ij}p_{\mu} + m \times \mathbf{1}_{ij})}{p^2 - m^2 + i\epsilon} = (i(\Omega^{\mu}_{ij}p_{\mu} + m \times \mathbf{1}_{ij}))^{-1}, \tag{5.9}$$

where $\Omega = \sigma$, $\bar{\sigma}$, and γ for $\omega = \psi$, χ, and Ψ, respectively. It is important to note that the propagator is matrix-valued, being a matrix in the representation space of the spinors. Especially, (5.9) is indeed the matrix-valued inverse of $(\Omega^{\mu} + m)$, as an explicit calculation shows. This is also the reason for the often used short-hand notation displayed as the last equality in (5.9). Notwithstanding, the pole of the propagator is found again at m, in accordance with the fact that every fermion spinor component satisfies the Klein-Gordon equation, and thus the relativistic energy-momentum relation on-shell.

5.3 Yukawa theory

The next step is to include interactions. So far, the existence of non-trivial representations of the Poincare group made the fermion action unique. However, already when deciding which types of fermions and how many of them should be included, empirical input is needed. There is a further issue.

There is a further issue. According to the rules set out in section 2.2, there can be no self-interactions of fermions, due to the dimensionality of the fermion fields, at least in four dimensions. While there exist theories, which ignore this and construct interacting theories made entirely from fermions, they have usually a very limited focus. They will not be considered here.

Rather, combining both the scalar field and the fermion field creates an interesting interacting theory, Yukawa theory, with Lagrangian

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}\bar{\Psi}(i\gamma^{\mu}\partial_{\mu} - m_{\Psi})\Psi + \frac{1}{2}\partial^{\mu}\phi\partial_{\mu}\phi - \frac{1}{2}m_{\phi}\phi^2 && (5.10)\\ &\quad -\frac{\xi}{3!}\phi^3 - \frac{\lambda}{4!} - g\phi\overline{\Psi}_i\Gamma_{ij}\Psi_j. && (5.11)\end{aligned}$$

The first line (5.10) contains the free, or sometimes also called kinetic, terms. They describe a fermion Ψ with mass m_{Ψ} and a scalar with mass m_{ϕ}. The second line (5.11) contains the interactions. Besides the usual self-interaction of the scalar field, a new interaction between the fermion and the scalar appears, with a new independent coupling

[1] To interpret this beyond perturbation theory as the full two-point matrix element, note the discussion in section 2.5, especially in the sense of being a shorthand notation in the spirit of (2.29).

strength g. Such an interaction is called a Yukawa coupling. It contains a matrix Γ, which acts in the spinor representation space, and which is needed to be such that the Lagrangian remains Lorentz symmetry invariant.

In case of $\Gamma = \mathbf{1}$, the Yukawa term looks like a mass term, but with the mass replace by the expression $g\phi$. It thus has the character of a scalar interaction. An alternative is $\Gamma = i\gamma_5$. The corresponding Feynman rule is

$$\Gamma^{\overline{\Psi}\Psi\phi} = ig\Gamma = -\Gamma^{\Psi\overline{\Psi}\phi}$$

where now the ordering of the fermion fields, and the corresponding derivatives is important, due to the anti-commutativity of the fermion fields.

Yukawa-type theories have been used quite successfully used to describe nuclear interactions by meson exchange, where two Ψ fields, introducing a SU(2) symmetry in analogy to the SO(2) symmetry of scalars in section 2.5.1, correspond to the nucleons. Likewise, Yukawa interactions play a central role for the mass generation for the fermions in the standard model by the Brout-Englert-Higgs effects, and many scenarios beyond the standard model. These applications will be investigated in detail in the corresponding lectures.

5.4 Perturbation theory beyond scalars

While most of the rules of perturbation theory do not change by the introduction of fermions, or other particles with spin larger than zero, there are two exceptions. These require additional rules.

The first originates with the Grassmann nature of fermion fields, and consequently of the derivatives. It affects loops of fermions. For this note that the interaction Lagrangian necessarily always contains an even number of Grassmann fields, as does the free Lagrangian. A loop emerges whenever a propagator, or chain of propagators, closes in on itself. This will only happen, if the derivatives acting are not from the external sources. Thus, it only affects expressions like

$$\int d^d z_1 ... z_n \frac{\delta^n}{\left(\delta\overline{K}\delta K\right)^m ...} \Pi_i^m \left(\int d^d y_i \overline{K} \Delta K \right) ... \tag{5.12}$$

where the spinor indices and arguments are suppressed, and the ordering of the sources follows from the ordering in the Lagrangian to obtain a Lorentz scalar.

Start with the case $m = 1$. Then it is necessary to exchange one derivative with respect to K with $\overline{K}$, giving a minus sign. The second derivative works without a minus sign. In this way, for a single closed loop of fermions an overall minus sign appears. For the

$m = 2$ case, there are two options. Either, the derivatives $(\bar{1}1\bar{2}2)$, counting the position arguments, act in the order $(\bar{2}2\bar{1}1)$ or $(\bar{1}1\bar{2}2)$. These creates two loops with one propagator each, having each one closed fermion loop, and thus an overall factor of $(-1)^2$. Or, they act as $(\bar{1}2\bar{2}1)$, or any even permutations thereof. Then, they create a single loop with two fermion propagators. But now one three exchanges of derivatives are necessary. Thus, the loop with two propagators gets an overall minus sign. By induction, it can then be shown thus that any closed fermion loops gets an overall minus sign. In case of nested loops, all possible closed loops have their own minus sign.

The other changes actually only affects the LSZ rule, so it is not part of the fermion rules proper for matrix elements. The LSZ rule requires to amputate the matrix elements. So far, these had been scalars, and thus this was a quite straightforward product, as the propagators were scalars. In fact, amputation originated in general from (3.23-3.24). However, there was an implicit step made when moving from (3.23) to (3.24): That the indices on the derivatives are just counting the fields. That is certainly true for scalars. But if the fields carry a representation, Lorentz or internal, and thus indices $a...$, the meaning of

$$\frac{\delta\Phi_i^{a\cdots}}{\delta j_i^{a\cdots}}\frac{\delta W}{\delta\Phi_i^{a\cdots}}$$

is different, as the summation is also over representation indices[2]. Thus, in the amputation process the propagators and the amputated vertex function are contracted over the symmetry indices in (3.24).

This requires then to understand (3.25-3.26). The propagators are matrix valued in the representation indices. However, the amputated vertex function are vector-valued. It is thus necessary to contract it with a vector in the representation indices to obtain a scalar S-matrix element, needed for a cross section.

It is again group theory, which yields the answer. A matrix-(or higher tensor-)valued representation can be build by a tensor product from vector representations. Thus, a free propagator can be written as

$$D^{ab}(p) = Z\frac{u_a(p)u_b^\dagger(-p)\gamma_0}{p^2 - m^2 + i\epsilon}$$

where u are the Bethe-Salpeter amplitudes[3] or, if the particle is elementary, the wave function. In case of a real, elementary scalar, these wave functions are 1. In other cases of elementary particles, they could be reverse engineered from the propagator. For bound states, it is more complicated, and this will be addressed in section 9.2. However, it is worthwhile to also pursue a different approach, which will be useful in chapter 10.

[2]Implicitly, this even hides more, the summation over invariant tensors of section 2.5.1.

[3]In case of higher tensors Faddeev amplitudes.

Because the propagator is the Green's function of the kinetic operator

$$\Delta^{ab} D^{bc} = \delta(x-y)\delta^{ac}$$

it follows that the wave functions need to fulfill correspondingly this equation. In case of fermions, this requires them to fulfill the Weyl equation or Dirac equation. Since the results for the Weyl equations can be read off from (4.13-4.14), it is sufficient to solve the Dirac equation

$$(p_\mu \gamma^\mu - m)\, v = 0,$$

where for convenience this was already changed to momentum space using place waves, as are fit for the free, elementary particle. Solving this algebraic equations yields

$$u = \begin{pmatrix} \sqrt{p_\mu \sigma^\mu}\xi \\ \sqrt{p_\mu \overline{\sigma}^\mu}\xi \end{pmatrix},$$

where ξ is an arbitrary two-dimensional unit vector. Because of the connection to the propagator, it immediately follows that

$$u u^\dagger \gamma_0 = p_\mu \gamma^\mu + m.$$

Likewise, the charge-conjugated wave function v is

$$v = \begin{pmatrix} \sqrt{p_\mu \sigma^\mu}\eta \\ -\sqrt{p_\mu \overline{\sigma}^\mu}\eta \end{pmatrix},$$

with an independent two-dimensional, normalized vector η.

Thus, ultimately, the Feynman rules have to be extended by demanding that external lines of the amputated matrix elements of the S matrix need to be contracted with wave functions. In this way, all Lorentz indices are contracted, creating a Poincaré-invariant matrix element.

Chapter 6

Discrete symmetries

There is an interesting detail in Yukawa theory. What is the difference between the choice $\Gamma = \mathbf{1}$ and $\Gamma = i\gamma_5$ in (5.11)? It turns out that this question leads t one of the most central theorems of quantum field theory.

6.1 Discrete symmetries

So far, both internal symmetries and external symmetries have been continous. But already the linear-σ model (2.2) allows at $\xi = 0$ a discrete symmetry by $\phi \to -\phi$. This is a Z_2 symmetry. While much of what has been said for internal symmetries in section 2.5 applies to discrete symmetries equally, there are two important differences. One is that constructing the equivalence to (2.29) can be a bit more involved. The other is that there are far-reaching consequences for space-time symmetries.

So far, only the SO(1,3) proper Lorentz group has been considered in section 4.1. But there are two more discrete symmetries[1], parity and time-reversal, both already known from quantum mechanics.

6.1.1 Parity

Parity is the more straight-forward one to understand. The parity operation P replaces $\vec{x} \to \vec{x}$ and $\vec{p} \to -\vec{p}$, but $\vec{l} \to \vec{l}$, if $\vec{l}$ is the angular momentum. It does not affect the time component, and thus acts only as $P = \mathrm{diag}(1, -1, -1, -1)$ in spatial subspace. Because $P^2 = \mathbf{1}$, it is a Z_2 symmetry, Z_2^P.

[1]Note that one decisive differences between the Euclidean case SO(4) and the Lorentz case SO(1,3) is that in the former case there is only one such discrete symmetry.

This leads to an important question: Who does this affect the fields? And does the classical action and/or the partition sum needs to be invariant under it, such that the field transformation properties can be inferred from it?

The answer to the second question is already provided by experiment: No. The weak interactions violate parity. Thus, it is not necessary, and should not be necessary for quantum field theory, to respect parity. This is as well, as parity is not an inherent part of Poincaré symmetry, but an optional tensorized subgroup.

This leaves to deal with the fields. As for any other symmetry, this poses the question, in which representation a field φ should transform under Z_2^P? This question can even be posed if parity is not a manifest symmetry of the theory. Since all possible irreducible representations of Z_2^P are either the trivial one or isomorphic to ± 1, there are only two possibilities to choose from,

$$\begin{aligned} P\varphi &= \varphi \\ P\varphi &= -\varphi \end{aligned}$$

up to a redefinition of the eigenvalue -1. Of course, a third possibility is that there is no defined behavior of a field at all. It turns out that this case has not been realized in nature,at least not for elementary fields, so it will be ignored for now.

Consider first scalar particle ϕ. It could indeed be in the trivial representation. Or it could change to $-\phi$. In the latter case it is called a pseudoscalar field. For a scalar field, the linear-σ model (2.2) is parity-invariant for all values of the coupling. In case of a pseudoscalar field, this is only true if $\xi = 0$ Otherwise, parity is explicitly broken. The assignment of the parity representation to the scalar field is part of the definition of the model, and can thus not be derived.

For fermions, the situation is a bit more involved. Because the matrix P in its spinor representations does not commute with the γ_μ matrices, this yields

$$P\Psi = \eta\gamma_0\Psi$$

where η is an arbitrary phase, which is the same for all spinors. It stems from the fact that Ψ is complex-valued, and thus an additional phase is possible. It can be chosen by convention, and is usually done to eventually have that the parity eigenvalue of the electron is $+1$. This is different to the case of the scalars, where no such choice exists, and stems from the properties of the spinor representation. This implies that parity exchanges necessarily left-handed and right-handed objects. This is not surprising. Spin is, after all, a kind of angular momentum. Thus, the projection of spin on the momentum changes, and hence the handedness. Therefore, a theory which contains Weyl fermions of one handedness only will not support parity.

For Dirac spinors, $P^2\Psi = -\eta^2\Psi$, since $P\gamma_0 = -\gamma_0 P$. The conventional choice is $\eta = 1$. Note that because the spinors are Grassman-valued, the parity of products of such spinors need to be taken into account.

It should be noted that under parity γ_5 is odd, i. e. $P\gamma_5 = -\gamma_5$. This follows as γ_μ acts like a vector under parity, and thus the definition (4.12) yields a factor $(-1)^3 = -1$ explicitly. Thus, an expression like $\overline{\Psi}\gamma_5\Psi$, like appearing in (5.11) for $\Gamma = i\gamma_5$, is odd under parity, in contrast to $\Gamma = \mathbf{1}$, which is even. Thus, if the scalar in (5.11) is even under parity, such a pseudoscalar coupling explicitly breaks parity. Alternatively, if setting $\xi = 0$ and $\Gamma = i\gamma_5$ in (5.11) and choosing ϕ to be a pseudoscalar field, will create again a parity-invariant theory. What is the correct choice is an input to theory building, and thus driven by experiment.

6.1.2 Time reversal

Time reversal appears at first to be just be the analogue of parity, just with a different matrix. However, it turns out to be more subtle than that, as was already the case in quantum mechanics. Of course, time appears to be no longer a special case, as spatial coordinates are now also just numbers, rather than operators. But the existence of cluster decomposition (3.17) and that correlation across time-like distance do not decay exponential, but over space-like distance does, (3.5), highlights already two differences.

In fact, more important is that in probabilities like (3.18) still matrix elements like $\langle i|f\rangle$ enter. But time reversal should replace initial states and final states. At the same time, if the theory is invariant under time reversal, which is not necessary, this requires that

$$\langle i|f\rangle = \langle PQ\rangle = \langle T^{-1}QPT\rangle = \langle T^{-1}QTT^{-1}PT\rangle \stackrel{!}{=} \langle QP\rangle = \langle f|i\rangle = \langle i|f\rangle^*.$$

This will work, if

$$\langle T^{-1}QTT^{-1}PT\rangle = \langle Q^*P^*\rangle = \langle (PQ)^*\rangle = \langle PQ\rangle^*.$$

Thus, as in quantum mechanics, time-reversal needs to be anti-unitary, and thus involve complex conjugation $\mathcal{C}$. Hence, in general $C = \mathcal{C}\Lambda_C$ with $\Lambda_C = \mathrm{diag}(-1,0,0,0)$. On coordinates, momentum, and angular momentum this has therefore the ordinary effect, as they are real.

However, this has a different effect on fields. Only for a real scalar field nothing changes, which includes the case with an SO(N) symmetry in section 2.5.1. Since still $T^2 = \mathbf{1}$ applies, the field ϕ could be assigned a definite time parity, in analogy to parity. For reasons to become clear in section 6.3, this is usually not detailed.

If, however, the field is complex, it follows

$$T\phi(x) = \phi^*(x). \tag{6.1}$$

The field is thus not an eigenstate of T, and is neither even or odd. The situation is more involved when considering fermion fields. Time reversal will send $\vec{p} \to -\vec{p}$, but spatial rotations are not affected, and also $\vec{x} \to \vec{x}$. Thus, the Pauli-Lubanski vector and thus spin flip as well. This leaves the helicity intact. Thus left-handed Weyl spinors need to be turned into left-handed complex conjugated Weyl spinors, and likewise for right-handed Weyl spinors. But it does not leave the spin component intact. If time is reversed, the spin components need to be reversed, which corresponds to flipping the components of the Weyl spinors, and likewise exchange the eigenvalue of the third component of the spin.

Noting that $\sigma_\mu\sigma_2 = -\sigma_2\sigma_\mu^*$, and thus for any arbitrary projection of σ_μ, it follows that

$$\begin{aligned} T\psi &= -i\sigma_2\psi^* \\ T\chi &= -i\overline{\sigma}_2\chi^*, \end{aligned}$$

and thus for a Dirac spinor

$$T\Psi = i\gamma_5\Psi^* = -\gamma_1\gamma_3\Psi^*$$

where the rewriting in terms of γ_1 and γ_3 is useful for the generalization to different dimensionalities.

Explicit calculation shows that the kinetic term in (5.10) is invariant under T. For the interactions in (5.11) the case $\Gamma = \mathbf{1}$ is likewise. However, for $\Gamma = i\gamma_5$, the term is odd. Thus, the term can only be T and P invariant if ϕ is a T-odd pseudoscalar.

6.2 Anti-particles and Majorana spinors

An interesting twist appears in the case of introducing a complex scalar (6.1). Consider the linear-σ model for $\xi = 0$, and take the interaction to be $(\phi^\dagger\phi)^2$. The theory is then invariant under the symmetry

$$\phi \to e^{i\alpha}\phi \approx (1 + i\alpha\mathbf{1})\phi + \mathcal{O}(\alpha^2) = \phi + i\alpha\delta\phi + \mathcal{O}(\alpha^2),$$

and thus a U(1) phase symmetry. As in quantum mechanics, a Hermitiean operator, here $\mathbf{1}$, is thus associated with the symmetry. In the more general case of SO(N) symmetry in section 2.5.1, these were the generators of the so(N) algebra, creating a different value for $\delta\phi$.

Assuming that it is a classical symmetry, and thus the Lagrangian itself is invariant, it follows that the variation of the Lagrangian under this change obeys[2]

$$0 = \delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\partial^\mu\phi}\partial^\mu(\delta\phi)$$

The equation of motion for ϕ

$$\frac{\partial\mathcal{L}}{\partial\phi} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\right) \tag{6.2}$$

holds classically. At the quantum level, the same result follows for arbitrary expectation values, and thus on the field level. This can be seen as follows. The path integral is by definition translational-invariant. Thus

$$0 = \int\mathcal{D}\phi\frac{\delta}{i\delta\phi}e^{iS+i\int d^dxj\phi} = \int\mathcal{D}\phi\left(\frac{\delta S}{\delta\phi}+j\right)e^{iS+i\int d^dxJ\phi} = \left\langle T\left(\frac{\delta S}{\delta\phi(x)}+j(x)\right)\right\rangle.$$

At vanishing source $j(x) = 0$, this implies that the expectation value of the action is extremal, thus yielding the quantum version of the equation of motion. It should, however, not been considered to be correct on the field-level or at finite source.

Having this at hand, yields

$$0 = \left(\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\right)\right)\delta\phi + \frac{\partial\mathcal{L}}{\partial\partial^\mu\phi}\partial^\mu(\delta\phi) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\delta\phi\right) = \partial_\mu J^\mu, \tag{6.3}$$

and thus the current J is conserved, at the very least on the level of expectation values. This consequence, the existence of a conserved current for any continuous symmetry, is known as Noether's theorem. Note that a discrete symmetry does not allow for a derivative in (6.2), and thus no such statement can be made.

It should be noted that this does not change, if a total derivative $\partial_\mu K^\mu$ is added to the Lagrangian, and thus a boundary term to the action. However, this will lead to a replacement of the Noether current,

$$J^\mu \to J^\mu + K^\mu = \tilde{J}^\mu.$$

as the total conserved current.

Given a conserved current, a total conserved charge can be defined as

$$Q(t) = \int d^{d-1}xJ_0(x),$$

which is temporally conserved

$$\partial_0 Q = \int d^{d-1}x\partial_0 J_0 = -\int d^{d-1}x\partial_i J_i = 0.$$

[2]Assuming that the Lagrangian is at most quadratic in derivatives.

The last equality follows either by requiring to have a vanishing charge density at infinity, or, using cluster decomposition (3.17), that fluctuations at infinite space-like distances need to be uncorrelated, and thus average out. A similar calculation will thus show that in the SO(N) there is a conserved charge Q^a for every dimension of the representation space of the field.

It also follows that in the present case

$$J_0 = -\phi^\dagger \mathbf{1}$$

where a partial integration has been performed. Thus the negative of the complex-conjugated field yields the charge density, and likewise for the field ϕ. Thus, ϕ and $\phi^\dagger$ carry the opposite charge, but are otherwise degenerate. Because of this, $\phi^\dagger$ is said to be the anti-particle of ϕ. As will be seen in section 6.3, the existence of a charge will imply the necessity that for every charged particle an anti-particle with the same mass as J^P needs to exist.

Remarkably, the existence of a particle and anti-particle gives rise to another symmetry, charge conjugation symmetry C. Defining in the present case that C is complex conjugation leads to

$$C\phi = \phi^\dagger,$$

under which indeed the complex linear-σ model is invariant. Again, a phase can be introduced, which is here set to 1. As P and T, it also applies that $C^2 = 1$. Note furthermore that if ϕ would be real, the action of complex conjugation is none. But it is still possible to define in that case that

$$C\phi = \pm\phi.$$

The field ϕ is then said to be even or odd under charge parity, and this quantum number is added to J^P as J^{PC}. The linear-σ model is then again invariant under this transformation for $\xi = 0$. But its breaking at $\xi \neq 0$ indicates that this symmetry does not need to be respected, again very much as T and P. It should be noted that the C-symmetry could also be interpreted as a discrete Z_2 symmetry of the theory, as such a symmetry has the same effect. This issue will be taken up again in section 6.3.

In the SO(N) case, where there is also a conserved charge but the fields are real, the question arises how to interpret C there. As U(1) is an Abelian symmetry, the solution lies again in group theory. For any non-Abelian Lie group, there exists a complex conjugate representation. Given the original representation of the generators T^a, it is obtained in the same representation space from $-T_a^*$. Charge conjugation can now be taken as the transformation, which maps states between the representation and its complex-conjugate representation. If the representation is real or pseudoreal, this will, up to a similarity

transformation, not change the fields. Then, a definite C quantum number can be assigned to the fields. Otherwise, both representations are distinct, and the concept of a particle and an anti-particle arises. Of course, when assigning the charge parity $+1$, this implies that particle and anti-particles can be also thought of as being identical. This is indeed what happens for the SO(N) model.

For fermions, the situation is again more subtle. Start with Dirac fermions. They are inherently complex, and to have a bosonic Lagrangian thus requires that both Ψ and $\overline{\Psi}$ are always appears in pair. Thus, there exists necessarily a global continous symmetry of type

$$\Psi \to e^{i\alpha}\Psi,$$

with a corresponding conserved charge. This charge is called fermion number, and thus implies that every Dirac fermion has an anti-particle[3], and that they can only be created in pairs of fermions and antifermions.

However, there is now the twist that fermions are also spinor representations of SO(1,3). Thus, this needs to be taken into account when defining C for Dirac fermions. The spinor representation is pseudo-real. There is thus the possibility to define a phase, and moreover the question, if spin-orientation should be altered by C. The standard definition is

$$C\Psi = -i\gamma_2\Psi^*,$$

which implies that $\overline{\Psi} = (C\Psi)^T\gamma_0$, thus maintaining the spin orientation, and thereby making $\overline{\Psi}\Psi$ invariant under C by construction. Therefore, the free Dirac fermion action is also C-invariant.

This implies for Weyl fermions

$$\begin{aligned} C\psi &= -i\sigma_2\chi^* \\ C\chi &= -i\overline{\sigma}_2\psi^*, \end{aligned}$$

thus C changes the chirality of Weyl fermions, and they are thus not invariant under C. Especially, only a theory which contains both left-handed Weyl fermions and right-handed complex-conjugated Weyl fermions can be C invariant, if the transformation is defined in this way. However, there is alternative option.

Define a Dirac spinor by

$$\Psi = \begin{pmatrix} \psi \\ -i\sigma_2\psi^* \end{pmatrix},$$

[3]Note this entirely avoids the discussion of the outdated concepts of Dirac sea or negative energy modes, which therefore will not be touched upon here.

and likewise for the other handedness. Due to (4.8), this will transform as a Dirac spinor, but satisfies by construction $C\Psi = \Psi$. Such a spinor is called a Majorana spinor[4], and is hence its own anti-particle. Note that such a particle cannot carry a charge, as, e. g., a phase rotation would act differently on the different components, and it would thus not be possible to build an invariant theory. This includes fermion number, as a particle can be turned into an anti-particle in this theory. Still, Lorentz symmetry forbids to have an odd number of fermionic fields in interaction vertices in the Lagrangian.

Since Majorana spinors transform otherwise like Dirac spinors, it is possible to build a theory in terms of them. However, there is a drastic difference. The number of degrees of freedom is halved with respect to a genuine Dirac spinor. Note that by construction both handedness are now included, and all terms constructed in this way are by construction C-invariant. Thus, insisting on Majorana fermions limits the number of possible expression in the Lagrangian. Phrased otherwise, there are terms, constructed from Weyl fermions, possible, which manifestly break C parity.

6.3 The CPT theorem and the spin-statistics connection

It is now an amazing consequence that the Poincaré group implies the Pauli principle only for half-integer spin particles. Before moving there, it is useful to first proof a related theorem, the CPT-theorem. It states that while each of the C, P, and T symmetry may be broken, as well as any of the combinations CP, CT, and TP, any quantum field theory needs to be invariant under the combined application of CPT. Because C is included, this implies that in the presence of charged particles also always their anti-particles need to be part of the theory.

The proof is relatively straightforward. Up to pre-factors, the combined application of the 8commuting) operators C, P, and T is, up to a pre-factor, essentially that all vectors change sign, and fermion pick up signs and factors of i. In addition, right-handed spinors pick up one more sign than left-handed ones. Thus, scalar spinor bilinears always lead to expressions like $-i \cdot i = 1$. Due to Lorentz covariance, both the classical action and the quantum effective action are scalars. Thus, all signs from scalar products of vectors also cancel, as do any factors from non-trivial representations of P and C. Thus, the only remaining part is that the arguments of the fields change from x to $-x$. But as the action is a space-time independent integral over all space, this will not have any consequence. Thus,

[4]Note that this construction, and thus equivalence of Weyl spinors and Majorana spinors, does not work in all dimensionalities. In some cases, both are distinct concepts.

any quantum field theory, for which a quantum effective action exists, will be invariant under CPT, if it is invariant under Poincaré symmetry, and vice versa.

To see the connection between spin and statistics, consider an arbitrary matrix element,

$$w(x_1, ..., x_n) = \langle \phi^1_{i_1}(x_1)...\phi^n_{i_n}(x_n) \rangle \tag{6.4}$$

which can be considered to be a function w of the space-time coordinates only, if the multi-indices i_j are suitable summed over. Complexify the arguments, $x_i \to z_i$. The so defined functions of n complex variables are called Wightman functions. This allows to group the complex space-time variables also in terms of the two SU(2) groups as a Dirac spinor.

It is worthwhile to note that the existence of the path integral would imply certain analyticity properties of the Wightman functions on certain patches of their arguments. While these will not be detailed here, the probably most important one is a reconstruction theorem, which implies that any Wightman function of real arguments x_i can be expressed in terms of algebraic combinations of Wightman function at analytically continued arguments $x_i \to i\tilde{x}_i$. As these would be the Wightman functions of an Euclidean space-time, this implies that it is, in principle, possible to do calculations of any desired correlation function in an Euclidean space-time, and then continue the result back to Minkowski space-time. In practice, the required algebraic combination is not always easily accessible, but that issue is often compensated by the fact that Euclidean space-time is often technically much more favorable than Minkowski space-time. This is physically remarkable, as the causal space-time structure of Minkowski space-time and Euclidean space-time are drastically different. Where in practice this is useful depends on the theory and the problem at hand. But especially for numerical evaluation, this reconstruction theorem is indispensable.

Perform now a full rotation in either of the two SU(2) subgroups of the Lorentz group. Spinors of this representation will then necessarily pick up a minus sign for each fundamental representation in a given tensor product, see section 4.2. All other fields will remain invariant, but this will also change the z. If the SU(2) is chosen for which $z \to -z$, and there are m fields in this SU(2), it follows that

$$w(z_1, ..., z_n) = (-1)^m w(-z_1, ..., -z_n) = (-1)^m (-1)^p w(-z_n, ..., -z_1),$$

where p is the number of anticommuting fields. Now, set all imaginary parts to zero[5] and

[5]Herein it is assumed that the Wightman function are, at least in some domain, without cuts and sufficiently smooth to allow for that kind of arguments. This can actually be asked as part of the axioms, but will also follow automatically if the path integral would be well defined.

select the case of a correlation function of two fields as

$$(-1)^m(-1)^p w(-x_2,-x_1) = w(x_1,x_2) = \langle \phi^1(x_1)\phi^2(x_2)\rangle = w(|x_1-x_2|) = w(-x_2,-x_1).$$

where translational invariance has been used. This implies that $m+p$ must be even, and thus spinors need to anticommute. As spinors steam from half-integer spins, this implies that half-integer spins are Grassman-valued, thus obey the Pauli principle, and hence are fermions, while integer spin particles need to be bosons and do not obey the Pauli principle.

It is thus once more the space-time structure which imposes physical properties on particles, especially tying their statistical nature to their space-time representation properties[6]. Since the space-time structure also provides causality with the light-cone structure and the existence of world-lines, the existence of causality, anti-particle and the spin-statistics relation are tightly intertwined. But this also implies that in a different space-time structure, e. g. 4-dimensional Euclidean space, things could be very different.

[6]Note again that 1+1 dimensions work differently, due to the Abelian nature of space-time.

Chapter 7

A first look at QED

So far, the particles had been restricted to ones with a spin of 1/2 or less. Going beyond that introduces the fact that particles with spin of one or larger necessarily introduce gauge freedom. Conceptually, this is similar, but usually technically more involved, to classical electrodynamics. Quantum gauge theories offer a whole new range of both interesting phenomena and technical complications. Thus, here, only the simplest possible case of a spin one particle will be introduced. Fortunately, this already yields a highly relevant theory, quantum electrodynamics.

7.1 Formulation and Maxwell theory

There are two possibilities to extend the representations of the Poincaré group beyond 1/2. One is to consider representations of the type $(1,0)$ or $(0,1)$. However, these do not offer anything new, as they turn out to be tensor products of the twice the corresponding fundamental representation. Physically, they can be interpreted as particles build from two Weyl fermions.

The more interesting option is the first case of a non-trivial representation for both SU(2) algebras, $(1/2,1/2)$. Incidentally, this is the same representations the momentum is in, and is thus called a vector representation. Depending on whether they a parity-even or odd, they are called vectors and axialvectors[1].

Thus, this is a tensor product of the two fundamental representations, and thus would, at first glance, yield four degrees of freedom, and at minimum a real-valued vector field A_μ. However, a Clebsh-Gordan construction shows that only three of the four degrees of freedom form a vector, $m = -1, 0, +1$. The other forms a scalar. This implies that a four

[1]It can be confusing that the additional 'axial' (and not pseudo) here signifies positive parity, while for scalars the additional 'pseudo' signifies negative parity. That is an historical oddity once more.

vector has a redundancy. That is already known from the momentum, as $p^2 = m^2$ fixed one of the components in terms of the others.

The same is now true here. Lorentz invariance fixes one of the components, such that $A_\mu A^\mu$ is a Lorentz-invariant quantity. However, this entails that it is possible to perform transformations of the four components, which only need to ensure this condition. In addition, if the field is massless, then the number of degrees of freedom needs to be reduced further, as instead of spin it now only carries helicity, see section 4.1. This implies that the vector A_μ is further restricted. In momentum space, $A_\mu(p)$, there is besides $A_\mu A^\mu$ only one other possible invariant, which can be build, $p_\mu A^\mu$, which therefore needs to be fixed as a second condition. Physically, the first condition is only a normalization of the field, which can be set arbitrary, and thus doe snot distinguish theories. Likewise, choosing $p_\mu A^\mu = 0$ is a unique choice, given that no other quantity exists except for $m = 0$, which can characterize the field as a representation. Thus, they different right-hand sides cannot describe different theories. The latter is especially visible when going beyond four dimensions. There, the number of vector components change. In fact, a more accurate description is to describe vector fields as anti-symmetric three-dimensional tensors, which make the number of degrees of freedom manifest, and where the SU(2) transformations act as left-multiplications and right-multiplications in the adjoint representation, respectively. Or as an $\mathrm{SU}(2)/\mathrm{Z}_2$ matrix, then independent under corresponding fundamental action. However, this has not been the option of choice, mainly again for historical reasons, the interplay with space-time coordinates, and the generalization to other dimensionalities.

That therefore a(n axial) vector field can only represent two degrees off freedom is encouraging. After all, this is exactly the number of degrees of freedom of a classical electromagnetic field. However, this is not manifest at first sight. Consider classical Maxwell theory. It is formulated in terms of the vector potential A_μ. However, rather than to implement the constraints directly, the theory is formulated in terms of the full vector field A_μ, which will become the photon field later, with its four component. This remaining freedom manifest by allowing for gauge transformations, as

$$A_\mu \to A^g_\mu = A_\mu + \partial_\mu \omega, \tag{7.1}$$

where ω is an arbitrary function. A defining property of such gauge transformations is the fact that they do not alter any measurable quantities. This is, as it needs to be, as they only encode information on the choice of representation, but no dynamical information

In particular, the electric and magnetic fields $\vec{E}$ and $\vec{B}$, which are obtained from the

gauge potential by

$$\vec{E}_i = -\frac{\partial}{\partial t}A_i - \partial_i A_0 \tag{7.2}$$
$$\vec{B}_i = (\vec{\nabla} \times \vec{A})_i, \tag{7.3}$$

are invariant under such transformations. From the vector potential, it is possible to form the field strength tensor

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

which is also invariant under gauge transformations. The Maxwell equations can then be written in the compact form

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= J^\nu \\ \partial^\mu F^{\nu\rho} + \partial^\nu F^{\rho\mu} + \partial^\rho F^{\mu\nu} &= 0, \end{aligned}$$

where J^μ is the electric current. These are the equations of motions of the classical Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - J^\mu A_\mu. \tag{7.4}$$

Setting the classical current J_μ to zero, this is known as Maxwell theory. This theory is non-interacting, as it is at most quadratic in the fields.

In the context of group theory, (7.1) can be considered to be in the adjoint representation of the U(1) algebra, with its only generator, unity. The gauge fields form therefore orbits under the gauge transformation (7.1). Different gauge orbits correspond therefore to distinct values of the electric field and magnetic field, and thus different physical situations.

At $J^\mu = 0$ the classical equations of motion for this theory are the vacuum Maxwell equations

$$\begin{aligned} \partial_\mu F_{\mu\nu} &= 0 \\ \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma} = \frac{1}{2} * F_{\mu\nu} &= 0, \end{aligned}$$

where the latter identity is sometimes also called (a) Jacobi identity. The quantity $*F_{\mu\nu} = \epsilon_{\mu\nu\rho\sigma}F_{\rho\sigma}$ is called also the Hodge dual or the dual field strength tensor. Both equations follow immediately from the antisymmetry of the field strength tensor. The only classical solutions of this theory is the vacuum as well as free electromagnetic waves, up to boundary conditions.

The gauge freedom (7.1) can be used to implement conditions, so-called gauge conditions, on the field. E. g., the Landau (or sometimes radiation gauge or in the classical case also Lorentz gauge) gauge is defined as

$$\partial_\mu A^\mu = 0. \tag{7.5}$$

I. e. the gauge freedom (7.1) is used to transform any given gauge field such that the condition (7.5) applies.

A gauge condition can always be written as a functional $C[A_\mu, x_\mu]$ such that $C[A_\mu, x_\mu] = 0$. It is not possible to demand any condition C. It is only possible, if for any arbitrary field configuration A_μ a function $g(x)$ exists such that $C[A^g_\mu, x_\mu] = 0$ can be satisfied. Depending on the form of $C[A_\mu, x_\mu]$ guaranteeing this can be an exercise in algebra, (partial) differential equations, or worse.

To see that the Landau gauge condition is well defined, consider

$$0 = \partial^\mu A^g_\mu = \partial_\mu A^\mu + \partial^2 g.$$

This is a wave equation for g with an inhomogeneity. Since the theory of partial differential equations implies that there exists a solution for any inhomogeneity, this implies that the Landau gauge is well defined.

Note, however, that it may happen that a gauge condition does not fully fixes the gauge freedom. E. g., the Landau gauge condition does not do so, as any gauge transformation satisfying a free-wave equation $\partial^2 g = 0$, a so-called harmonic gauge transformation, does not change it. Imposing as boundary conditions that the (gauge) fields vanishes at (spatial) infinity removes also this ambiguity, and fixes the gauge completely.

In fact, it can be useful to leave a certain degree of gauge freedom. E. g., the Coulomb gauge condition $\partial_i A_i = 0$ is only defined up to purely time-dependent gauge transformations. There is thus a whole class of functions satisfying the Coulomb gauge condition. Another examples are axial gauges with some fixed vector n_μ such that $n_\mu A^\mu = 0$. Such gauges can already simplify calculations considerably in certain cases.

It should be noted that a mass for the gauge bosons, e. g. $m A_\mu A^\mu$, cannot be added to the Lagrangian, as it will immediately break gauge invariance, as can be seen immediately by calculation. In fact, (7.4) is the unique gauge-invariant Lagrangian with terms up to power four in the fields in four dimensions. Thus, any additional interactions compatible with gauge invariance would lead to a much more involved theory. Thus, Maxwell theory is non-interacting and exactly massless without matter. It turns out that it is not possible to have a massive spin one particle as an elementary degree of freedom without involving either higher-order polynomials in the fields or introducing further fields.

7.2 Quantization

The gauge freedom creates a genuinely new challenge when quantizing Maxwell theory. The naive approach would be to quantize Maxwell theory by writing down the path integral

(2.7) and use (2.10) to calculate correlation functions. That's it. Unfortunately, there is a twist to this for gauge theories.

Start with the naive quantization of the free Maxwell theory with the classical Lagrangian (7.4) at $j_\mu = 0$ by writing down the generating functional

$$\begin{aligned} Z[j_\mu] &= \int \mathcal{D}A_\mu \exp\left(iS[A_\mu] + i\int d^dx j_\mu A^\mu\right) \\ S[A_\mu] &= \int d^dx \mathcal{L}, \end{aligned} \tag{7.6}$$

where the normalization has been absorbed into the measure for convenience. It is important to note that, despite the similarity, j_μ is now not the physical electrical current, but rather an auxiliary field introduced to more easily compute correlation functions.

The integral (7.6) is a Gaussian one. Hence, it should be possible to integrate it. Explicitly, it takes the form

$$Z[j_\mu] = \int \mathcal{D}A_\mu \exp\left(i\int d^dx \left(\frac{1}{2}A^\mu(g_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu)A^\nu + j^\mu A_\mu\right)\right). \tag{7.7}$$

However, it is not possible to perform this integral, since this would require the matrix

$$g_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu \tag{7.8}$$

to be invertible, which is not the case. This can be seen directly by the fact that its momentum-space version $g_{\mu\nu}k^2 - k_\mu k_\nu$ is a projection operator which vanishes when contracted with k_μ.

An alternative way to understand the problem is to note that any gauge transformation (7.1) leaves S invariant, and, as a shift, also does not influence the measure. Thus, there are flat directions of the integral, and thus the integral diverges when integrating along this direction. Thus, at first sight, it seems to be impossible to find a well-defined path integral for Maxwell theory.

There are only few possibilities to escape. One is to perform the quantization on a discrete space-time grid in a finite volume, determine only gauge-invariant observables and only after this take the continuum and infinite-volume limit. This is in most cases only feasible numerically, but then a rather successful approach, so called Lattice Monte Carlo simulations of lattice gauge theory. This approach is a subject of a separate lecture.

Another one is to determine only quantities which are invariant under gauge transformations. Classically, this is perfectly feasible, as an explicit formulation only in terms of electric and magnetic fields, both manifestly gauge-invariant, shows. At the quantum level, however, this is not so easy. After all, this amounts to replacing the vector potential

by the electromagnetic fields. The Jacobian then involves solving the partial differential equations (7.2-7.3), yielding non-trivial space-time integrals, and ultimately leading to a non-local action. Still, this is possible, though very, very painful. It is, however, far too painful for the kind of generalization of gauge theories needed in particle physics, and even for QED in practical problems far too complicated. Thus, while reassuring that it would be conceptually possible, is not a practical solution.

In this context one may wonder why not directly integrating over the electromagnetic fields in the path integral (7.7) instead of the gauge fields. This is an inequivalent path integral to the one obtained by solving (7.2-7.3) as a variable transformation in (7.7). Interestingly, while mathematically perfectly feasible and doable, this does not yield a quantum theory compatible with experiments. Why this is so is a very good, and unsolved, question. It is a purely empirical outcome. This is even more surprising as the gauge-fields are classically an irrelevant technical tool, and actually introduce redundancies. They are thus quite the opposite to what is usually done in classical mechanics, the reduction to generalized coordinates. Still, if one wants to base the construction of the theory on experiment, one is therefore stuck with (7.7).

There is a third option, and the one to be used in this lecture. As the possibility to introduce gauge conditions such as (7.5) shows, the additional degrees of freedom over which are integrated do not carry physical information[2]. The aim is therefore to separate off this unphysical degrees of freedom in a way which allows to remove it. In the end, this will lead to a(n infinite) prefactor of the path integral, which drops out when normalizing the partition function at zero source to unity. Even though unphysical information is removed, this is not equivalent to introduce a gauge-invariant formulation. Any gauge condition will define one distinct way of removing the superfluous degrees of freedom, but eventually one still works with the gauge fields A_μ. But this field will differ for every gauge condition, and therefore depends on the gauge condition. But this difference depends only on the gauge condition, and will therefore drop out in gauge-invariant quantities, making them physical.

After these initial remarks, it is now time to actually perform this gauge-fixing, and quantize Maxwell theory. To this end select, as in classical electrodynamics, a gauge condition $C[A_\mu, x] = 0$ which selects uniquely the gauge field. I. e., for a set of gauge-fields related by gauge transformations (7.1) there is one and only one, but also at least one, which satisfies the condition C when going through all possible gs. This is actually not a necessary condition to make progress, and it will be relaxed later on. But for the moment,

[2]While this is obvious at the classical level, this remains true at the quantum level because of the demand that also at the quantum level physics is independent of the observer's choice of gauge.

it makes the calculation simplest.

An example of such a condition is, e. g., the Landau gauge (7.5), which in this language takes the form, suppressing from now on possible explicit dependencies on x_μ,

$$0 = C[A_\mu] = \partial^\mu A_\mu. \tag{7.9}$$

To make the path integral well-defined, it is then necessary to get rid of all the gauge transformed fields which do not satisfy this condition. Then just one representative for each physically inequivalent field configuration is left[3].

The question is now how to enforce this in practice. To do this consider the functional generalization of the Dirac-δ function. The expression

$$\Delta[A_\mu]^{-1} = \int \mathcal{D}g\delta(C[A^g_\mu])$$

contains an integration over all gauge-transformations g for a fixed physical field configuration A_μ, but by the δ-function only the weight of the one configuration satisfying the gauge condition is selected. Hence, when performing the change of variables $g \to g + g'$ with some gauge transformation g' it remains unchanged by definition: The functional integral is translationally invariant. As a consequence, Δ is actually gauge-invariant. Evaluating it at the gauge-transformed configuration $A^{g'}_\mu$ yields

$$\begin{aligned}\Delta[A^{g'}_\mu]^{-1} &= \int \mathcal{D}g\delta(C[A^{g+g'}_\mu]) = \int \mathcal{D}(g-g')\delta(C[A^g_\mu]) \\ &= \int \mathcal{D}g\delta(C[A^g_\mu]) = \Delta[A_\mu]^{-1}.\end{aligned}$$

Inverting Δ, the relation

$$1 = \Delta[A_\mu] \int \mathcal{D}g\delta(C[A^g_\mu]) \tag{7.10}$$

is found.

Inserting this into the functional integral yields

$$\begin{aligned}Z &= \int \mathcal{D}A_\mu \Delta[A_\mu] \int \mathcal{D}g\delta(C[A^g_\mu]) \exp(iS[A_\mu]) \\ &= \int \mathcal{D}g \int \mathcal{D}A^{g'}_\mu \Delta[A^{g'}_\mu]\delta(C[A^{g+g'}_\mu]) \exp(iS[A^{g'}_\mu]) \\ &= \int \mathcal{D}g \int \mathcal{D}A_\mu \Delta[A_\mu]\delta(C[A_\mu]) \exp(iS[A_\mu])\end{aligned} \tag{7.11}$$

[3]Note that this does not fix the harmonic part of the gauge transformations. However, they form a measure zero part of the whole set of gauge transformations, and therefore do not matter. Also, in any perturbative calculation the fields drop to zero at infinity, and thus in this case by definition they are excluded.

In the second line, a gauge transformation of the integration variable A_μ is performed. In the last line the inner variables of integration have been changed from $A_\mu^{g'}$ to $A_\mu^{-g-g'}$ and it has been used that all expressions, except the δ-function, are invariant. Hence, the integral over g is not influencing anymore the remaining integral, and contributes only a factor, which can be removed by appropriate normalization of the functional integral. In addition, it would have been possible to also replace the action by any gauge-invariant functional, in particular expressions involving some observable f in the form $f[A_\mu]\exp(iS[A_\mu])$. Thus, gauge-fixing is not affecting the value of gauge-invariant observables. Due to the δ-function, on the other hand, now only gauge-inequivalent field configurations contribute, making the functional integral well-defined.

It remains to clarify the role of the functional Δ. It was demanded that it is always possible to resolve the condition $C[A_\mu^g] = 0$ to obtain g as a function of C. Then, by exchanging C and g as variables of integration, it is found that

$$\Delta[A_\mu]^{-1} = \int \mathcal{D}C \left(\det \frac{\delta C}{\delta g}\right)^{-1} \delta(C) = \left(\det \frac{\delta C[A_\mu, x]}{\delta g}\right)^{-1}_{C=0},$$

The appearing determinant is just the corresponding Jacobian. Thus, the function Δ is given by

$$\Delta[A_\mu] = \left(\det \frac{\delta C[A_\mu, x]}{\delta g(y)}\right)_{C=0} = \det M(x, y). \tag{7.12}$$

The Jacobian has the name Faddeev-Popov operator, abbreviated by M, and the determinant goes by the name of Faddeev-Popov determinant.

A more explicit expression is obtained by using the chain rule

$$\begin{aligned} M(x,y) &= \frac{\delta C[A_\mu, x]}{\delta g(y)} = \int d^d z \frac{\delta C[A_\mu, x]}{\delta A_\mu(z)} \frac{\delta A_\mu(z)}{\delta g(y)} \\ &= \int d^d z \frac{\delta C[A_\mu, x]}{\delta A_\mu(z)} \partial_\mu^y \delta(y-z) = -\partial_\mu^y \frac{\delta C[A_\mu, x]}{\delta A_\mu(y)}. \end{aligned} \tag{7.13}$$

To proceed further, a choice of C is necessary. If C is a local gauge condition, i. e. only involving the fields and their derivatives, the Faddeev-Popov operator will be local, i. e.

$$M \sim \delta(x-y).$$

If this is not the case, the resulting theory will be inherently non-local. Such a local choice is always possible in perturbation theory. Choosing, e. g., the Landau gauge $C = \partial^\mu A_\mu = 0$ yields

$$M(x,y) = -\partial^2 \delta(x-y). \tag{7.14}$$

Due to the presence of the δ-function the functional $\det \Delta$ can then be replaced by $\det M$ in the path integral. Note that the result (7.14) is independent of the field variables. Thus,

this factor can be absorbed in the normalization constant. But then the original problem is solved and everything is complete. However, the resulting integral has always the implicit Landau gauge condition to be taken into account. Especially, this implies that the gauge field is always transverse.

This is implicit, and thus somewhat cumbersome. It can be made more explicit by taking a detour. To do so select as gauge condition

$$C = D[A_\mu, x] + \Lambda(x) \tag{7.15}$$

for some arbitrary function Λ. In general, this will make Lorentz symmetry not manifest. This can be recovered by integrating the path integral over all possible values of Λ with some arbitrary integrable weight function. Since the path integral will not depend on Λ, as this is a gauge choice, the integration is only an arbitrary normalization. Using a Gaussian weight, the path integral then takes the form

$$\begin{aligned} Z &= \int \mathcal{D}\Lambda\mathcal{D}A_\mu \exp\left(-\frac{i}{2\xi}\int d^dx\Lambda^2\right)\det M\delta(C)\exp(iS) \\ &= \int \mathcal{D}A_\mu \det M \exp\left(iS - \frac{i}{2\xi}\int d^dx D^2\right), \end{aligned}$$

where the δ-function has been used in the second step. For the most common choice $D = \partial_\mu A^\mu$, the so-called covariant gauges or R_ξ gauges, this yields the final expression

$$Z = \int \mathcal{D}A_\mu \exp\left(iS - \frac{i}{2\xi}\int d^dx(\partial_\mu A^\mu)^2\right). \tag{7.16}$$

This additional term has the consequence that the Gaussian integral is now well-defined, since the appearing matrix is changed to

$$g_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu \to g_{\mu\nu}\partial^2 - \left(1 - \frac{1}{\xi}\right)\partial_\mu\partial_\nu, \tag{7.17}$$

which can be inverted. The appearing parameter ξ, the so-called gauge parameter, is arbitrary, and can be chosen at will, defining the gauge. Furthermore, the ever-so popular Landau gauge corresponds to the limit $\xi \to 0$, as this is corresponding to the case where all of the weight of the weight-function is concentrated only on the gauge copy satisfying $\partial^\mu A_\mu = 0$. However, in principle this limit may only be taken at the end of the calculation, as it appears to recover the non-invertible original operator (7.8) in Maxwell theory. The choice $\xi = 1$ is known as Feynman gauge.

There is now one subtlety. What is the propagator of the photon, the answer to the question that started all of this? Because the matrix (7.17) is now invertible, this seems to be obvious. But it is not.

To understand what is happening, it is best to go back to the beginning. The photon propagator is the expectation value

$$\langle A_\mu(x)A_\nu(y)\rangle = \int \mathcal{D}A_\rho A_\mu A_\nu e^{iS}.$$

Irrespective of the issues with gauge symmetry, this expression will yield zero. The reason is, as noted before, that because space-time is isotropic, there is no preferred direction. Thus, a tensor cannot be non-zero. To avoid this problem, it is necessary to contract $A_\mu A_\nu$ with an invariant tensor of the Lorentz group. There are two symmetric rank two tensors suitable for the task, $g_{\mu\nu}$ and $p_\mu p_\nu$.

Denoting either of them as $P_{\mu\nu}$, the actual object to calculate is

$$\langle P^{\mu\nu}A_\mu(x)A_\nu(y)\rangle = \int \mathcal{D}A_\rho P^{\mu\nu}A_\mu A_\nu e^{iS}, \tag{7.18}$$

twice, once for each possible choice of $P_{\mu\nu}$. This object is well-defined from the space-time perspective, but not from the gauge symmetry perspective.

But before gauge-fixing, it is an interesting question to ask what happens if one tries to calculate the expression (7.18) by brute force. This can be done, and the result is actually $\sim \delta(x-y)$. The reason is that for any gauge field configuration with some value $A_\mu(x_0)$ at the fixed position x_0, there exists a gauge transformation, which is only non-vanishing at x_0, such that the value of the gauge transformed gauge field is $-A_\mu(x_0)$. In this way, any integration over the full gauge group yields zero. The only exception can happen if $x = y$, because then this is essentially the square of the gauge field, and thus positive. Hence, the propagator is only non-vanishing at coinciding space-time points. This argument can be extended to any gauge-dependent correlation function. Thus, without gauge-fixing all gauge-dependent quantities vanish up to expressions proportional to $\delta(x_i - x_j)$ for its arguments.

But this should be remedied by gauge-fixing. But it is not as simple. The decisive step is the expression (7.11). Here, the integral $\int \mathcal{D}g$ was absorbed in the normalization, because none of the remaining expressions depended on them, because all were gauge-invariant. This is no longer true, if the integral is taken over gauge-dependent quantities, like the photon propagator. Thus, the gauge transformation integral can then no longer be separated as a factor, and be removed. Thus, the whole procedure of gauge-fixing seems to break down for gauge-dependent quantities.

There if fortunately a solution to this. The expression (7.11) can also be interpreted differently for any gauge-invariant observable $f(A_\mu)$. Denoting the set of all gauge field configurations, including all gauge-transformed field configurations, as Ω, the following

expressions are identical

$$\begin{aligned} & \frac{1}{\mathcal{N}} \int_{\Omega} \mathcal{D}A_\mu f(A_\mu) \exp(iS[A_\mu]) \qquad (7.19) \\ = & \frac{1}{\mathcal{N}} \int_{\Omega/\Omega_c} \mathcal{D}g \int_{\Omega_c} \mathcal{D}A_\mu \Delta[A_\mu] \delta(C[A_\mu]) f(A_\mu) \exp(iS[A_\mu]) \\ = & \frac{1}{\mathcal{N}'} \int_{\Omega} \mathcal{D}A_\mu \Delta[A_\mu] \delta(C[A_\mu]) f(A_\mu) \exp(iS[A_\mu]) \\ = & \frac{1}{\mathcal{N}''} \int_{\Omega_c} \mathcal{D}A_\mu \Delta[A_\mu] f(A_\mu) \exp(iS[A_\mu]), \end{aligned}$$

where Ω_c is the set of all gauge field configurations satisfying the gauge condition C. Thus, gauge-invariant expectation values are the same, whether they are calculated over the whole gauge field configuration space or whether over a restricted gauge-field configuration space, provided the Faddeev-Popov determinant is included to compensate for the geometric structure of the restricted space. Thus, on the level of gauge-invariant quantities, all expressions describe the same theory. The normalization constants differ, but can always chosen such that $\langle 1 \rangle = 1$.

However, for gauge-dependent observables the expressions are not equivalent, for the arguments given above. In the first two lines, they vanish, but not in the third and fourth lines. Thus, from a purely mathematical point of view, these theories are distinct. From the point of physics, there is just an infinite number of equivalent quantum theories, the one without gauge fixing and the infinitely many choices of Ω_c, which all give the same observables.

Alternatively, this can also be taken to imply that any choice of theory with the action $S' = S - i \ln \Delta_c$ gives the same observable quantities, provided they are integrated over the corresponding configuration set Ω_c, either directly implemented as integration range or by a δ-function. In either way, this leads ultimately to the expression (7.16). Note that gauges like the linear-covariant gauges then can be considered to be averages over theories with different Ω_c.

This infinite degeneracy of quantum theories is a consequence of working with redundant variables. If it would be technically possible to go to generalized variables, just one theory would remain. Hence, the degeneracy should rather be considered to be a choice of suitable variables for technical purposes than any physical meaning. At any rate, for the purpose of this lecture, expressions like (7.16) will be used as the definition of the quantum theory, while, e. g., lattice calculations rather start at (7.19). But, as was just shown, both yield eventually the same results for observables.

7.3 Implications of gauge freedom

Are now the photons physical, or what is their status? The answer is, once more, subtle. Choose, e. g., Feynman gauge, i. e. $\xi = 1$. The corresponding propagator is then given by Gaussian integration as

$$\langle A^\dagger_\mu(x) A_\nu(y)\rangle = g_{\mu\nu}\int \frac{d^4p}{(2\pi)^4}\frac{e^{ip(x-y)}}{p^2+i\epsilon} = -g_{\mu\nu}\int \frac{d^3p}{2(2\pi)^3|\vec{p}|}e^{-ip_i(x-y)_i}.$$

The norm of a state

$$\Psi(x) = \int d^4x f(x) A_0(x)\,|0\rangle = \int \frac{d^4x d^4p}{(2\pi)^4} e^{ip_0x_0 - p_ix_i} f(p) A_0(x)\,|0\rangle\,,$$

with $f(x)$ an arbitrary weight function, created from the vacuum by the operator A_μ then reads

$$|\Psi|^2 = \int d^4x \int d^4y \langle A^\dagger_0(x) A_0(y)\rangle f^\dagger(x) f(x) = -\int \frac{d^3p}{2|\vec{p}|} f^\dagger(p) f(p) < 0.$$

Hence, there are negative (and zero) norm states present in the state space. Especially, the photon field has negative norms, and the corresponding space cannot be a Hilbert space.

The reason for this, is, of course, that A_0 itself is not gauge-invariant, and thus does not need to obey the usual properties expected for a physical observable. But it is possible to remedy the situation. As noted, the physical degrees of freedom can only be two, which can be chosen to be the transverse polarized ones. Consider therefore to contract the gauge field with a transverse projector. This this yields for the gauge-transformed field in momentum space

$$\left(g_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}\right) A^{\mu'} = \left(g_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}\right)(A^\mu + p^\mu g(p)) = \left(g_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}\right) A^\mu.$$

Thus, the transverse projected gauge-field is invariant under a gauge transformation. The longitudinal part behaves like

$$\frac{p_\mu p_\nu}{p^2} A^{\mu'} = \frac{p_\mu p_\nu}{p^2} A^\mu + \frac{p_\nu}{p^2} g(p)$$

and thus carries all the changes, but does also include (longitudinal) information from the gauge field.

To get rid of this part, define a physical photon field A^P_μ as

$$A^P_\mu = \left(g_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}\right) A^\nu,$$

which is manifestly transverse, and thus has only two degrees of freedom, and is invariant under gauge transformations. Seems to be pretty simple, but there is a catch.

This catch appears when trying to invert the Fourier transform, and return to position space. This yields

$$\int d^4p A^P_\mu(p) e^{ipx} = \int d^4p \left(g_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}\right) A^\nu(p) e^{ipx}.$$

Now, this is no longer a simple expression, because of the factor $1/p^2$. Rather, what is needed is a an integral over all of space of $A_\mu(x)$,

$$\begin{aligned} A^P_\mu(x) &= g_{\mu\nu} \int d^4y \delta(x-y) A^\nu(y) - \frac{1}{4\pi} \int d^4y \frac{\partial_\mu \partial_\nu}{|x-y|} A^\nu(y) \\ &= g_{\mu\nu} A^\nu(x) - \frac{1}{4\pi} \int d^4y \frac{\partial_\mu \partial_\nu}{|x-y|} A^\nu(y). \end{aligned} \tag{7.20}$$

Because

$$\begin{aligned} &\left(g_{\mu\nu} \int d^4y \delta(x-y) - \frac{1}{4\pi} \int d^4y \frac{\partial_\mu \partial_\nu}{|x-y|}\right) \partial^\nu g(y) \\ = \; &\partial_\mu g(x) - \frac{1}{4\pi} \int d^4y \frac{\partial^2 \partial_\mu g(y)}{|x-y|} = \partial_\mu g - \partial_\mu g = 0, \end{aligned} \tag{7.21}$$

the so defined field is indeed gauge-invariant. However, it is not a local object anymore, as the second integral is not reduced. Thus passing to a physical vector field implies even in QED giving up locality. Note that the physical field is entirely transverse, as any other contraction of it vanishes by construction. Thus, it carries only two degrees of freedom (in four dimensions). In fact, as the non-local term is essentially a solution to the partial differential equations determining the vector potential at fixed electric and magnetic field, it is possible to regain an expression in terms of these fields.

If, as in perturbation theory as a small-field expansion it is required, the integrated amplitude is small, the field is well localized and has also locally a small amplitude. Then the gauge field can be approximated by a δ-function, for which the integral can be solved, yielding just the local field. Hence, the physical field becomes the unphysical field in the small amplitude limit. Thus, it is possible to still use the gauge field to denote external states in perturbation theory.

The projection to physical states by a suitable operator is known as the Gupta-Bleuler condition in canonical quantization. Matter and the electric charge need to be constructed similarly, after adding it in the next section. This will not be detailed here.

7.4 Adding matter

To add matter, it remains to construct the electric current. As the current carries a conserved charge, this implies the existence of a continous symmetry by Noether's theorem.

Based on (7.1), this can be identified to be a U(1) symmetry, as the generator is the unit operator. However, it is a local symmetry, in contrast to the global symmetries encountered so far, as its parameters are functions of space-time. This has many further far-reaching consequences, which are the subject of quantum field theory II, and will be largely, and tacitly, ignored here. Due to the Abelian nature of U(1), QED is called an Abelian gauge theory.

As a U(1) symmetry, this implies the existence of a (local) transformation of a charged field ϕ

$$\phi \to \phi^g = e^{i\omega}\phi \tag{7.22}$$

and thus consequently the existence of an anti-particle. The conserved current can then be constructed from (6.3).

Consider an electron represented by a Dirac spinor Ψ. Under a gauge transformation (7.22) the spinors change as

$$\Psi \to \exp(-ie\omega)\Psi,$$

where the same function ω appears as for the vector potential in (7.1). For latter convenience, a constant e has been factored out, which will turn out to be the electric charge. Since ω is a function, the kinetic term for an electron is no longer invariant under a gauge transformation, and has to be replaced by

$$i\bar{\Psi}(\gamma^\mu(\partial_\mu + ieA_\mu))\Psi.$$

This replacement

$$\partial_\mu \to \partial_\mu + ieA_\mu = D_\mu$$

is called minimal coupling[4], and D_μ the covariant derivative. This is gauge invariant, as an explicit calculation shows,

$$\begin{aligned} i\bar{\Psi}'(\gamma^\mu(\partial_\mu + ieA'_\mu))\Psi' &= i\bar{\Psi}\exp(ie\omega)\gamma^\mu(\partial_\mu(\exp(-ie\omega)\Psi) + \exp(-ie\omega)(ieA_\mu\Psi + ie\partial_\mu\omega\Psi)) \\ &= i\bar{\Psi}\exp(ie\omega)\gamma^\mu(\exp(-ie\omega)(\partial_\mu\Psi - ie\partial_\mu\omega\Psi) \\ &\quad + \exp(-ie\omega)(ieA_\mu\Psi + ie\partial_\mu\omega\Psi)) \\ &= i\bar{\Psi}(\gamma^\mu(\partial_\mu + ieA_\mu))\Psi. \end{aligned}$$

Thus, the (gauge-invariant) Lagrangian of QED is given by

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\Psi}(i\gamma^\mu D_\mu - m)\Psi,$$

[4]It is possible to also formulate consistent theories with non-minimal coupling. However, none of these have so far been compatible with experiment, and they will therefore be ignored here.

where a mass term has been added, which is trivially gauge-invariant. By comparison, the term $-e\overline{\Psi}\gamma_\mu\Psi$ is thus the electric current J_μ in (7.4). This incidentally shows that $\overline{\Psi}\gamma_\mu\Psi$ transforms as a vector under Lorentz transformations. It is visible that $-e$ can then be interpreted as the charge carried by the field Ψ. By convention, for electrons e is taken to be positive, and thus the current is negative. Note that the electric current is by construction gauge-invariant, and thus observable. But it is also not a single field , but a composite expression made from two fermion fields.

When adding further fields, e. g. a proton field to be able to describe a hydrogen atom, the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \sum_f \bar{\Psi}(i\gamma^\mu(\partial_\mu + ie_f A_\mu) - m_f)\Psi, \tag{7.23}$$

where the e_f can be arbitrary for each fermion species, as long as the same e_f appears for them in the covariant derivative and the gauge transformation (7.22). It is noteworthy that QED therefore does not imply that, e. g., a hydrogen atom is electrically neutral.

7.5 Feynman rules

They Feynman rules for QED have the conventional fermion (or other matter) propagator. The propagator for the photon field is given by (7.17). The fermion-matter vertex is obtained in the usual way, and given by

$$\Gamma_\mu^{A\Psi\overline{\Psi}} = -ie_f\gamma_\mu. \tag{7.24}$$

No further additions are necessary.

The vertex, as the fermion propagator, does not depend on the gauge parameter ξ, in contrast to photon propagator. This implies that somehow the dependence on ξ needs eventually to drop out in observables. How this happens will be shown in an explicit calculation in the section 7.6.1.

In addition, as in case of the Yuakaw theory in section 5.4, the LSZ construction requires to introduce wave-functions for external photon lines. These are also called polarization vectors ϵ_μ, given their connection to the polarization of electromagnetic fields. Since these are solutions to the corresponding equations of motion, the Maxwell equations, there are two solutions $\epsilon_\mu^\pm$, corresponding to the two polarization directions. Furthermore, as they represent on-shell particles, the wave-functions need to respect for massless photon $p^\mu\epsilon_\mu^\pm = 0$. If they are to be transverse, they can then be written as

$$\epsilon_\mu^\pm = \begin{pmatrix} 0 \\ \vec{\epsilon}^\pm \end{pmatrix},$$

where $\vec{\epsilon}^{\pm}$ is a normalized unit vector satisfying $\vec{p}\vec{\epsilon}^{\pm} = 0$. E. g., for a momentum in the z direction a suitable choice is

$$\epsilon_{\mu}^{\pm} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ \pm i \\ 0 \end{pmatrix} \tag{7.25}$$

which correspond to right-handed and left-handed circular polarization, respectively, due to their phase shifts. Having ± 1 instead of $\pm i$ gives transverse polarization. Which of both is the correct choice depends on the experimental setup.

In general, if the initial state polarization is not determined, it is necessary to average over all possibilities. Likewise, if the final polarization is not measured, it is necessary to sum over all possible polarizations of the final state, by the rules of quantum mechanics.

7.6 QED processes

7.6.1 The electric charge

Considering a theory as well known classically as QED, the immediate question arises, how the concepts emerge from the full treatment. One of the obvious concepts is the electric charge. In section 7.4 the coupling constant e was introduced, and called already electric charge. It remains to be shown, what justifies such a statement.

For this, it is necessary to ask how the electric charge is actually classically defined. That is usually done in terms of the Coulomb force, i. e. the force which a particle experiences when acted upon by an electromagnetic field. Since the electromagnetic field in QED is described by the photon field A_{μ}, this translates to the question how a particle interacts with the photon field. A possible way is to discuss the absorption of a photon, i. e. the process $\Psi A \to \Psi$, in which kinetic energy will be transferred. This respects all conservation laws, and is thus a valid process.

As QED is time-reversal invariant, this can also be considered as the reverse process, in which the electron emits a photon. This can then be treated using the procedures for a decay of section 3.2,

$$\frac{d\Gamma^{\Psi \to \Psi A}}{d\Omega} = \frac{1}{32\pi^2 m} \frac{|\vec{q}|}{m} |\mathcal{M}_{\Psi \to \Psi A}(p,q)|^2 \theta(m - m_1 - m_2).$$

This requires to determine the matrix element. At tree-level the amputated matrix element is just the electron-photon vertex (7.24). However, in contrast to the case in section 3.3.1, the external wave-functions are not trivial. Thus, the total matrix element is, when not

measuring spin

$$\begin{aligned}\mathcal{M}_{\Psi\to\Psi A} &= -\frac{ie_e}{2}\sum_{\pm}\epsilon^i_\mu(q)\bar{u}(p)\gamma^\mu u(k) = -\frac{ie_e}{2}\sum_{s=\pm}\epsilon^s_\mu(q)\bar{u}_i(p)\left(\delta_{ij}\frac{k^\mu+p^\mu}{2m_e}+\frac{i\Gamma^{\mu\nu}_{ij}q_\nu}{2m_e}\right)u_j(k)\\ \Gamma_{\mu\nu} &= \frac{i}{2}\left[\gamma_\mu,\gamma_\nu\right],\end{aligned}$$

where use was made of features of the spinors and Dirac equation, as well as of the Dirac matrices. Note that the spinors are orthonormalized.

While this case can now be treated in full generality, a very interesting limit is that of vanishing photon momentum $q \to 0$, with the photon being on-shell $q^2 = 0$. In that case the expression $e\bar{u}(p)\gamma^\mu u(p)$ appears, which can be recognized as to be the electric current coupling to A_μ in the QED Lagrangian (7.23). Thus, this vertex really is the manifestation of describing the classical interaction of the electromagnetic current with the electromagnetic field. This expression will only vanish if the electron's momentum has a component along the direction of the electromagnetic field, e. g. $p = (E, p_1, 0, 0)$. The Dirac equation then yields

$$\mathcal{M}_{\Psi\to\Psi A} = -ie\sum_{s=\pm}\epsilon^s_\mu(q)p^\mu = -\sqrt{32}e\frac{p_1\sqrt{m^2+p_1^2}}{m^2}$$

which finally yields, after squaring,

$$\frac{d\Gamma^{\Psi\to\Psi A}}{d\Omega} = \frac{e^2}{\pi^2}\frac{p_1^3(m^2+p_1^2)}{m^4} = 4\pi\alpha\frac{p_1^3(m^2+p_1^2)}{m^4}$$

in which the fine structure constant $\alpha = e^2/(4\pi)$ was defined, which is a characteristic combination appearing throughout. This is the Thomson-limit, and it most closely realizes a classical limit.

It is also interesting to consider the case, that the relevant momenta are all large enough to neglect the electron mass, and have equal sized components, and thus the same four-momentum squared. The problem then has only a single scale, given by the absolute-valued squared of this component Q^2. The momentum then satisfy $q + p + k = 0$ and $p^2 = q^2 = k^2$, the so-called symmetric point. The matrix element is then a function of this single quantity, but it is possible to factor out the value of e still,

$$\mathcal{M}_{\Psi\to\Psi A} = -ief(Q^2).$$

As the limit of the zero energy photon had been used to define the electric charge, the quantity $ef(Q^2)$ can be considered to be the electric charged when probed at the energy scale given by Q^2. If $f(Q^2)$ is non-constant, this implies that the effective electric charge

changes with energy, it is said to be a running coupling. This is then a characteristic feature of the theory. Especially, the quantity $\alpha(Q^2) = (ef(Q))^2/(4\pi)$ reappears usually in many places, and thus many quantities can be expressed in terms of the running coupling. measuring cross sections allow then to infer the behavior of the running coupling.

However, there was nothing special about choosing this quantity to define the quantum field theory version of the charge, nor of the running one. The only motivation was to have an experiment with a classical counterpart. Thus, while the running charge can be define din this way, it does not need to be. This issue will be taken up again in section 8.3.

7.6.2 Compton scattering

Another process, which is familiar from classical physics, is the scattering of two different charged particles of each other. This is, e. g., Compton scattering of an electron off a proton.

Because both particles are distinct, there is only a single diagram, the exchange of a photon in the t-channel,

$$\mathcal{M}^{ep\to ep} = \frac{e^2}{4}\sum_{\text{spins}} \bar{u}^e(p_1)\gamma^\mu u^e(q_1) D_{\mu\nu}(k) \bar{u}^p(p_2)\gamma^\nu u^p(q_2)$$

By comparison, this show that the process can be interpreted as two electric current of opposite charges interacting by a photon exchange. While the calculation is simplified by using Feynman gauge, it is instructive to work in an arbitrary covariant gauge, and thus using the propagator (7.17)

$$D_{\mu\nu}(k) = \left(g_{\mu\nu} - \left(1 - \frac{1}{\xi}\right) k_\mu k_\nu\right) \frac{1}{k^2}$$

to see how the ξ-dependence is removed in the measurable final cross section. Since the process is t-channel, this yields

$$\begin{aligned}\mathcal{M}^{ep\to ep} &= \frac{e^2}{4t}\sum_{\text{spins}} \Bigg(\bar{u}^e(p_1)\gamma^\mu u^e(q_1)\bar{u}^p(p_2)\gamma_\mu u^p(q_2) \\ &\quad - \left(1 - \frac{1}{\xi}\right)\bar{u}^e(p_1)k_\mu\gamma^\mu u^e(q_1)\bar{u}^p(p_2)k_\nu\gamma^\nu u^p(q_2)\Bigg).\end{aligned}$$

Now it is true that

$$k = p_1 - q_1 = p_2 - q_2.$$

Thus, each of the generate four terms contains a contribution $p_\mu\gamma^\mu u(p)$, or the corresponding conjugated. Since the spinors fulfill the corresponding Dirac equation, it follows that $p_\mu\gamma^\mu u(p) = mp$.

$$\bar{u}^e(p_1)k_\mu\gamma^\mu u^e(q_1) = m\bar{u}^e(p_1)u^e(q_1) - m\bar{u}^e(p_1)u^e(q_1) = 0,$$

and likewise for the proton current. Thus, the term containing the ξ-dependence vanishes exactly on the level of the matrix element, yielding a gauge-parameter-independent[5] result, and thus is physical.

It remains to calculate the first term. Switching to the center-of-mass frame (3.20) with scattering angle θ and 3-momentum size p, the momentum transfer is

$$t = (0, 0, p\cos\theta, p(1-\sin\theta))^2 = -2p^2(\sin\theta - 1) = Q^2. \tag{7.26}$$

and thus space-like. The latter is typical for such exchanges, and the label Q^2 is often used as well. This yields ultimately

$$\begin{aligned}
\frac{d\sigma}{d\theta} &= 8\pi^2\alpha^2\frac{\left((2m_e^2+q^2)(2m_p^2+q^2) - 2q^2E_eE_p - q^4(q^2\cos\theta + (m_e^2+m_p^2+(E_e-E_p)^2\sin\theta\right)^4}{m_e^2m_p^2(E_e+E_p)^2\left(m_e^2+m_p^2-2E_eE_p+2q^2\sin\theta\right)^2} \\
E_e &= \sqrt{m_e^2+q^2} \\
E_p &= \sqrt{m_p^2+q^2}
\end{aligned} \tag{7.27}$$

which is a highly in transparent result. The reason is the complicated phase-space structure with the two different masses.

It thus becomes more transparent in various limits. The first is the very low-energy one, in which the particles are taken to be effectively infinitely heavy, and thus both electron and proton essentially static. This yields

$$\frac{d\sigma}{d\theta} = \frac{8\pi^2\alpha^2m_em_p}{q^4(\sin\theta-1)^2}. \tag{7.28}$$

which has the typical divergence for forward scattering, which stems from a neglect of the extension of the projectiles. It also shows a typical decay with $1/q^4$, or actually $1/t^2$. It thus behaves like just the photon propagator. Physically, this makes a lot of sense, as the only kinematics in this case comes from the photon, which is basically exchange between unmovable, static sources.

[5] It can be shown that the result is even gauge-independent.

Likewise, the ultrarelativistic limit is happening when both masses can be neglected, yielding

$$\begin{aligned}\frac{d\sigma}{d\theta} &= \frac{4\pi^2\alpha^2 q^2\cos^4\theta}{2m_e^2m_p^2(\sin\theta-1)^2}\\ \sigma &= \frac{12\pi^3q^2\alpha^2}{4m_e^2m_p^2}.\end{aligned}$$

In that case, there is no forward singularity, and the cross section increase with exchange energy. That there is no forward divergency is because massless particles cannot be localized. That the cross section increases is partly due to the nature of QED, as discussed in quantum field theory II, but will also addressed again in section 8.5.

Finally, there is interesting intermediate range, where the fact that the electron is much lighter than the proton allows to neglect the electron mass, and send the proton mass to infinity. This yields

$$\begin{aligned}\frac{d\sigma}{d\omega} &= \frac{16\pi^2\alpha^2q^4(3-\cos(2\theta)+4\sin\theta)}{m_p^4m_e^2}\\ \sigma &= \frac{48\pi^3\alpha^2q^4}{m_p^4m_e^2}.\end{aligned}$$

In this case, there is a strong asymmetry in the scattering angle, typical for vector interactions between particles with spin.

7.6.3 The potential

There is, however, a way how to make expression (7.28) even more useful. Noting that (7.28) scales with the momentum transfer (7.26) like $1/Q^4$, or $(1/Q^2)^2$ leads to a remarkable insight. $(1/Q^2)$ is the Fourier-transform of $1/r$. Thus, the cross section scales like $(\alpha/r)^2$, and the matrix element behaves like α/r. Since the matrix element has been seen to be the coupling of two electromagnetic currents by $g_{\mu\nu}\alpha/Q^2$ this shows that at tree-level the interactions between charges in QED is indeed the classical electromagnetic interactions of particles. These are considered as test masses, i. e. not interacting on the field, as in the classical case of calculating the Coulomb force. Only if the masses, and thus phase space, is explicitly taken into account in the full expression (7.27), this becomes obscured.

Likewise, if one would do a similar analysis in the case of Yuakaw theory of section 5.3, the corresponding potential would be found to be

$$g^2e^{-m_\phi r}/r \tag{7.29}$$

for the exchange of the scalar particles between fermionic currents. Only between scalar currents, the appearance of further terms like the genuine four-point interaction, spoils this classical analogue, as does the exchange graphs.

This already shows that the interpretation in terms of a potential does not hold for all theories. In fact, it works only for those theories, where currents can interact by a single-particle exchange within a single channel. However, classical electrodynamics, as well as the leading part of nuclear interactions, belong to this class. Other theories, e. g. as noted the linear σ-model, do not provide such a simple interpretation.

Finally, as will be seen in section 8.6, higher-order corrections, and thus quantum effects, will spoil the analogy. Moreover, even the interpretation as a potential becomes untenable as soon as inelastic processes become possible. This will again be taken up in section 10.3.

7.7 Beyond vectors

It appears now straightforward to extend the concept further to ever higher-spin elementary particles. Of course, the same issues as with vectors will arise with higher spin. Especially, any representations in terms of 4-tensors will suffer from the same problem, and necessitate the introduction of gauge fields.

Consider the next case of spin 3/2. It turns out that this is easiest achieved by considering a direct sum $(1, 1/2) \oplus (1/2, 1)$ representation, which thus contains both a left-handed part and a right-handed part. This is mainly due to the CPT theorem, and the fact how left-handed and right-handed map into each other. The so obtained field is call ed a Rarita-Schwinger field Ψ^i_μ, which is a vector with vector components μ being (Dirac-)spinor-valued fields with spinor components i.

In this way, the Rarita-Schwinger field is behaving similarly as the gauge field of QED, and transforms under a local (Abelian) gauge transformation as

$$\Psi^i_\mu \to \Psi^i_\mu + \partial_\mu \epsilon^i$$

where ϵ is, however, now a spinor-valued function. This structure was to be expected, as this couples effectively a spin 1 and a spin 1/2 object to create spin 3/2.

Since the transformation is linear, it is an Abelian gauge theory, and the corresponding field strength tensor

$$\Omega^i_{\mu\nu} = \partial_\mu \Psi^i_\nu - \partial_\nu \Psi^i_\mu$$

is therefore gauge invariant, but carries also a spinor index i.

It is still necessary to postulate a Lagrangian for the theory, which is gauge-invariant. Introducing $\bar{\Psi} = \Psi^\dagger \gamma_0$, a possibility is

$$\begin{aligned} \mathcal{L} &= -\bar{\Psi}_\mu \gamma^{\mu\nu\rho} \partial_\nu \Psi_\rho . \\ \gamma^{\mu\nu\rho} &= \frac{1}{2i} \{\gamma^\mu, \Gamma^{\rho\sigma}\} \end{aligned}$$

As for the Maxwell case, there are no gauge-invariant, perturbatively renormalizable further interaction terms possible. Without interactions, only non-interacting Rarita-Schwinger fields are possible. The equation of motion is, similar to the Dirac equation,

$$\gamma^{\mu\nu\rho} \partial_\nu \Psi_\rho = 0.$$

It follows that the Rarita-Schwinger field can have (classically) physical modes only for $d > 3$, similar like the vector potential only for $d > 2$. This equation of motion also implies

$$\gamma^\mu \Psi_{\mu\nu} = 0,$$

which is Rarita-Schwinger form of the homogeneous Maxwell equations. The equations of motions can be solved in a similar way as the free Dirac equation, and creates the free-field solutions. It is possible to add a mass term, yielding

$$\mathcal{L} = -\bar{\Psi}_\mu (\gamma^{\mu\nu\rho} \partial_\nu - m \gamma^{\mu\rho}) \Psi_\rho ,$$

in contrast to the vector gauge fields.

The primary problem is that, in contrast to the Maxwell case, no interaction terms can be found, which are at most power four. Hence, implementing this restriction, there is no interacting theory with Rarita-Schwinger field available. However, effective theories, e. g. of hyperfine-interactions in atoms or of nuclei, do exist, but they cannot be fundamental.

The situation persist to higher spins. At spin 2, there are two choices, symmetric tensor fields or anti-symmetric tensor fields, both of which need to be gauge fields. But here also the nature of the gauge symmetry changes. In the Maxwell case, the coupling was to a vector current. This was sufficient to find a Poincaré-invariant interaction term, (7.4). A similar construction for a tensor field requires a tensor current. However, by Schur's lemma, there is only one Abelian tensor current, the energy-momentum current. But this implies that a tensor field will automatically gauge energy-momentum. This is then necessarily a (quantum) gravity theory, like general relativity. Thus, flat-space quantum field theory seems to be only possible with elementary particles up to spin 3/2, if any meaningful theory, even a free one, should exist. Of course, composite particles of higher spin are possible, as will be discussed in section (9.2). But for elementary particles, any further steps will be relegated to the lecture on advanced general relativity and quantum gravity.

Chapter 8

Beyond tree-level perturbation theory

Tree-level perturbation theory is often giving already the right idea. But as a first-order in quantum effects, it cannot bee expected to yield the full result. E. g., as has been seen in section 7.6.2, tree-level generates for QED only the Coulomb potential. Thus, deviations like the observed Lamb-shift could not be explained at tree-level. It is hence necessary to go beyond tree-level.

8.1 The scalar propagator at one-loop and regularization

However, beyond tree-level, many new issues arise, mostly connected to the appearance of loops. A full treatment will only be possible in the lecture quantum field theory II lecture, but already now it is possible to give the basic ideas. In fact, they can be illuminated with one of the most simple theories and quantities already, the propagator of the scalar in the Yukawa model at order λ.

To this process only a single diagram contributes to the 1PI matrix element the self-energy $\Pi_\lambda^{\phi\phi}$. This diagram is a loop attached at the propagator, a so-called tadpole diagram. Its value is, up to $\mathcal{O}(\lambda^2)$,

$$\Pi_\lambda^{\phi\phi} = -\frac{\lambda}{2}\int \frac{d^4p}{(2\pi)^4}\frac{1}{p^2 - M^2 + i\epsilon}, \tag{8.1}$$

where the factor $1/2$ is a symmetry factor. The integration over p_0 can be performed first

by contour-integration and using the Cauchy theorem, since

$$\begin{aligned}\Pi_\lambda^{\phi\phi} &= -\frac{\lambda}{2}\int\frac{d^3\vec{p}}{(2\pi)^4}\int dp_0\frac{1}{p_0^2-\vec{p}^2-M^2+i\epsilon}\\ &= -\frac{\lambda}{2}\int\frac{d^3p}{(2\pi)^4}\int_{-\infty}^{\infty}dp_0\frac{1}{(p_0+\sqrt{\vec{p}^2+M^2})(p_0-\sqrt{\vec{p}^2+M^2})+i\epsilon}\end{aligned}$$

This has a pole in the upper half-plane, and vanishes sufficiently fast on a half-circle at infinity. The residue at the simple poles $p_0 = \pm\sqrt{\vec{p}^2+M^2}$ is $1/(p_0 \mp \sqrt{\vec{p}^2+M^2})$, dropping the small contribution of $i\epsilon$, which only served to not have the pole on the axis. The Cauchy theorem then yields, using polar coordinates in the final expression,

$$\Pi_\lambda^{\phi\phi} = \frac{i\lambda}{4\pi^2}\int_0^\infty \vec{p}^2 d|\vec{p}|\frac{1}{\sqrt{\vec{p}^2+M^2}}.$$

This expression shows the generic problem: It is infinite, when integrated, with the infinity originating from the upper limit of the integral, as so-called ultraviolet divergency[1]. This is an issue, which arises for many, though not all, loop integrals.

At first sight, this result may seem to be a catastrophe, and to invalidate the theory. However, it turns out that it is possible to make the integrals convergent without introducing additional parameters into the theory, albeit at the price that the theory still looses its validity at some high energy scale. Since this scale can be pushed to practically very high energies, this is of little practical importance, as it can anyway not be assumed that conventional quantum field theories could be theories of everything, since they do not include gravity. And, as discussed in section 7.7, this anyhow entails to leave the arena of flat-space quantum field theory. Therefore, theories with such a feature are considered today to be 'low-energy' effective theories. It is generally assumed that a hierarchy of such effective theories will continuously approach the point where quantum gravity takes over.

Still, even if this is correct, it is required to make sense out of such a theory. This implies to deal with the appearing divergency. This requires two basic steps. One is a prescription how to regularize integrals, i. e., how to map their divergent value to a finite value. For this purpose of regularization the integrals are made convergent by the introduction of some parameter, and the original divergence is recovered when sending this parameter to a particular limit. As a result, all quantities calculated will depend on this parameter.

The second step, the so-called renormalization, gives a prescription how to redefine the theory such as to loose the dependence on this extra parameter, the so-called renormaliza-

[1]There can also be infrared divergencies, but they are mostly due to massless particles. They are therefore due to a completely different origin. They can be systematically treated. As they mainly, but not exclusively, appear in gauge theories, this issue is relegated to quantum field theory II.

tion scheme, without recovering the original divergence. The consequence of this program, and the particular renormalization scheme used, is that quantities like masses or coupling constants can no longer be interpreted as static quantities, but will depend on the scale at which they are measured. It is said that they become running. However, measurable quantities, like a cross-section, turn out not to depend on the measurement scale, at least for an exact calculation. Unfortunately, most calculations are not exact in general, and in particular for the standard model. As a consequence, a dependence on the scale may be left.

These steps will now be implemented in turn in this and the following section. The first step is regularization. The technically simplest, but by no means only, possibility in the present context is the cutoff regularization. It works by replacing the upper integral limit ∞ by a large, but finite number Λ, the so-called cutoff.

The integral can then be calculated explicitly to yield

$$\Pi_{\lambda}^{\phi\phi} = \frac{i\lambda}{4\pi^2}\left(\Lambda^2\sqrt{1+\frac{M^2}{\Lambda^2}} - M^2\ln\left(\frac{\Lambda+\Lambda\sqrt{1+\frac{M^2}{\Lambda^2}}}{M}\right)\right). \tag{8.2}$$

As can be seen, the integral diverges with the cutoff Λ quadratically, and has in addition a sub-leading divergence logarithmically in Λ. Still, as long as the limit is not performed, the result is finite, independent of the momentum, but explicitly dependent on Λ. The maximum power of Λ appearing is called the degree of divergency of the diagram, in thus case quadratically divergent. An upper limit of the divergency, the so-called superficial degree of divergency, can be obtained by the formula

$$d + \left(n\frac{d-2}{2} - d\right)v - \frac{d-2}{2}n,$$

where d is the space-time dimension, n is the highest power of fields appearing in the Lagrangian, v is the number of vertices in the diagram, and n is the number of external lines. E. g., here this yields $4+(4-4)1-2=2$, as required. This formula can be derived by just counting powers of momenta in the diagrams, which will not be done here. Note that the actual divergence, e. g. due to cancellations, can be lower.

It is valid to ask, whether this could not have been done already for the four-dimensional integral (8.1). The answer is both yes and no. The expression (8.1) is still in Minkowski space-time. Thus, the concept of a large energy scale is much harder to define, because even for every component of a four-momentum large, the four-momentum square can be zero. To still work with the explicit four-dimensional quantity would thus require an alternative approach, to make sense of what a high-energy cutoff means. The most straightforward one would be to analytically continue $p_0 \to ip_0$ (and also all other momenta), and thus

effectively from Minkowski space-time to Euclidean space-time. Since all possible perturbative integrals are indeed analytic functions in the decisive quarter-plane, this would be possible. Then, an Euclidean cutoff could be introduced. In the present case, there is no external momentum, but if there was, this would be needed to be followed by an analytic continuation back into Minkowski space-time.

Both approaches are valid regularization procedures. However, at the level of expressions like (8.2), they yield different results. It is the task of renormalization to ensure that both yield the same measurable consequences. If this is possible, the theory is said to be renormalizable. If not, the theory is called non-renormalizable. In the latter case, the choice of regularization scheme influences the outcome. Usually, the effects will be suppressed like energy scale/Λ, but that can still be relevant. Such theories are then genuine low-energy effective theories.

Within the context of perturbation theory, it can be shown that any theory, which has positive-energy or dimensionless coupling constants only is renormalizable. This is the origin of the condition to have only such coupling constants in theories, as was introduced in section 2.2. Beyond perturbation theory, anything may be possible.

It should be noted that theories exists, in which either divergencies cancel always between diagrams order-by-order, or all diagrams are non-divergent. Such theories are called super-renormalizable and finite, respectively. They become more common the lower the dimensionality, but are in four dimensions the very rare exception.

8.2 Renormalization and counter terms

To remove the dependence of the regularization, it is worthwhile to investigate the total structure of the two-point function $\Gamma^{\phi\phi}$, which is given to this order by,

$$\Gamma^{\phi\phi}_{\lambda} = p^2 - M^2 + \Pi^{\phi\phi}_{\lambda} + \mathcal{O}(\lambda^2, y) \tag{8.3}$$

As is seen from the result (8.2), the contribution $\Pi^{\phi\phi}$ is momentum-independent and dependent on the cutoff Λ. If it would be finite, it could be interpreted as a change of the mass M, since then the expression would have the form

$$p^2 - M^2 - \delta M^2 \to p^2 - M_R^2$$

with the so defined renormalized mass

$$M_R = \sqrt{M^2 + \delta M^2}.$$

The actual mass of a ϕ particle, which would be measured in an experiment, would then be M_R, instead of the so-called bare mass M. In fact, since the experimental measurement

is the only knowledge available on the theory, it is mandatory that the bare parameters of the theory, like the bare mass M, are adjusted such that the resulting renormalized mass M_R agrees with experiment[2].

Now, since the actual bare parameters cannot be measured, there is nothing which prevents us to set it to

$$M^2 = M_R^2 - \delta M^2,$$

with the experimental input M_R. This automatically fulfills the requirement to reproduce the experiment. In particular, since M is not an observable quantity, there is no reason for it to be finite, and independent of the cutoff Λ. Thus, it is possible to absorb the infinity of the divergent integral in unobservable bare parameters of the theory. This is the renormalization process: The absorption of the divergencies into unobservable quantities in the definition of the theory. In a way, M could be interpreted as the mass at the cutoff scale Λ, which is infinite. The actual mass measured at low energies, and thus in the sense of a low-energy effective theory, is the renormalized mass M_R.

This can be arranged already at the level of the Lagrangian by replacing

$$\frac{M}{2}^2 \phi^2 \to \frac{M_R}{2}^2 \phi^2 - \frac{\delta M^2}{2}\phi^2.$$

The second term is a so-called counter-term, and it depends on the actual order of the calculation. E. g., at tree-level, it would be zero. In this way, when performing a calculation, the results are already finite, and in agreement with the observed mass.

It is actually not the the only contribution which appears. If the calculation is extended to also include corrections up to $\mathcal{O}(\lambda, y^2)$, there is a second diagram contributing to the self-energy, which is due to a loop of the fermions. The expression then takes the form

$$\Pi^{\phi\phi}_{\lambda,y^2} = \Pi^{\phi\phi}_{\lambda} + \Pi^{\phi\phi}_{y^2},$$

with the fermionic contribution given by

$$\Pi^{\phi\phi}_{y^2} = -\frac{y^2}{2}\int \frac{d^4p}{(2\pi)^4}\frac{\mathrm{tr}((\gamma_\mu p^\mu + M)(\gamma_\nu(p^\nu - q^\nu) + M))}{(p^2 - M^2 + i\epsilon)((p-q)^2 - M^2 + i\epsilon)}.$$

Using the trace identities tr1 = 4, trγ_μ=0, and tr$\gamma_\mu\gamma_\nu = 4g_{\mu\nu}$ this simplifies to

$$-\frac{y^2}{2}\int \frac{d^4p}{(2\pi)^4}\frac{p(p-q) + M^2}{(p^2 - M^2 + i\epsilon)((p-q)^2 - M^2 + i\epsilon)}.$$

[2]This implies that the bare parameters have to be adapted at each order of perturbation theory calculated.

Since the numerator scales with p^2, the integral is quadratically divergent. Suppressing the $i\epsilon$, the expression can be rewritten by introducing a zero and then shifting the integration argument, as

$$\begin{aligned} & -\frac{y^2}{2}\int\frac{d^4p}{(2\pi)^4}\frac{(p^2-m^2)+((p-q)^2)-m^2)-q^2+4m^2}{(p^2-m^2)((p-q)^2-m^2)} \\ = & -\frac{y^2}{2}\int\frac{d^4p}{(2\pi)^4}\left(\frac{1}{(p-q)^2-m^2}+\frac{1}{p^2-m^2}+\frac{4m^2-q^2}{(p^2-m^2)((p-q)^2-m^2)}\right) \\ = & -\frac{y^2}{2}\int\frac{d^4p}{(2\pi)^4}\left(\frac{2}{p^2-m^2}+\frac{4m^2-q^2}{(p^2-m^2)((p-q)^2-m^2)}\right), \end{aligned}$$

Such integrals can be performed using a number of analytical tricks. However, for the present purpose this will not be necessary. It is sufficient to observe that the result, just by counting powers of integration momenta, will have the form

$$\Pi^{\phi\phi}_{y^2} = c_1\Lambda^2 + (c_2m^2 + c_3q^2)\ln\frac{\Lambda}{m} + f(m^2,q^2),$$

where f is some finite function when Λ is send to infinity, and depends on the momentum q and the coupling constant y, as do the constants c_i.

The first two terms have again the same structure as the tadpole contribution (8.1). However, the third term is different, as it does depend explicitly on the momentum. Therefore, it cannot be absorbed into a mass renormalization. However, it can be absorbed in a renormalization of the kinetic term. If in the Lagrangian the modification

$$\partial_\mu\phi\partial^\mu\phi \to \partial_\mu\phi\partial^\mu\phi + \delta Z_\phi\partial_\mu\phi\partial^\mu\phi = Z_\phi\partial_\mu\phi\partial^\mu\phi,$$

is performed, the kinetic term of the field ϕ has been renormalized by a the wave-function renormalization $\sqrt{Z_\phi}$. Choosing

$$\delta Z_\phi = -c_3\ln\frac{\Lambda}{m},$$

this will remove the divergence. By this the field amplitude is modified to yield the correct special relativistic dispersion relation. Note that quantity also appears asymptotically, and thus is the relevant quantity in the context of the LSZ formalism in (3.25-3.26).

Performing further calculations, it turns out that similar changes have to be performed for the remaining bare parameters m, λ, and y, yielding a renormalized fermion mass m_R, and renormalized couplings λ_R and y_R. Thus, including these counter-terms yields the renormalized Lagrangian

$$\begin{aligned} \mathcal{L}_R = & \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \bar\chi i(\gamma^\mu\partial_\mu - m_R)\chi - \frac{{M_R}^2}{2}\phi^2 - \frac{\lambda_R}{4!}\phi^4 - y_R\phi\bar\chi\chi \\ & + \frac{\delta Z_\phi}{2}\partial_\mu\phi\partial^\mu\phi + \bar\chi i(\delta Z_\chi\gamma^\mu\partial_\mu - \delta m)\chi - \frac{\delta{M}^2}{2}\phi^2 - \frac{\delta\lambda}{4!}\phi^4 - \delta y\phi\bar\chi\chi. \end{aligned}$$

It should be noted that always certain products of fields appear together with a parameter of the theory. Thus, often explicit factors of various Zs are introduced such that not kinetic terms are renormalized, but rather the field itself, in the sense of an amplitude renormalization. In this case, explicit factors of $Z_i^{1/2}$ are multiplied for each field in the counter-term Lagrangian, and the counter-terms δM, δm, $\delta\lambda$, and δy are redefined by appropriate factors of $Z_i^{-1/2}$. This is, however, conventional, but the more common case.

Also, it is usual that δx is rather defined as

$$\delta x = Z_x x = (1 + \delta Z_x)x,$$

i. e. as a multiplicative factor to the original quantity. However, Z_x may then depend again on x, even in the form of $1/x$.

In case of QED, the same is also necessary, and leads to the renormalized Lagrangian

$$\begin{aligned}\mathcal{L}_{QED} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial^\mu A_\mu)^2 + \bar{\psi}(i\gamma_\mu + m)\psi + eA_\mu\bar{\psi}\gamma^\mu\psi \\ &\quad -\frac{\delta Z_A}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\frac{Z_A}{Z_\xi} - 1}{2\xi}(\partial_\mu A^\mu)^2 + \delta Z_\psi\bar{\psi}i\gamma^\mu\partial_\mu\psi - (Z_\psi Z_m - 1)m\bar{\psi}\psi \\ &\quad +(Z_e Z_A^{-\frac{1}{2}} Z_\psi - 1)eA_\mu\bar{\psi}\gamma^\mu\psi.\end{aligned}$$

In this case all parameters, m, e, and ξ, as well as A_μ and ψ have been multiplicatively renormalized.

The remaining question is then whether this is sufficient, or whether further terms, e. g. a sixth power of the fields, would be necessary, or whether non-multiplicative terms would appear, beyond the additive mass shift renormalization for the scalar in Yukawa theory. The answer depends on the theory. However, as long as only a finite number of independent counter terms are needed[3], the theory is renormalizable. Because then with a finite number of experimental inputs in form of numbers, all divergencies are absorbed, and all parameters fixed. Everything else then becomes a genuine prediction of the theory. Whether this number is smaller, the same, or larger than the number of parameters and fields at tree-level depends on the theory. It also depends on the dimensions of space-time, and theories, which are not renormalizable in higher dimensions may be so in lower dimensions.

If not, the theory is not renormalizable. However, in perturbation theory, the number of counter-terms may still be finite at any given order in perturbation theory, but may grow indefinitely with the order. In that case, some predictions are still possible, if uncertainties on the order of the higher-order corrections are acceptable, and the theory becomes an

[3] An infinite number of dependent counter terms is allowed, without altering the statement.

effective theory. In perturbation theory also all renormalization is found to be either additive or multiplicative.

Beyond perturbation theory, all of these statements can change in any direction qualitatively. However, in many, if not most, cases the perturbative result statements about renormalizability carry over to the full theory. There is one special case, which needs to be mentioned. There are some theories, which can be fully renormalized, but only if all renormalized couplings are set to zero, i. e., the renormalized theory is free. Such theories are called trivial. However, this also implies that when keeping the cutoff, they may still be interacting, and serve as effective theory. To all intent in purposes, the linear-σ model in four dimensions, but not in lower dimensions, seems to be of the class.

8.3 Renormalization schemes

So far, the counter terms have been identified by direct comparison. However, assume that the propagator has finally the form

$$D = \frac{c^2 - d^2 + 2p^2}{p^4 + (d^2 - c^2)p^2 - c^2 d^2}.$$

Such a propagator has no longer the form of a conventional free particle. It is thus not clear how to determine, e. g., δm, such that it represents the mass of a particle. Thus, it is necessary to give a more precise definition of what physical mass means. Since such a mass would be expected as a pole, one possibility would be to choose it as the smallest momentum at which the propagator has a simple pole. In this case, this would imply

$$m_R = d,$$

and thus the counter-terms can be arranged such that this equality holds. This is called a pole scheme.

It becomes much more ambiguous for the coupling constants, as they are not associated with some pole. For the electromagnetic charge, it still seems reasonable to choose its macroscopic value determined in the Thomson limit in section 7.6.1. Another possibility is, e. g., to choose[4]

$$\Gamma^{A\bar{\psi}\psi}(p, q, p+q) \stackrel{!}{=} e,$$

for two arbitrarily chosen momenta. This already shows that a certain ambiguity is introduced, because a scale μ is introduced, which is proportional to p^2, at a fixed ratio of p and q. This scale is called a renormalization scale. An often found choice is the

[4]The Thomson limit is this expression at $p = q = 0$.

symmetric point, at which all involved momenta have size $\mu = p^2$. This generalizes the ideas from section 7.6.1, and especially connects the idea of a running coupling with the renormalization process. It is even more ambiguous when it comes to identify conditions for the wave-function renormalization.

Of course, these choices are made by humans, and hence physics should not depend on it[5]. Therefore, any set of choices, which is sufficient to fix all independent counter terms, sets up a fixed scheme, a so-called renormalization scheme. Similar to a choice of gauge or a coordinate system, but for entirely different reasons, it defines a technical frame in which to work. And just as for coordinate systems, renormalization scheme transformation could be made to change between them, though they are technically more involved.

Thus, any choice will do. For QED, e. g., a popular scheme is

$$\begin{aligned}
\left(p^2 g_{\mu\nu} - p_\mu p_\nu\right)_{p^2=\mu^2} D^{\mu\nu}_{AA}(\mu^2) &= i && (8.4)\\
\mu^2 g_{\mu\nu} D^{\mu\nu}_{AA}(\mu^2) &= i\xi && (8.5)\\
\mathrm{tr} D_{\psi\psi}(\mu^2) &= i m(\mu^2) && (8.6)\\
(\mathrm{tr}\gamma_\mu p^\mu D_{\psi\psi})(\mu^2) &= i 16\mu^2 && (8.7)\\
\Gamma^{A\bar{\psi}\psi}(\mu^2,\mu^2,\mu^2) &= i e(\mu^2) && (8.8)
\end{aligned}$$

Note that there is no condition that involves a mass of the photon. The corresponding counter-term would violate gauge-invariance, as a term $\sim A_\mu^2$ is not gauge-invariant.

The appearing renormalization scale μ is arbitrary, but it cannot be removed. Of course, it would be possible to choose for each of the five conditions (8.4-8.8) a different scale, but these would then differ only by constant prefactors multiplying the single scale. Since this scale is arbitrary, nothing which is observable can depend on it. This observation is the basis for the so-called renormalization-group approach, which uses this knowledge and by forming derivatives on renormalization-scale-invariant quantities determines (functional) differential equations, which are useful for determining properties of correlation functions. This will be taken up in the lecture quantum field theory II.

There is a further consequence of this scale. If a theory like massless QED is taken, there appears no dimensionful parameter at the classical level, and the theory is classically scale-invariant. However, when the renormalization conditions are imposed, this is no longer the case, since they involve this scale. Since this scale is a manifestation of the ultraviolet divergences, and thus incompleteness of the theory, it is thus created in the quantization

[5]Actually, any quantity which depends on the renormalization scheme, and thus also the renormalization scale, can hence not be measured directly. The only direct measurements possible are in fact either cross sections, masses, or decay rates in one form or the other, which are all independent of the scheme in an exact calculation.

process. It is thus said that the classical scale invariance is broken by quantum effects, a process also referred to as dimensional transmutation. In a sense the quantization process itself is breaking the classical scale symmetry[6], a so-called anomaly. This perspective will be further in the lecture quantum field theory II.

A second feature is that the mass of the electron and the electric charge now depend on the renormalization, and thus energy scale, by virtue of the renormalization conditions (8.6) and (8.8), they are running as noted for the charge before in section 7.6.1. Thus, the parameters of the theory become energy-dependent, and out of a set of theories with fixed parameters e and m a single theory with energy-dependent parameters emerges.

Note that, as discussed already in section 3.2, masses manifest themselves as poles in propagators, and only in the pole scheme this explicitly enters, while (8.6) is actually only indirectly related to this pole. In the end, everything which depends on the choice of scheme and scale are not physical. However, measurements can nonetheless be used to determine them within a fixed scheme as a function of the scale, using a prescription like (8.4-8.8). And thus one can, e. g., plot the energy-dependence of such a quantity. However, the plot is only meaningful after fixing the renormalization scheme.

As a consequence, standardized renormalization schemes have been developed, which are commonly used, and are therefore usually not made explicit. These schemes have been tailored for particular purposes, and must be looked up, if a calculation is to be compared to preexisting results. Some of those require other regularization techniques and renormalization schemes then those presented here. They will be introduced in the lecture quantum field theory II.

8.4 Dyson-Schwinger equations and resummation

There is now a discrepancy about dealing with masses. On the one hand, in section 3.2, the mass of a particle was associated with a pole in the propagator. On the other hand, the mass was here discussed by the relation (8.3).

There is, of course, the relation (2.20), which implies that the 2-point function (8.3) is the inverse of the propagator. Thus, for the propagator necessarily

$$D = \frac{1}{p^2 - m^2 + i\epsilon + \Pi_2} \tag{8.9}$$

where Π_2 is the perturbative expression for the 2-point function. That resolves the discrepancy. It is, however, useful to put this into a larger perspective. This will lead to the

[6]As a side remark, it should be noted that the exact masslessness of the photon can be shown to be a consequence of this broken scale symmetry in massless QED.

context of resummation.

To this end, the quantum equivalent to the equation of motions are helpful, the so-called Dyson-Schwinger equations. They follow from the fact that the path integral is translationally invariant. Hence, the path integral of any total derivative vanishes. Take for simplicity a theory with fields A, which are characterized by (multi-)indices a,

$$\begin{aligned} 0 &= \int \mathcal{D}A^a \frac{\delta}{\delta A^a(y)} e^{iS+i\int d^dx A^a(x) j^a(x)} \\ S &= \int d^dx \mathcal{L}. \end{aligned}$$

Here, S is the usual action, j^a is the source of A^a. Performing the derivative and pulling the resultant factor out of the integral by replacing A^a with $\delta/\delta j^a$, the equation

$$\left(\left. \frac{\delta S}{\delta A^a(x)} \right|_{A^a(x)=\frac{\delta}{\delta j^a(x)}} + j^a(x) \right) Z[j^a] \Big|_{j^a=0} = 0 \tag{8.10}$$

is obtained. Performing further derivatives will create a sequence of equations.

To establish the meaning of (8.10), it is necessary to recast it further. Reexpressing it in terms of the free energy W changes (8.10) to

$$\frac{\delta S}{\delta A_a}\left[\frac{\delta W}{\delta j} + \frac{\delta}{\delta j}\right] + j_a = 0.$$

Switching by a Legendre transformation to the quantum effective action Γ yields

$$\frac{\delta \Gamma}{\delta A_a} + \frac{\delta S}{\delta A_a}\left[A_a + \frac{\delta^2 W}{\delta j_a \delta j_b}\frac{\delta}{\delta j_b}\right] = 0.$$

This is the master equation to determine the Dyson-Schwinger equation of vertex functions.

To see how this works, consider the case of the Yukawa theory, and determine the Dyson-Schwinger equation for the fermion vertex functions. This requires to introduce anti-commuting sources for the fermion. This modifies the above to

$$Z = \int \mathcal{D}\overline{\Psi}\Psi\phi e^{iS+\int d^dx(\bar{\eta}(x)\Psi+\overline{\Psi}\eta(x)+j\phi)}.$$

The general procedure to obtain the corresponding Dyson-Schwinger equations is to start from (8.10) and then derive once more with respect to the field or with respect to the anti-field in case of anti-commuting fields for a two-point vertex, and correspondingly more often for an n-point function. The additional source term then yields the propagator while the right-hand-side of the equations are found by the derivative of the action.

Since in the course of the derivation, the source in equation (8.10) becomes the inverse full propagator, it makes sense to already rewrite (8.10) as[7]

$$j^a(x)Z = \frac{\delta S}{\delta \phi^a(x)}\Big|_{\phi^a(x)=\frac{\delta}{\delta A^a(x)}} Z \tag{8.11}$$

at the sources set equal to 0.

Using the action (5.10-5.11) with a single flavor and the scalar case, and differentiating with respect to $\bar{\eta}(x)$ yields

$$\left((i\gamma^\mu \partial_\mu^x - m_\Psi)\Psi(x) - g\phi(x)\Psi(x) + \frac{\delta\Gamma}{\delta\overline{\Psi}(x)}\right) e^W = 0 \tag{8.12}$$

where the x-index on a ∂ indicates the variable with respect to which to derive. This intermediate step shows why the Dyson-Schwinger equations are called the quantum equation of motion. Replacing everything with their classical counter-part at this stage, this is just the Lagrange equation of motion for the field Ψ. As W and the quantum effective action are both of order $\hbar$, they introduce the quantum information into these classical equation.

Replacing the fields in (8.12) by their respective derivatives and divide, after performing the derivation, by $\exp(W)$ yields

$$(i\gamma^\mu \partial_\mu^x - m_\Psi)\Psi(x) - g\left(\frac{\delta W}{\delta j(x)}\frac{\delta W}{\delta \bar{\eta}(x)} + \frac{\delta^2 W}{\delta j(x)\delta\bar{\eta}(x)}\right) + \frac{\delta\Gamma}{\delta\overline{\Psi}(x)} = 0$$

As a general feature of such derivations, terms containing products of only single derivatives of W appear at this stage. Here, the aim is only an equation for the propagator. Hence, when deriving such terms again with respect to the fields, always at least one single derivative remains, which can be replaced by a classical field. When setting the classical sources to zero at the end, also the classical fields are set to 0 and therefore these terms always vanish. Hence they can be neglected already at this stage of the calculation, and will not appear furthermore. This is not true when deriving equations for higher n-point functions. E. g., when determining the equation for a four-point function, each single derivative may be acted upon, yielding a non-vanishing term.

So the remaining expression is

$$(i\gamma^\mu \partial_\mu^x - m_\Psi)\Psi(x) - g\frac{\delta^2 W}{\delta j(x)\delta\bar{\eta}(x)} + \frac{\delta\Gamma}{\delta\overline{\Psi}(x)} = 0$$

To obtain the equation for the fermion propagator, this equation is derived once more with respect to $\Psi(y)$ which leads to

$$(i\gamma^\mu \partial_\mu^x - m_\Psi)\delta(x-y) - g\frac{\delta^3 W}{\delta\Psi(y)\delta j(x)\delta\bar{\eta}(x)} + \frac{\delta^2\Gamma}{\delta\Psi(y)\delta\overline{\Psi}(x)} = 0$$

[7]It should be noted that the equations may differ in form depending on the order of derivatives, though of course this only corresponds to rearrangements due to identities relating different vertex functions.

The last term is already the inverse propagator. The second term yields the interaction part. Using the inverse chain rule yields

$$\frac{\delta^2 W}{\delta j(x)\delta\bar{\eta}(x)} = -\int d^d z d^d w \frac{\delta^2 W}{\delta j(z)\delta j(x)} \frac{\delta^2 \Gamma}{\delta\overline{\Psi}(w)\delta\phi(z)} \frac{\delta^2 W}{\delta\eta(w)\delta\bar{\eta}(x)}$$

where the minus arises due to the anti-commuting derivatives. Using further the fact, that

$$\frac{\delta^2 \Gamma}{\delta\overline{\Psi}(w)\delta j(z)}|_{j=\eta=\bar{\eta}=0} = 0,$$

which generalizes to all mixed two-point functions, it is possible to write down the result in position space

$$D^{-1}_{\overline{\Psi}\Psi}(x-y) = i\gamma^\mu \partial^x_\mu - m_\Psi + g\int d^d z d^d w D_{\phi\phi}(x-z) D_{\overline{\Psi}\Psi}(x-w)\Gamma^{\overline{\Psi}\Psi\phi}(y,w,z).$$

where inverse 2-point functions have been replaced by propagators, and 3-point functions by the vertex.

Replacing all expressions with their Fourier-transformed[8] and afterwards dropping $\int d^d p/(2\pi)^d \exp(-ip(x-y))$, produces the result in momentum space as

$$D^{-1}_{\overline{\Psi}\Psi}(p) = i\gamma^\mu p_\mu + m_\Psi + \int \frac{d^d q}{(2\pi)^d}(g) D_{\phi\phi}(p-q) D_{\overline{\Psi}\Psi}(q)\Gamma^{\overline{\Psi}\Psi\phi}(-p,q,p-q), \tag{8.13}$$

where momentum conservation at the vertex has been used. The expression (g) (with suppressed unit matrix in Dirac space) can be identified to be the tree-level vertex. Using this, allows to rewrite (8.13) as

$$D^{-1}_{\overline{\Psi}\Psi}(p) = i\gamma^\mu p_\mu + m_\Psi + \int \frac{d^d q}{(2\pi)^d}\Gamma^{\overline{\Psi}\Psi\phi}_{\text{tree-level}}(-q,p,q-p) D_{\phi\phi}(p-q) D_{\overline{\Psi}\Psi}(q)\Gamma^{\overline{\Psi}\Psi\phi}(-p,q,p-q),$$

which is the final form.

In a similar way, all Dyson-Schwinger equations can be derived. As is visible, this is algorithmic, and can therefore be automatized. Note that the equations couple different correlation functions of different order. In the present case, the equation for the fermion propagator is coupled to the scalar propagator and the scalar-fermion vertex, and thus an n-point vertex of higher order. Generically for a theory which has at most quartic terms in the Lagrangian the equations for an n-point function involves the $n+1$ and $n+2$ point functions. Furthermore, in general in such theories also expressions with two integrals in

[8] All momenta are always defined incoming and momentum conservation at the vertices is taken into account. Hence in principle one of the arguments of the vertices could be dropped, but since this depends on conventions, all are kept.

momentum space appear. The Dyson-Schwinger equations therefore are a coupled system of non-linear integral equations, whose exact solutions would yield the exact result for all correlation functions.

More importantly, the structure shows that if all correlation functions would be expanded in Taylor series in the coupling this is another way to generate perturbation theory. Thus, this reproduces the relation (8.9). However, this also shows now how the concept can be generalized beyond propagators.

A general Dyson-Schwinger equation will have the form

$$\Gamma_{\text{full}} = \Gamma_{\text{tree-level}} + \Pi$$

where the inversion only happens for the propagator. Π is some involved self-energy. Thus, this shows how perturbative contributions build up, order by order, contributions to the exact Π.

It is now an interesting possibility to partially resum the perturbative series. Generically, due to the 1PI nature, perturbation theory will have th structure

$$\Pi = g\pi^1_{\text{1PI}} + (g\pi^1_{\text{1PI}})^2 + g^2\pi^2_{\text{1PI}} + (g\pi^1_{\text{1PI}})^3 + g^3 \sum_{\text{permutations}} \pi^1_{\text{1PI}}\pi^2_{\text{1PI}} + g^3\pi^3_{\text{1PI}} + \mathcal{O}(g^4)$$

where the upper index of the π implies the number of vertices. Thus, there are new contributions at every order, but also contributions, which are just products of lower-order contributions. It is thus possible to reorder the perturbative series, assuming it is convergent, as

$$\Pi = \sum_{i=0}^{\infty} (g^1_{\text{1PI}})^i + \mathcal{O}(\pi^{n>1})$$

which is called a resummation of a certain subclass of diagrams with one vertex, which often can be done analytically. Of course, it is not guaranteed that this improves the result, as it mixes order in the original Taylor series.

E. g., in case of the scalar propagator in section 8.1, this would be a summation of the tadpole diagrams. Since they are constant, this would just absorb all of them in the mass renormalization, and would certainly not alter the situation. However, for the fermion propagator in section 8.2, where the lowest order is already momentum-dependent, this is not a-priori clear. Often, in turns out that the result improves in certain momentum regimes, but does degrade in others.

The Dyson-Schwinger equations can improve the situation, if self-consistency is required to a given order in the coupling, as their non-linearity makes this a non-trivial constraint. However, this is beyond the scope of this lecture.

8.5 What is a particle?

As has been seen, the analytic structure and the asymptotic features of correlation functions start to become more and more important. It is therefore worthwhile to formalize these questions to some extent. For this, it is useful to first introduce the concept of a spectral density, and the Källen-Lehmann representation.

The propagator plays a central role in the description of physics, as the LSZ-formalism of section 3.2 has shown. Consider for simplicity a scalar particle. Start by rewriting its propagator as

$$D(p) = i\int_0^\infty dM^2 \frac{\rho(M^2)}{p^2 - M^2 + i\epsilon}, \tag{8.14}$$

which defines the spectral density ρ. This is known as the Källen-Lehmann representation. For a free particle of mass m,

$$\rho(M^2) = \delta(M^2 - m^2) \tag{8.15}$$

this will reproduce the free propagator. Thus, the spectral density describes the deviation of the propagator from the free one. Because of[9]

$$0 < i\langle\phi^\dagger(0)\phi(0)\rangle = iD(0) = \int_0^\infty dM^2 \frac{\rho(M^2)}{M^2 + i\epsilon}, \tag{8.16}$$

it follows that the spectral density needs to be positive. This implies that

$$\frac{\partial^n D(-p^2)}{\partial(p^2)^n} = (-1)^n i\int_0^\infty dM^2 \frac{\rho(M^2)}{(p^2 + M^2)^n} \tag{8.17}$$

and thus that the propagator and its derivatives are monotone functions of space-like momenta.

If the propagator contains information about an asymptotic state[10] in the sense of section 3.2, this part can be explicitly extracted

$$D(p) = \frac{iZ}{p^2 - m^2 + i\epsilon} + i\int_{m_t^2}^\infty dM^2 \frac{\sigma(M^2)}{p^2 - M^2 + i\epsilon},$$

with

$$\rho(M^2) = Z\delta(M^2 - m^2) + \sigma(M^2), \tag{8.18}$$

and thus σ is also positive. The quantity m_t is the smallest value for which σ needs to be non-vanishing such that (8.14) still holds. It is called the threshold mass, and will be seen

[9] This follows also from cluster decomposition.

[10] Multiple asymptotic states with the same quantum numbers would provide a sum over free-particle propagators.

in the following to encode the energy to obtain the simplest multiparticle state with the same quantum numbers as the single-particle state.

Combining (8.17) and (8.18) yields

$$1 = Z + \int_{m_t^2}^{\infty} dM^2 \sigma(M^2)$$

implying the Oehme-Zimmermann superconvergence relation

$$0 \leq Z < 1. \tag{8.19}$$

Thus, the asymptotic residuum is bounded from above by the free-particle residuum.

To continue, it is useful to relate these statements to further properties of a quantum field theory. Any matrix element can be considered to be an S-matrix element, if the arguments are not taken to be asymptotic. Furtheremore, as the S matrix is considered to encode time evolution, it is necessarily a unitary operator. It thus follows that

$$S^\dagger S = \mathbf{1} = \mathbf{1} + i(T - T^\dagger) + TT^\dagger.$$

The T-matrix is related to the matrix element. Thus, this provides the following implication for any matrix element with differing initial state and final state

$$\mathcal{M}_{f\to i} - \mathcal{M}^*_{i\to f} = i \sum_n (2\pi)^4 \delta^4(Q_n - P_i) \mathcal{M}^*_{n\to f} \mathcal{M}_{n\to i} \tag{8.20}$$

where Q_n and P_i are the total momenta in the states n and i. This implies a relation between any given process and all possible processes, which satisfy the conservation laws. Of course, any such state n will necessarily also appear as an intermediate state in the full process $i \to f$. This suggest already that there is a possibility to cut diagrams in half in a certain sense, and relate them. Especially, in a diagrammatic expansion, like perturbation theory, the decomposition into amplitudes $\mathcal{M}_{n\to i/f}$ indeed is equivalent to cutting a diagram in two. As the left-hand side will be related to cross sections for certain processes below, this often implies that diagrams can be calculated in perturbation theory by decomposing them. This is especially useful if this approach cuts a loop, and thus makes an integration unnecessary. The formalization of this is known as Cutosky rules.

To give an example, the optical theorem, consider the elastic case of $i = f$, but having no identical fields in the initial (or final) state. That can be relaxed, but requires some more work. In that case

$$\Im \mathcal{M}_{i\to i} = \frac{i}{2} \sum_n (2\pi)^4 \delta^4(Q_n - P_i) \left|\mathcal{M}_{n\to i}\right|^2. \tag{8.21}$$

However, the right-hand side is nothing but the total cross section, i. e. the possibility that the initial state goes to anything. Taking the kinematics from (3.20), this yields

$$\sigma_{\text{total}} = \frac{1}{2\sqrt{s^2+m_{i1}^4+m_{i2}^4-2(sm_{1i}^2+sm_{2i}^2+m_{1i}^2m_{2i}^2}}\sum_n (2\pi)^4\delta^4(Q_n-P_i)\left|\mathcal{M}_{n\to i}\right|^2. \tag{8.22}$$

Herein, the very useful relation

$$\begin{aligned}\int \frac{d^3\vec{p}_1 d^3\vec{p}_2}{4(2\pi)^2E_1E_2}\delta(p_1+p_2-Q) &= \frac{\sqrt{s^2+m_1^4+m_2^4-2(sm_1^2+sm_2^2+m_1^2m_2^2}}{8s\pi^2}d\Omega \\ &= \frac{\sqrt{\lambda(s,m_1,m_2)}}{8s\pi^2}d\Omega\end{aligned}$$

has been used. Combining (8.21) and (8.22) yields the optical theorem,

$$\Im\mathcal{M}_{i\to i} = \sqrt{s^2+m_{i1}^4+m_{i2}^4-2(sm_{1i}^2+sm_{2i}^2+m_{1i}^2m_{2i}^2}\sigma_{\text{total}}$$

which thus relates the imaginary part of the elastic matrix with the total cross section, the latter including all inelastic options.

Conversely, this implies a bound on the total cross section in the forward direction,

$$\left.\frac{d\sigma_{\text{elastic}}}{dt}\right|_{t=0} = \frac{1}{16\pi}\left(\frac{(\Re\mathcal{M}_{i\to i})^2}{s^2+m_{i1}^4+m_{i2}^4-2(sm_{1i}^2+sm_{2i}^2+m_{1i}^2m_{2i}^2}+\sigma_{\text{total}}^2\right) \geq \frac{\sigma_{\text{total}}^2}{16\pi}.$$

and thus σ_{total} is bound from above by an elastic process.

There is a further useful possibility to obtained from (8.20). If the energy is small enough to be below the inelastic threshold, in the sum in (8.20) only the initial state contributes, as the matrix elements connect on-shell amplitudes. If the initial state has total angular momentum j, the corresponding case can be made inclusive, and the equation then becomes

$$2i\Im\mathcal{M}_{i\to f}^j = 2i\frac{\sqrt{\lambda(s,m_1,m_2)}}{s}\left|\Im\mathcal{M}_{i\to f}^j\right|^2$$

where the left-hand side was simplified from (8.20) by assuming T-invariance.

In general, there is a very useful insight for an arbitrary $2\to 2$ cross section to be gained from this. Because the only dependence is on the scalar s and the scattering angle, it is possible to expand the cross section in Legendre polynomials P_l, leading to the following list of equalities

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \frac{1}{64\pi s\vec{p}^2}\left|\mathcal{M}_{2\to 2'}\right|^2 \\ \mathcal{M}_{2\to 2'} &= 16\pi\sum_{l\geq 0}(2l+1)F_l(s)P_l(\cos\theta) \\ F_l(s) &= \frac{\sqrt{s}}{2|\vec{p}|}e^{i\delta_l}\sin\delta_l\end{aligned} \tag{8.23}$$

where F_l are the so-called partial waves, and δ_l is the phase shift. The whole expression is called a partial wave decomposition. Especially, each partial wave includes the case that an s-channel exchange of an intermediate particle of spin l happens. Of course, as soon as in the inelastic case final states with more than two particles become available, no such expansion is anymore possible. But generalizations using other base systems with more angles, e. g. hyperspherical coordinates are still available, especially in cases where the final state masses can be neglected.

Using the previous statements also yields

$$\begin{aligned}\sigma_{\text{total}} &= \frac{8\pi}{\sqrt{s}|\vec{q}|}\sum_{l\geq 0}(2l+1)F_l(s)\\ |F_l|^2 &\leq \frac{\sqrt{s}}{2|\vec{q}|}\Im F_l\end{aligned}$$

where in the second line equality holds in the elastic case. Conversely, this implies the phase shift becomes complex above the inelastic threshold. As it is a phase, this implies that there is a unit circle, the Argand circle, on which the phase shift moves in the elastic case, and can move inside above the inelastic threshold. As angular dependence can be measure, this provides a valuable approach to understand experiments.

It has now been visible that a particle-like object is obtained by a pole in the (asymptotic) spectral density, and allows for the appearance of a resonance in a partial wave of its angular momentum. Thus, quantum field theory exhibits as asymptotic states objects which are localized in position space, and have definite spin. In particular, cluster decomposition (3.17) implies that there asymptotic states are localized.

It is possible to proceed with further characterization. The first is that it has been implicitly assumed that matrix elements are tempered, i. . polynomial bounded in momentum space up to measure-zero singularities like Dirac functions. Otherwise, Fourier transformation would not be possible. Indeed, before taking the limit the path integral construction ensures this feature.

Consider now again a $2 \to 2$ process for two real, distinct scalars A and B. They can then be combined into a single complex field, e. g. $\phi = A + iB$. Non-zero spin will only add logistics, but no physics. Factoring off the momentum conserving δ-function, and introduce the change of momenta

$$\begin{aligned}q &= \frac{1}{2}(p_1+q_1)\\ p &= \frac{1}{2}(p_2+q_2).\end{aligned}$$

These two quantities are already more than enough to fully characterize the process, due to Lorentz invariance. In fact, the dependence of p an be entirely dropped, as only two

masses, the center of mass energy, and the scattering angle are ultimately enough to characterize the process. It is then possible to consider the matrix element as an enlarged object by allowing q and p to be complex. Consider now the explicit form

$$\mathcal{M}^{AB\to AB}(q) = i\int d^4z e^{iqz}\mathcal{M}^{AB\to AB}(z).$$

Because the matrix are tempered, this implies that $\mathcal{M}$ is an analytic function if $(\Im q)^2 > 0$ and $\Im q_0 > 0$, i. e. the vector $\Im q$ is forward time-like. This is the so-called forward tube, essentially the future light cone defined by $\Im q$.

From this follows, stated without proof, the crossing symmetry. Replacing particle A by its anti-particle $\overline{A}$, the difference for flipping the momentum is given by

$$\mathcal{M}^{AB\to AB}(q) - \mathcal{M}^{\overline{A}B\to\overline{A}B}(-q) = iC(q),$$

where $C(q)$ is some function. The statement of the crossing symmetry is that $C(q)$ vanishes outside the regions given by the physical mass-shells $p^2 = \pm m^2$, but coincide on the real boundaries of these regions. In analogy to analytical continuation in a single complex variable the so-called edge-of-the-wedge theorem then implies that both amplitudes are then analytically continuation of each other, at least in the absence of massless particles.

These features allow to conclude that the amplitude is an analytic function, and can be written either as

$$\mathcal{M}^{AB\to AB}(q) = -\frac{1}{2\pi}\int d\mu^2\int_\Omega d^4k d\mu^2\frac{\rho(k,\mu^2)}{(q-k)^2-\mu^2} \tag{8.24}$$

or as

$$\mathcal{M}^{AB\to AB}(q) = -\frac{1}{2\pi}\int d\mu^2\int_\Omega d^4k d\mu^2\frac{\rho(k,\mu^2)}{(q-k)^2-\mu^2}\left(\frac{(q-k)^2+\mu_0^2}{\mu^2+\mu_0^2}\right)^n + P(q). \tag{8.25}$$

These are the so-called Jost-Lehmann-Dyson representation, reminiscent and generalizing the Källen-Lehmann representation for the propagator (8.14). The version (8.25) needs to be used, if the integral in (8.24) is not converging. In that case n is a suitable parameter and $P(q)$ a suitable polynomial. The region of integration Ω for the four-momentum k is given by hyperboloid, which is define in terms of the external momentum p as well as μ, and is defined implicitly. This will not detailed here. However, this concludes that also the amplitude $\mathcal{M}^{AB\to AB}$ is an analytic function in the physical region. This can now be sued to determine the actual analytic structure.

In fact, it will be seen that $\mathcal{M}^{AB\to AB}$ is analytic up to isolated poles of order one and a cut. These features can then be interpreted in terms of the physics, and especially the

appearance of intermediate state particles. For this it is useful to define the incoming energy as

$$\nu = \frac{qp}{m_B} = \frac{s - m_A^2 - m_B^2}{2m_B}$$

where the normalization is chosen for convenience, to obtain a dimensionless quantity. However, the same statements can be cast also in terms of the center-of-mass energy s of t, which are then also taken to be complexified. It is furthermore useful to illustrate the situation with the case of a single exchange diagram.

In fact, the analyticity properties imply

$$\mathcal{M}^{AB\to AB}(\nu) = Q + \frac{1}{\pi}\int_{(m_A+m_B)^2}^{\infty} \Im\mathcal{M}^{AB\to AB}(\nu')\left(\frac{1}{\nu'-\nu} + \frac{1}{\nu'+\nu}\right).$$

Thus, the amplitude has cuts in both directions. Also, it is implied that the amplitude cannot grow faster than a power in ν (or s). In fact, this yields, together with the optical theorem, the so-called Froissart bound

$$\begin{aligned} \left|\mathcal{M}^{AB\to AB}\right| &\leq \text{const} \times s\ln^2(s) \\ \sigma_{\text{total}} &\leq \frac{1}{q^2}\ln^2(s). \end{aligned}$$

Thus, interactions can grow with s, but not arbitrarily so.

Reformulating the same analyticity in terms of s yields the cuts start at the elastic threshold $(m_A + m_B)^2$ in the real s plane. In terms of diagrams, this corresponds to the case that an intermediate state particle can split into both particles on-shell. Thus, a cut signals the appearance of decay thresholds.

The term Q contains isolated poles. There are three different cases. One is that there is a pole between zero and the cut on the real axis. Diagrammatically, this appears when there is a single line, and thus a propagator. The pole is then at the mass of the particle. Hence, real poles can be interpreted as the appearance of a single particle in the intermediate state. Note that these are the same masses as appear in asymptotic states. It should be noted that this pole appears at positive energy.

Another pole appears on the second Riemann sheet in the complex plane, i. e. at a complex value of the center-of-mass energy, but with positive energy. This is a structure, which does not naturally arise at tree-level. However, if the propagator would take the form

$$D(p) = \frac{i}{p^2 - m^2 + im\Gamma}, \tag{8.26}$$

which has a pole at $p^2 = m^2 - im\Gamma$, in which Γ is the width. A resummed propagator with

a total width determined by (3.19) would show such a structure[11]. Thus such a structure can be interpreted as the presence of a resonance.

Finally, there can be poles on the unphysical, i. e. negative energy, Riemann sheet, on the real axis below threshold. Such states are so called virtual resonances. Such virtual resonances can become real resonances if the parameters of the theories change. They can thus be interpreted as the potential of a theory to sustain bound states and resonances in general, but not at the given parameter values in actuality.

This concludes the list of possible structure in the amplitude. Going beyond $2 \to 2$ processes, similar considerations apply, but the number of Riemann sheets increases, and the logistics do so as well.

There is one more issue to make the statement of particle more precise, and these are super selection sectors. It has been seen in section 6.2 that particles carrying a conserved quantum number can be distinguished. Furthermore, Noether's theorem showed that there is a conserved charge associated with symmetries. Because of the conservation, time-evolution cannot change the charge of a system. Thus, any transition amplitude $\mathcal{M}_{i\to f}$ will neither. Hence, any physical process will not change the total charge in a system, and there is no unitary evolution operator which connects sectors of different charge. Therefore, any underlying space decomposes into separate, non-communicating subspaces of asymptotic states. This is not in a one-to-one relation to particles, though, as only the total charge is conserved. Thus, an asymptotic state with a single unit of charge connects to any state where there is one more particle than anti-particle. These isolated subspaces are thus super-selection sectors. This also applies, e. g. to tonal four momenta, as well as generalized symmetries, e. g. those based on non-Abelian Lie groups.

In the context of the path integral, this is an interesting situation. Charge is obtained from integrals over field configurations. Thus, the set of all field configurations decompose into distinct holonomy classes corresponding to the superselection sectors, but the path integral itself is a sum over all possible superselection sectors.

Together with cluster decomposition, this implies that an isolated particle is a field configuration, which belongs to a superselection sector which allows for one unit of charge, and in which there is an appreciable field amplitude which contains an integrated amount of one unit of charge, in a region of space-time, which is disconnected in all space-like directions from any other field configurations, which thereby need to have zero charge. It does not need to be time-like isolated, though. This is the closest concept for a single-particle field configuration possible. If along a time-like interval the situation remains,

[11] Perturbation theory is here hitting its limits. After all, all elementary particles should appear as asymptotic states, but also can decay. A full resolution needs to go beyond perturbation theory.

the propagator evaluated on this sub space-time volume will be just the asymptotic free particle propagator.

Of course, field configurations with such a configuration are rare compared to all possible field configurations in the path integral. The fact that these are still typical situations on long distance scales implies that they have correspondingly a large weight. Intuitively, the reason can probably be best understood by the fact that either interactions are screened in the sense of a Yukawa potential (7.29) or that different charges attract each other in the Coulomb potential, and thus tend to screen charges to zero. But this is not necessary, as other theories, especially the non-Abelian gauge theories to be encountered in the lecture quantum field theory II, will demonstrate.

8.6 Vertex corrections

Having the developments now pushed to this point, it is useful to investigate another process to illustrate the concepts further. That will be done with considering the one-loop correction to a vertex. Given its importance for the definition of the electric charge and the concept of running, a suitable objct to investigate is the photon-electron (or any other charged fermion) vertex.

At the 1PI level, there is only a single diagram contributing, the one in which a photon is exchanged between the electron and the positron leg. Choosing as kinematics the incoming and outgoing electron momentum to be p and p', and the photon momentum to be $q = -p' - p$. Since gauge-invariance would turn out to be maintained in the same way as in section 7.6.2 in the following already Feynman gauge will be chosen. The resulting one-loop expression then becomes

$$\mathcal{M}_\mu^{\gamma e^+ e^-} = (-ie\gamma_\mu) + \int \frac{d^4k}{(2\pi)^4} \quad \frac{-ig_{\nu\rho}}{(k-p)^2+i\epsilon}(-ie\gamma^\nu)\frac{i(\gamma_\alpha(-k-q)^\alpha + m_e}{(k+q)^2-m_e^2+i\epsilon} \times \\ \times(-ie\gamma_\mu)\frac{ik_\beta\gamma^\beta + m_e}{k^2-m_e^2+i\epsilon}(-ie\gamma^\rho).$$

It is visible, how a dependence only on the Lorentz index μ arises, and how this reproduces the general e^2-type correction.

Performing all possible contraction yields

$$\mathcal{M}_\mu^{\gamma e^+ e^-} = (-ie)\left(\gamma_\mu + 2ie^2\int\frac{d^4k}{(2\pi)^4}\frac{-k_\alpha(k+q)_\beta\gamma^\alpha\gamma_\mu\gamma^\beta + m_e^2\gamma_\mu + 2m_e q_\mu}{((k-p)^2+i\epsilon)((k+q)^2-m_e^2+i\epsilon)(k^2-m_e^2+i\epsilon)}\right)$$

where it has been used that

$$\gamma^\alpha\gamma_\mu\gamma_\alpha = -2\gamma_\mu$$

as a consequence of the Clifford algebra.

There are now two elements, which need to be taken care of. One is that there is still a non-trivial tensor structure in the nominator. The other is that the integral is relatively involved.

Start with the latter. The central problem is the highly non-trivial angular dependence. To get rid of it, so-called Feynman parameter can be used. In their simplest form, they are introduced using the identities

$$\frac{1}{AB} = \int_0^1 \frac{dx}{(xA+(1-x)B)^2} = \int_0^1 dxdy \frac{\delta(x+y-1)}{(xA+yB)^2}$$

where x and y are the Feynman parameters. This generalizes to n Feynman parameters x_i by

$$\frac{1}{A_1^{m_1}...A_n^{m_n}} = \int_0^1 \delta\left(\sum x_i - 1\right) \frac{\Gamma\left(\sum m_k\right)}{\left(\sum x_j A_j\right)^{\sum m_l}} \Pi_r \frac{x_r^{m_r-1} dx_r}{\Gamma(m_r)}.$$

That this is indeed a useful approach cannot be guessed, but can be seen in the following, when contrasted to not using Feynman parameters.

In the present case, there will be three Feynman parameters necessary. Considering only the denominator yields

$$\frac{1}{((k-p)^2+i\epsilon)((k+q)^2-m_e^2+i\epsilon)(k^2-m_e^2+i\epsilon)} = \int_0^1 dx_1 dx_2 dx_3 \delta\left(\sum x_i - 1\right) \frac{2}{D^3}$$

$$D = x_1(k^2-m_e^2) + x_2((k+q)^2-m_e^2) + x_3((k-p)^2) + i\epsilon\left(\sum x_i\right) \tag{8.27}$$

The whole expression is still to be integrated over k. This allows to perform a linear variable transformation[12] for k

$$k = l - x_2 q + x_3 p$$

which allows, using that the x_i sum to 1 due to the δ-function, to simplify D to

$$\begin{aligned} D &= l^2 - \Delta + i\epsilon \\ \Delta &= -x_1 x_2 qr + (1-x_3)^2 m_e^2 > 0 \end{aligned}$$

where the inequality holds because of the kinematic situation that the external particles are all on-shell. By this manipulation, the denominator no longer depends on any angle, but only on the absolute momentum squared. The numerator, is, of course, also affected by the necessary shift in the momentum. Tedious algebra yields the numerator to be

$$\gamma_\mu \left((1-x_1)(1-x_2)q^2 + (1-4x_3+x_3^2)m_e^2 - \frac{l^2}{2}\right) + \frac{i\Gamma_{\mu\nu}q^\nu}{2m_e}(2m_e^2 x_3(1-x_3)),$$

[12] If the integral is divergent, it is not obvious that this is allowed. In the present case, it turns out to be correct. A more general investigation will be provided in quantum field theory II, but fortunately in most cases it is correct.

and thus the numerator contains either terms independent of l^2 or linearly dependent on l^2. This allows for a substantial simplification.

Going back to the full integral, note that Lorentz invariance implies a decomposition in invariant tensors, and thus

$$\int \frac{d^4l}{(2\pi)^4} l_\mu f(l^2) = 0 \tag{8.28}$$

$$\int \frac{d^4l}{(2\pi)^4} l_\mu l_\nu g(q^2) = \frac{g_{\mu\nu}}{4} \int \frac{d^4l}{(2\pi)^4} l^2 g(q^2). \tag{8.29}$$

The original problem can be decomposed entirely into prefactors time these integrals. Thus, it is sufficient to solve such, very general, integrals to solve the original problem. Therefore, such expressions are also called master integrals, and it is a standard strategy to decompose complex expressions into known master integrals.

Of course, this implies to solve such integrals. To do so, in this case performing a wick rotation, i. e., performing the variable transformation $l_0 \to il_0$ is useful, as this turns the integral into four-dimensional Euclidean ones. As a consequence, the integral can be solved in four-dimensional hyperspherical coordinates.

Consider first (8.28) with the corresponding choice of f, but slightly generalized

$$\frac{i}{(-1)^m(2\pi)^4} \int d\Omega_4 d|l| \frac{|l|^3}{(|l|^2+\Delta)^m} = \frac{i(-1)^m}{(4\pi)^2} \frac{1}{(m-1)(m-2)\Delta^{m-2}}.$$

Likewise (8.29) yields

$$\frac{i}{(-1)^{m-1}(2\pi)^4} \int d\Omega_4 d|l| \frac{|l|^5}{(|l|^2+\Delta)^m} = \frac{i(-1)^{m-1}}{(4\pi)^2} \frac{1}{(m-1)(m-2)(m-3)\Delta^{m-3}}.$$

Note that the results are, however, only valid if m is sufficiently large, larger than 2 in the first case and larger than 3 in the second case. Considering (8.27), the latter is however not the case. Thus, there is again an ultraviolet divergency.

While this could be regularized as in section 8.2, it is useful to introduce a different approach here, which allows also to demonstrate another important concept, decoupling. This concept rests on the Appelquist-Carrazone theorem, which states that in perturbation theory in a perturbatively renormalizable theory the effects of a heavy particle of mass M in a process with maximum energy scale E, e. g. the center-of-mass energy in an s-channel process, is suppressed as a power of E/M, and thus becomes irrelevant at energies $E \ll M$. The proof follows essentially by expanding the general structure of perturbative matrix elements in this quantity. Thus, heavy degrees of freedom decouple from the low-energy dynamics. While a proof beyond perturbation theory is in general not possible, it remains at least a very good qualitative description for the theories relevant in nature.

Otherwise, a so-called ultraviolet-infrared mixing would make processes measured at low energy stronger than power-suppressed dependent on the properties of particles which are much heavier than the relevant energy scales. Conversely, this implies that heavy degrees of freedom can be detected if measurements are sensitive enough to detect power-suppressed contribution, a fact utilized successfully many times.

This idea underlies now the Pauli-Villar regularization. It introduces a heavy particle, or multiple heavy particles, artificially into the theory. Besides giving them a very large Λ, they are also equipped with an opposite sign kinetic term, and thus propagator. Thus, they are not physical. However, in the end, the limit $\Lambda \to \infty$ will be taken in the renormalization process, and thus this does not matter.

In principle, it can be necessary to introduce multiple of such particles. In the present case, it is sufficient to do so for the photon. Thus, any diagram including a photon is doubled, by replacing

$$\frac{g_{\mu\nu}}{q^2+i\epsilon} \to \frac{g_{\mu\nu}}{q^2+i\epsilon} - \frac{g_{\mu\nu}}{q^2-\Lambda^2+i\epsilon}.$$

For (8.29), this yields for the currently relevant case of $m = 3$

$$\begin{aligned}\frac{d^4l}{(2\pi)^4}\left(\frac{1}{(l^2-\Delta^2)^3} - \frac{1}{(l^2-\Delta_\Lambda^2)^3}\right) &= \frac{1}{(4\pi)^2}\ln\frac{\Delta_\Lambda}{\Delta} \\ \Delta_\Lambda &= \Delta + z_3\Lambda^2.\end{aligned}$$

The expression is only regularized, as it still diverges logarithmically for $\Lambda \to \infty$.

Putting everything together yields an expression of the form

$$\mathcal{M}_\mu^{\gamma e^+e^-} = (-ie)\left(\gamma_\mu F_1(q^2) + \frac{\Gamma_{\mu\nu}q^\nu}{2m_e}F_2(q^2)\right)$$

The quantities F_1 and F_2 are called form factors, as they describe the deviation of the vertex from the tree-level one, which is $F_1 = 1$ and $F_2 = 0$.

To renormalize it, it is necessary to introduce a counter-term. Given how the electric charge has been defined in the Thomson limit, it is reasonable to require $F_1(0) = 1$. This is also consistent, as F_2 does not contain the divergent term. There is, however, another issue arising in F_1 only, as the remaining integral in x_3 is not convergent. Tracing back the origin of the problem, this comes from the fact that the photon propagator has an infrared singularity. It can be traced back to the fact that the photon is massless. This issue can be regularized by introducing a small photon mass μ, which requires to replace

$$\Delta \to \Delta + x_3\mu^2.$$

The origin of this problem is somewhat subtle. A hand-waving argument is that the asymptotic states need to be wave-packets. But for a wave-packet the energy is not well-defined.

Because of that, it is not possible to distinguish the situation of a single electron and a single electron plus any number of very low-energy photons. This issue can be resolved by either defining suitable asymptotic states or by taking corresponding cancellations on the external lines explicitly into account by resumming subsets of diagrams. This leads to far astray for the current purpose. As an operative definition, μ can be taken as a cutoff which models the energy resolution of a detector for the asymptotic particles. A more refined discussion will be given in quantum field theory II.

This yields

$$\begin{aligned}
F_1(q^2) &= 1+\frac{\alpha}{2\pi}\int_0^1 (\Pi dx_i)\,\delta\left(\sum x_i-1\right)\left(\ln\frac{m_e^2(1-x_3)^2}{m_e^2(1-x_3)^2-q^2x_1x_2}\right.\\
&\left.+\frac{m_e^2(1-4x_3+x_3^2)+q^2(1-x_1)(1-x_2)}{m_e^2(1-z_3)^2-q^2x_1x_2+\mu^2x_3}-\frac{m_e^2(1-4x_3+x_3^2)}{m_e^2(1-x_3)^2+\mu^2x_3}\right)\\
F_2(q^2) &= \frac{\alpha}{(2\pi)}\int_0^1 (\Pi dx_i)\,\delta\left(\sum x_i-1\right)\frac{2m_e^2x_3(1-x_3)}{m_e^2(1-z_3)^2-q^2x_1x_2},
\end{aligned}$$

admittedly not especially transparent results. Note that the fact that they do seem to obey the Appelquist-Carrazone theorem is because the external momentum of the electrons is always of the same size as the electron mass, because the electron is not purely internal.

Since by renormalization the electric charge is defined at zero momentum transfer, the quantity $eF_1(q^2)$ can be considered to be the momentum-dependence of the electric charge, and thus its running.

It is worthwhile to note that this also modifies the Coulomb potential. While in the static case $\vec{p}\approx\vec{p}'\approx 0$ this makes itself felt by a change in the electric charge, the situation is also interesting when considering a moving electron. In this case, also a Lorentz force appears. Taking again the classical limit of $q\to 0$, the interaction strength is governed by a magnetic moment in the direction of the electron spin. The magnetic moment is given by

$$\frac{e}{2m_e}\left(2\left(F_1(0)+F_2(0)\right)\right).$$

There are no two facts. One is the tree-level effect that there is a factor 2 inside the outer parantheses. This doubles the magnetic moment compared to the one expected for a classical particle of charge e and mass m_e. It stems entirely from the spin of the electron. Moreover, in the present case, it is modified by the contribution in the inner parantheses. Since $F_1(0)=1$ due to renormalization, only $F_2(0)$ can change the result. Since the integral is finite, it can be evaluated, and indeed in this limit analytically, to yield $F_2(0)=\alpha/(2\pi)$. Thus, the quantum effects at NLO yield that the magnetic moment of the electron changes, though admittedly by only about 0.1%. However, this effect was measured.

Chapter 9

Beyond perturbation theory

So far, most of the actual calculations centered around perturbation theory. And while perturbation theory is and remains one of the most vital tools in quantum field theory, it is not capable to describe all of its phenomena. But, as in quantum mechanics, once the systematic idea of physics being expressible in terms of analytic functions is left, or needs to be left, much less becomes accessible in simple calculations. Thus, most of treating quantum field theory must be postponed to various other lectures. Here, however, a few general remarks will be made.

9.1 Subtleties and Haag's theorem

The first point will be to understand why quantum field theory cannot be the final answer. This can be illustrated by example of an ordinary integral, which has, however, the same structure as the perturbation theory and the path integral developed previously.

Consider a pseudo action

$$S = -x^2 + \lambda x^4.$$

In this case, the first, quadratic term, corresponds to the kinetic term. Here, an Euclidean situation has been chosen for ease of reference. The quartic term then represents self-interactions, with a coupling constant λ.

The corresponding pseudo-path integral can, in fact, be evaluated explicitly. This yields

$$\int dx e^{-x^2-\lambda x^4} = \frac{e^{\frac{1}{\lambda}}}{2\sqrt{\lambda}} K_{\frac{1}{4}}\left(\frac{1}{8\lambda}\right), \tag{9.1}$$

where K_n is the modified Bessel function of the second kind. This expression is finite for any value of $\lambda > 0$.

If the same logic is applied to this expression as when deriving perturbation theory in chapter 3, this yields

$$\begin{aligned}\int dx e^{-x^2-\lambda x^4} &= \int dx \sum_i \frac{(-1)^i \lambda^i x^{4i}}{i!} e^{-x^2} \neq \sum_i \frac{(-1)^i \lambda^i}{i!} \int dx x^{4i} e^{-x^2} \\ &= \sum_i \frac{(-1)^i \lambda^i \left(1+(-1)^{4i}\right)}{i! 2} \Gamma\left(2i+\frac{1}{2}\right) = \sum_i \frac{\sqrt{\pi}(-1)^i \lambda^i (4i-1)!!}{i! 2^{\frac{2i+1}{2}}} \\ &= \sqrt{\pi}\left(1 - \frac{3\lambda}{2\sqrt{2}} + \frac{105\lambda^2}{8\sqrt{2}} - \frac{3465\lambda^3}{16\sqrt{2}} + \frac{675675\lambda^4}{128\sqrt{2}} + ... + \mathcal{O}\left(\frac{(\lambda i)^i}{\sqrt{i}}\right)\right) = \infty.\end{aligned}$$

Thus, the result differs from the exact result (9.1), and is not even well defined. The reason can be seen from two different perspectives. On the one hand, the exact result (9.1) has an essential singularity for $\lambda = 0$. Thus, an expansion in λ around $\lambda = 0$, as is performed in perturbation theory, cannot work. The other perspective is given where the unequal sign appears. Because the series is not absolutely convergent, exchange the summation and the integration, as is vital for the perturbative calculation, is actually not correct. It is also visible that the pre factor increases factorial with the order. This is actually something which also can happen in quantum field theory, as graph-theoretical consideration show that the number of Feynman diagrams increase factorial with the order. However, this does not preclude the possibility of cancellations.

Of course, in a genuine path integral neither an exact solution nor an explicit check of absolute convergence is usually possible. However, the issue can be recast into a different statement, known as Haag's theorem. The perturbative expansion is around the free theory. Especially, it requires a smooth limit to the free theory, when the coupling is switched off. But if this would be the case, there should exists a smooth, unitary field transformation, which is parametrized by the coupling, by which the theory can be transformed into a free theory of new fields. This is exactly what is forbidden by Haag's theorem: The free theory and the interacting theory are not unitarily equivalent.

To get an intuitive, hand-waving picture. Consider a scattering of two non-identical, real fields. In a free theory

$$\langle \phi(x)\psi(y)\phi(z)\psi(s)\rangle = \langle \phi(x)\phi(z)\rangle \langle \psi(y)\psi(s)\rangle = D_\phi(x-z) D_\psi(z-s).$$

The expression is independent of the distance $x - z$, and thus a constant as a function of it. If the correlation function should be an analytical function, in which it make sense to take the distance smoothly to any distance $|x - z|$, analyticity implies that the function is necessarily identical to the constant function in $|x - z|$, even if $|x - z|$ is timelike.

In the interaction theory, however,

$$\langle \phi(x)\psi(y)\phi(z)\psi(s)\rangle = \langle \phi(x)\phi(z)\rangle \langle \psi(y)\psi(s)\rangle = D_\phi(x-z) D_\psi(z-s) + \Gamma_{\phi\phi\psi\psi}(x-y, x-z, y-s)$$

If Γ should be non-trivial, it cannot be a constant in $|x-z|$, and much less so, if the distance is time-like. Especially, if the theory could be expanded in g, then there would always be combinations $g^n|x-z|^m$ appear in any such expansion. But then the radius of convergence could always be made arbitrarily small by varying the distance such as to offset any change in g. Hence, the radius of convergence can always be made arbitrarily small, and thus, for any infinitesimal g, there is always interaction remaining, and the theory is thus in the limit of $g \to 0$ always interacting. In contrast, at $g = 0$, it is non-interacting. All of this can be framed mathematically more precise, but this is beyond the scope of this lecture.

9.2 Bound states

However, Haag's theorem yields interesting question when it comes to asymptotic states. After all, they had been defined as non-interacting states, which emerge as the result of time-evolution. But since time-evolution is unitary, this cannot be exactly true. As a consequence, cluster decomposition (3.17) cannot be literally correct in an interacting theory.

Of course, this does not imply that this is a quantitative relevant issue. Indeed, theories are known for which this is both quantitatively relevant and irrelevant. After all, even exponentially suppressed interactions are sufficient to satisfy Haag's theorem, as long as they are not zero. Then, neglecting them can likely be a negligible effect for all but exact calculations.

An alternative to such a way out is that the asymptotic states are not, in fact, the elementary states of the Lagrangian, but rather composite states, made from more then one elementary particles. The simplest such case is actually QED with electrons and protons. The hydrogen atom is itself a stable bound state, and thus can appear as an asymptotic state. Insisting only on free states as asymptotic states would miss it, and thus grossly misrepresent the physics of all of QED, which includes chemistry. However, this also implies that perturbation theory, as developed in chapter 3, is necessarily unable to describe a stable hydrogen atom.

Thus, it is necessary to describe such bound states. At the same time, it appears than reasonable to allow that for bound states cluster decomposition can be applied, as the bound state's constituents still interact. Thus, using bound states, and only bound states, as asymptotic states satisfies Haag's theorem.

But to formulate cluster decomposition requires that the bound state itself can be represented by an object depending only on the position of the bound state. This is

achieved by employing composite operators. E. g., for the hydrogen atom, this composite operator would be given by

$$H(x) = \overline{\Psi}_p(x)\Psi_e(x)$$

where the resulting operator is a scalar. Thus, the composite operator consists of two (or more) operators evaluated at the same space-time point[1].

In momentum space, this yields

$$H(p) = \int \frac{d^d q}{(2\pi)^4} \overline{\Psi}_p(p-q)\Psi_e(q).$$

Thus, while the bound state depends on a single momentum only, allowing it to fulfill its only energy-mass relation $p^2 = m_H^2$, the actual structure corresponds to the constituents having all possible relative momenta, over which they are averaged. Especially, the constituents are not necessarily on-shell, and therefore in general the mass of the bound state is different from the sum of masses of the elementary state, a phenomena known as mass defect.

Calculating the properties of the bound states, like its mass, is in most theories a very hard problem, which requires non-perturbative methods to solve. However, eventually they will show up in correlation functions of the corresponding quantum numbers in the same way as elementary particles, i. e. as poles or as scattering states as cuts.

9.3 Scattering of bound states

The idea of bound states as more appropriate asymptotic states in many cases than elementary states leads ultimately to the concept of superselection sectors. A superselection sector is characterized by the set of all conserved quantum numbers, including e. g. spin and fermion number.

The statement is now that any physical process can only happen within a fixed superselection sector. That is, time evolution cannot change the superselection sector. Especially, this implies that for a scattering process that the initial quantum numbers and final quantum numbers coincide. However, as noted, this has nothing to do with particles, as this is really the integrals over all conserved currents in terms of the underlying fields on a spatial hypersurface. Thus, an initial state of a single fermion as asymptotic state is still

[1]This can be extended to multiple operators. It is also possible to build bound state operators without both fields at the same point, as long as any spatial variation is averaged in such a way as the final operator is only depending on a single space-time point. Furthermore, such coinciding operators provide new challenges for renormalization. All of these issues are beyond the scope of this lecture.

connected to a scattering state of any number of fermions and antifermions, as long as there is always one more fermion than anti-fermion.

This also allows the formation or break-up of bound states in a similar manner. Thus, as with elementary states, it is possible to prepare the initial state and final state in terms of bound states. Almost the complete LSZ formalism of section 3.2 will go through unchanged. The differences is mainly that the composite operator, rather than the elementary operator, will appear in the matrix element. This has also consequences for the external wave-function amplitude, which needs now to be replaced by a composite amplitude, called Bethe-Salpeter amplitude (or, for fermionic composite particles, Faddeev amplitude).

While conceptual straightforward extensions, this involves substantially more complications at the technical level. This is therefore beyond the scope of this lecture.

9.4 Universal threshold expansion

There are many ways in which bound states differ from point particles. Most pronounced is the existence of an extension. However, this cannot be interpreted as a radius. As noted, particles, and thus also bound states, need really to be though of as localized excitations of the field, or fields. Thus, it is better to think of the term as an effective range or size parameter, which determines roughly the effective domain of influence. However, this size depends on the interaction.

Thus, determining such a domain of influence is best seen as probing a particle with some force (carrier), not unlike the situation with electric charge in section 7.6.1 and 8.6. Therefore, three-point functions can be used for this purpose.

There are two possibilities to do so. One is probing. In that case, to probe the size of a particle ϕ under interaction mediated by a particle ψ, the 1-PI three-point function $\langle\phi\psi\phi$ is used. In quantum mechanics or classical physics, this corresponds to a probe particle current of *phi*, which interacts with a potential due to ψ. In analogy to the Fourier transform of a potential due to some spherical symmetric body, a size parameter is defined as

$$\langle r^2\rangle = -6\left.\frac{\partial\mathcal{M}^{\phi\psi\phi}(0,q,0}{\partial q^2}\right|_{q^2\to 0+} \tag{9.2}$$

and thus corresponds quantum mechanically to the expectation value of the average size squared.

Likewise, if a stable particle can be formed in an intermediate stage when two (identi-

cal[2]) particles collide elastically, there exists a so-called universal threshold expansion of the scattering phase δ_l from (8.23) as

$$|\vec{p}| \cot \delta_l = \frac{1}{a_0} + \sum a_i^{2i-1} |\vec{p}|^{2i} \tag{9.3}$$

where $|\vec{p}|$ is the three momentum of the interacting particles, and the radius of convergence is usually smaller than the distance to the inelastic threshold or the next resonance.

The quantity a_0 is the scattering length. It is negative if there exists a bound state below the elastic threshold, and positive if there is none or only point-like particles. For the same interaction, its square is usually of the same size as $\langle r^2 \rangle$. Its definition stems from the fact that for a classical particle, the hard-sphere cross-section of the bound state would be πa_0^2.

Such calculations and measurements) allow to identify bound states, and separate them from point-like particles.

[2]There are also more complicated formulas for the case of more involved initial states.

Chapter 10

Canonical quantization

The discussion so far has concentrated on a path-integral formulation, giving its predominance in modern relativistic quantum field theory. While also non-relativistic quantum field theory can be formulated in terms of a path integral, this is still less familiar, as it does not lend itself as easy as the operator formalism to non-relativistic calculations. However, non-relativistically, both are equivalent. It thus stands to reason that also relativistically such a treatment should be possible. And indeed, this is possible, and historically this was the first version of relativistic quantum field theory.

In fact, many problems can be equally well be dealt with in an operator formalism. Especially, it will be seen that the perturbation theory using Feynman rules emerges in an identical form. Hence, once these have been derived, it becomes essentially indistinguishable in perturbation theory which underlying formalism is used. That is, of course, different beyond perturbation theory.

Where using operators becomes difficult is when dealing with gauge theories. As has been seen in chapter 7, even the simplest gauge theory involves massless, and thus ultra-relativistic, particles. The formulation of operators uses a generalization of time evolution, and thus a Hamilton operator. The latter is, however, not relativistically invariant. The combination of massless particles without rest mass and a dynamical, frame-dependent principle does not work well together.

The path integral formulation, due to its Poincaré-invariant nature, lends itself more easily, and thus became more important. However, in the non-relativistic limit, there again exists a preferred frame, and thus the operator formalism remains very important in that case.

10.1 Canonical quantization

One issue, which needs to be taken into account in a relativistic treatment is that the Poincaré group necessarily treats space and time on the same footing. Thus, the distinct treatment of quantum mechanics, one as parameter the other as operator, is fundamentally impossible in a relativistic quantum theory. Moreover, the energy-mass relation necessarily implies that a particle concept cannot be suitable, as do classical mechanics and electrodynamics strongly suggest to pass to a field formulation.

All of that together leads to postulates[1], which eventually turn out to be adequate.

The starting point is to degrade space again to a parameter. Thus, any quantity is parametrized by a (four-)vector in Minkowski space-time, or, more accurately, is a map starting in Minkowski space-time. Furthermore, they map into linear operators acting in a Hilbert space (or vector space in general), $x \to \Phi(x)$, where the operators can carry in addition a representation of the Poincaré symmetry, just like the fields in the path integral formalism in chapter 4.

In addition, it is required that the operators fulfill canonical equal-time (anti-)commutation relations,

$$[\Pi(t,\vec{x}),\Phi(t,\vec{y})]_{\pm} = i\delta^3(\vec{x}-\vec{y})$$

and the (anti-)commutator of Π and Φ vanish at the same time. The operator Π is the canonical conjugate momentum to Φ. It is at this level postulated to exist. Later, it can be determined in the same way as in classical mechanics, by postulating a Hamiltonian operator. Given a Lagrangian formalism, this Hamiltonian operator can be obtained from the Lagrangian in the usual way, keeping the commutation relations in mind. Thus, the existence of a suitable Hamiltonian $\mathcal{H}$ is assumed in the following.

This result immediately shows that the whole formalism is not explicit frame-invariant. However, it also shows that it very much embeds the fact that space-like distances and time-like distances are different.

It should be noted that, in contrast to quantum mechanics, these operators are not associated to measurements, nor are they analogous to wave-functions. Even though they are often Hermitiean. A measurement in a quantum-field theory will rather produce a classical field $\phi(x)$ and canonically conjugated momentum field $\pi(x)$, from which the value

[1]Sometimes, this is called second quantization. This has been a historical name, which based on the idea that in quantum-field theory operators become quantized, and thus is a second level of the quantization process. As will be seen, this stems from the fact that operators can be expanded in other operators. However, as this is also possible in quantum mechanics, and indeed quantum mechanics is unnecessary to formulate an operator-based version of quantum field theory, this name becomes less and less used, as it is really a misnomer.

of actual observables can then be determined. But there is no analogue to the Born rule, because the fields are extended over all space and time, and thus lack an outside. Indeed, it is only possible to construct operators in terms of the fields and the symmetries of which they carry representations. There are no other operators. Rather, the probability to find a given field configuration requires to determine the correlation functions of all possible number of combination of fields, which absolute squares then give the probability to find a certain field carrying such correlations.

Similarly to quantum mechanics, these postulates only provide the kinematics. There is still a dynamical principle missing. This comes in terms of the Hamiltonian, which describes the (infinitesimal) evolution of the operators in terms of the Heisenberg equation

$$\partial_t \mathcal{O}(t,\vec{x}) = -\left[\mathcal{H}, \mathcal{O}(t,\vec{x})\right],$$

which describes the change when moving from an initial spatial hypersurface. Like in quantum mechanics, a time evolution operator can be build from the Hamiltonian, with the usual complications if the Hamiltonian, which is necessarily given in terms of the fields, does not commute with itself at different time,s if it is time-dependent.

Since the Hamiltonian is Hermitiean, the states in the underlying Hilbert space can be given in a basis of the eigenstates of the Hamiltonian. As usual, it relates to the four-momentum operator $P_0 = \mathcal{H}$ by virtue of being, by the Heisenberg equation, the generator of time translation. However, P^0 is the total four-momentum operator, and thus does not need to describe a fixed particle content.

As usual, by a Legendre transformation a Lagrangian can be constructed, which is then necessarily the same as the one appearing in the path integral formulation. Just that, like in quantum mechanics, operators need to be replaced by functions. The equivalence of this formulation to the path integral can be proven in essentially the same way as in quantum mechanics, except for the subtleties to be described in section 10.2. It will therefore not be detailed here.

To have a stable system, the spectrum of the Hamiltonian needs to be bounded from below. The state of lowest energy is the vacuum. Note that the vacuum is not particle free. After all, there is also no particle concepts, just the field operators. The idea of particles need to be reconstructed in the same way as in the path-integral representation by localized field excitation. Also, the vacuum is a state, and there is no vacuum situation for the fields, just for their expectation value. After all, the fields are operator-valued.

Finally, expectation values can be constructed as in quantum mechanics, by performing generalized vector-matrix-vector products. But there are some subtleties.

10.2 Normal ordering, time ordering, and Wick contractions

.

It is instructive to see how perturbation theory works. For this consider first again the situation of a single, free, scalar field. Since the theory is non-interacting, the state space will be made up of states, which can be interpreted as n-particle states with three-momenta $\vec{p}_i$, which fulfill the energy-momentum relation $p_0^2 - \vec{p}^2 = m^2$, and are thus necessarily all on-shell.

Define now operators[2] $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$, which add one quantum of three momentum $\vec{p}$ to any given state, with the lowest-energy state, the vacuum (which, as usual, is not the zero vector), having no such quanta. The total energy is given by the corresponding relativistic dispersion relation. Therefore, such states could be interpreted as single-particle states. However, this is only possible because the theory is free. Otherwise, the dispersion relation changes because of the interaction energy, and furthermore off-shell particle allow for violation. Thus, its is convenient to think of particles as quanta of excitations, but it is only tenable in the free case. If there would be multiple particles or multi-component fields, each would require its own set of such operators.

These creation and annihilation operator therefore allow to change every state into every other state, and therefore can be used to describe every operator, and its action, in the Hilbert space. To have this action, they need to fulfill the canonical commutation relation

$$\left[a_{\vec{p}}, a^\dagger_{\vec{q}}\right] = (2\pi)^3 \delta(\vec{p} - \vec{q}). \tag{10.1}$$

This allows to reexpress the field operators as

$$\phi = \int \frac{d^4p}{(2\pi)^4}(2\pi\ \delta(p^2 - m^2)\big|_{p_0>0}\,(a_p + a^\dagger_{-p})e^{ipx} = \int \frac{d^3\vec{p}}{(2\pi)^3\sqrt{2E_{\vec{p}}}}\left(a_{\vec{p}} + a^\dagger_{-\vec{p}}\right)e^{ipx} \tag{10.2}$$

$$\pi = \int \frac{d^4p}{(2\pi)^4} i(2\pi\ \delta(p^2 - m^2)\big|_{p_0>0}\,(a_p - a^\dagger_{-p})e^{ipx} = \int \frac{d^3\vec{p}}{i(2\pi)^3\sqrt{2E_{\vec{p}}}}\left(a_{\vec{p}} + a^\dagger_{-\vec{p}}\right)e^{ipx} \tag{10.3}$$

where it needs to be noted that the field operators therefore create simultaneously field with opposite momenta of all values, weighted by the place-wave and the energy. This can

[2]It should be noted that in early developments one of these operators was interpreted as to create a particle with negative energy, necessitating the introduction of (Dirac) sea of particles, over which excitation happen. While this picture is still useful in the context of solid state physics to understand particle-hole dynamics, in relativistic field theories that state can actually be identified as an anti-particle, making the whole construction superfluous. It will thus not be detailed here.

be interpreted by noticing

$$\langle 0|\phi|\vec{p}\rangle = e^{i\vec{x}\vec{p}}.$$

Thus the field operator project out a place wave from a state with fixed three-momentum.

More directly, the vacuum expectation value of the two-particle operator remains the propagator, which necessarily has the same form as in section 2.4. But the commutator allows to notice

$$\langle 0\,|[\phi(x),\phi(y)]|\,0\rangle = D(x-y) - D(y-x).$$

This appears to be zero at first. However, as the propagators are really defined as limits, which depend on the causal structure of space-time, this needs to be carefully checked. It is actually found that this is zero only for space-like separation, while an oscillatory contribution, with the mass given the oscillation frequency, remains at time-like separation. In this way, the operator formalism implements the causal structure of Minkowski space-time.

This leads to the concept of the time-order propagator from section 2.4. For this, the advanced and retarded propagators are defined to be

$$\begin{aligned} D_R(x-y) &= \theta(x_0-y_0)\langle 0\,|[\phi(x),\phi(y)]|\,0\rangle \\ D_A(x-y) &= \theta(y_0-x_0)\langle 0\,|[\phi(x),\phi(y)]|\,0\rangle, \end{aligned}$$

respectively, and the Feynman propagator

$$D_F(x-y) = \theta(x_0-y_0)D(x-y) + \theta(y_0-x_0)D(y-x) = \langle 0\,|T\phi(x)\phi(y)|\,0\rangle. \qquad (10.4)$$

These are useful abbreviations in the following.

One of the dramatic differences between the canonical quantization and the path-integral quantization is the commutativity of fields. Since in canonical quantization they are non-commuting operators, their order matters. Especially, given their expansion in terms of oscillator-like operators (10.1) this immediately yields problems as in infinite chains of harmonic oscillators in quantum mechanics. To deal with this, the concept of normal-ordering is introduced.

A normal-ordered product requires all annihilation-type operators are placed to the right, and are thus applied first, and all creation operators last. Thus, a normal-ordered structure will only survive if there are at least sufficient quanta in it to still yield non-zero. Of course, operators at different times commute. Working out the algebra, this leads to Wick's theorem,

$$T(\phi_1(x_1)...\phi_n(x_n)) = N\left(\phi_1(x_1)...\phi_n(x_n) + \text{contractions}\right) \qquad (10.5)$$

where a contraction is defined by a recursive (Wick) contraction operator

$$\begin{aligned} C_m(\phi_1(x_1),...,\phi_m(x_m)) &= \sum_{i=1}^{m-2} C_{m-1-i}(\phi_1(x_1),...,\phi_{i-1}(x_{i-1}))\phi_i(x_i)C_i(\phi_{i+1}(x_{i+1}),...\phi_m(x_m)) \\ C_2(\phi_1(x_1),\phi_2(x_2)) &= D_F^{12}(x_1-x_2) \\ C_1 = C_0 &= 0. \end{aligned}$$

This propagator arises from the fact that any commutation induces a $\mathbb{C}$-number with two operators less, until only two operators are left. But this is then a propagator. The usefulness arises from

$$\langle 0|T\phi_1(x_1)...\phi_n(x_n))|0\rangle = \sum_{\text{permutations}} D_F(x_{i_1}-x_{i_2})...D_F(x_{i_{m-1}}-x_{i_m}).$$

Thus, all vacuum expectation values of time-ordered field operators are equivalent to products of propagators, as the normal-ordered operators will vanish.

This formulation comes in its own in perturbation theory. In that case, the interaction picture of quantum mechanics is particularly useful. Taking the free theory with its known free particles as the unperturbed Hamiltonian, the interaction potential and the time evolution operator in this picture can be determined as usual. Especially, the time evolution operator is now given in terms of the Dyson series

$$U_I = 1+\sum_{n=1}^{\infty}(-i)^n\int d^4z_1\int d^4z_2...\int d^4z_n V_I(z_1)...V_I(z_n). \tag{10.6}$$

But from this, perturbation theory from Feynman rules emerge almost as in section 3.1.

Consider

$$\langle x|y\rangle = \langle\alpha|\phi(x)\phi(y)|\alpha\rangle = \langle 0|U(\infty,x_0)\phi(x)U(x_0,y_0)\phi(y)U(y_0,-\infty)|0\rangle.$$

This is the transition amplitude from a particle located at y to x, which is created by the operators ϕ. More complicated states can be created from sums and products. The state α is the interacting vacuum, which is assumed in perturbation theory to be described by the non-interacting vacuum at infinite times in past and future. This is like the asymptotic non-interacting states in section 3.2. A suitable normalization is assumed here, just as in section 3.1. However, this normalization will turn out to work exactly in the same way, so it will not be considered explicitly.

This statement is true for arbitrary time-orderings. However, the actual answer can be reconstructed from the time-ordered ones in the same way as in (10.4). Also, the full time

evolution operators appear, not the interaction ones. Rewriting tediously everything, the only necessary ingredient will be to determine

$$\langle\alpha|T\phi(x)\phi(y)|\alpha\rangle = \langle 0|T(\phi(x)\phi(y)U_I)|0\rangle.$$

But then the combination of Wick's theorem (10.5) and the expansion of(10.6) automatically recreates the same expression in terms of integrals and propagators as in the path integral formalism, starting with (3.7). The rest follows analogously, and thus the same set of Feynman rules arises, leading to the same perturbation theory.

10.3 Non-relativistic limit and quantum mechanics

How non-relativistic quantum mechanics arises from the quantum field theory has two components. One is the, fairly straightforward, kinematic aspect. Just like in mechanics, momenta, energy, Lorentz transformations etc. will become their non-relativistic counterpart at kinetic energies small compared to the rest mass.

The second aspect is more involved. Quantum mechanics is a theory of point particles, and position operators and momentum operators take over the role of fields. To understand, how this emerges, switch over first to the operator-time-independent picture. Then consider first the basis of creation operator and annihilation operators, as this is a universal basis independent of the underlying theory.

Given (10.1) consider its Fourier transformation,

$$\int \frac{d^3\vec{p}}{(2\pi)^3} e^{-i\vec{x}(\vec{p}-\vec{q})} \left[a_{\vec{p}}, a^\dagger_{\vec{q}}\right] = 1$$

Average in addition Gaussian in momentum space with a suitable width and a linear shift

$$\int \frac{d^3\vec{p}d^3\vec{q}}{\sqrt{\pi^3}(2\pi)^3} e^{-i\vec{x}(\vec{p}-\vec{q})-(\vec{p}-\vec{q})^2} \left[a_{\vec{p}}, a^\dagger_{\vec{q}}\right] = 1.$$

If the a and $a\dagger$ would be independent of the momentum, the left-hand-side would result in Gaussian in $\vec{x}$, centered at zero. Thus, such a smearing of the operators can be interpreted as that if the fields are smeared in position space in a Gaussian way, they regain their non-relativistic commutation relations. The same is, of course, true for the field operators, as (10.2-10.3) are invertible.

This implies that a correlation function like $\langle 0|\phi(x)\phi(y)|0\rangle$ will give a result, which is (time-independent) and, for free particles, Gaussian in $|\vec{x}-\vec{y}|^2$. But this this is the shape for the probability density of a wave-packet positioned at this position. Thus, the interpretation in terms of wave-packets arises again in such an approximation.

Hence, point-particle wave-function emerges at the level of expectation values quite directly, if the fields are smeared with a Gaussian weight. As was noted in section 8.5, this is anyhow the way in which particle-like excitations need to be viewed in quantum-field theory. Since this is done with time-independent operators spatially, this is also preserve the causal structure, if the smearing is located at sufficiently small distances. The time-dependence is then correctly given, once the time-evolution operator is also Gaussian-smeared.

Since the operators a and $a^\dagger$ form a full basis, so do their Gaussian smeared ones. Thus, as for the harmonic oscillator or the Schwinger representation of angular momentum, they can be used to reconstruct operators like the momentum operator and the position operator, as well as their commutation relations. In this way, the basic postulates of quantum mechanics would best be reformulate in terms of this basis.

The one thing, which doe snot necessarily arises directly in this context are again measurements and the Born rule. As noted in section 10.1, this can be reconstructed from the measurement of field operators. However, this lacks the importance of overlaps and the definition of observables in terms of Hermitiean operators. It is not entirely clear how this can arise in a straightforward way. Otherwise, the measurement postulate could be derived from quantum field theory. But it can be heuristically guessed in the following way.

The Gaussian smearing necessarily a sub-region of space-time. In the normal setup of an experiment, this smearing will create an outside in which all operators are exponential suppressed. This patch can be taken to be well approximated by non-relativistic physics, by assumption. All measurement machines reside actually outside of this space-time patch. Thus, their interference with the system can be taken to be instantaneous, but build from operators, which do not belong to the same subspace, giving the appearance of additional operators characterizing the measurement process. Due to their very small temporal overlap, they will instantaneously affect the system by moving around the quanta of the states, which do have overlap, and yielding in this way Born's rule. How exactly this works remains unresolved at the moment, but appears to involve also the fact that this is in reality always an interaction, which still involves many particles, and thus many interactions.

10.4 Axiomatic field theory

After all of these arguments, the question arises, to which extent this could be put also as a foundation, rather than consequence. This is the aim of axiomatic (quantum) field theory,

with the aim to start with a minimal set f postulates, and then derive the remainder. In the same vein, there is also often the attempt to leave a Lagrangian formalism, in favor of starting from the matrix elements, thus to access more general field theories.

The result of these attempts is axiomatic field theory. It starts from the following postulates, the so-called Wightman axioms:

- Like in quantum mechanics, the state space is a Hilbert space, which carries a unitary representation of the Poincaré algebra, and in which the spectrum of the four-momentum operator is time-like with positive energy.
- Fields are maps from Minkowski space-time to a distribution-valued operators in the Hilbert space, and for any such field also its Hermitiean conjugate exists, and they transform as representations of the Poincaré group. Any operator on the Hilbert space can be written in terms of the fields.
- The fields fulfill canonical (anti-)commutator relations at space-like distances.
- A dynamical principle (time evolution operator or transfer matrix) allows to determine fields at infinitesimal time-like distances from fields on a reference spatial hypersurface.

These requires some comments.

In principle, this program formulates more precisely the canonical quantization of section 10.1, by sharpening the concepts of field operators. However, the requirement of a Hilbert space is already for QED problematic, due to consequences of gauge freedom discussed in section 7.3. The next, probably a bit more problematic, statement is that Poincaré-type causality is essentially required as an axiom. Again, this runs afoul with gauge symmetries, and requires a very careful elimination of the latter. In general, also the dynamical principle, to some extent, will run into problems when it comes to questions of space-like correlations.

These questions have only been partially overcome, even in QED. But especially giving up a Hilbert space for the fields to act in turns out to be very complicated without compromising the required locality. It is at the moment not clear, whether giving up some of the axioms or technical issues let the program essentially grind to a halt. And, of course, it not clear a-priori whether such an axiomatic approach can be consistent with nature, even though it had been very successful in non-relativistic physics. Thus, the more ad-hoc approach of the path integral has become the mainstay of relativistic quantum field theory today, and to an appreciable extent also in non-relativistic case.