

Lectures on quantum field theory 1

Stefan Floerchinger and Christof Wetterich

Institut für Theoretische Physik, Philosophenweg 16, D-69120 Heidelberg, Germany

ABSTRACT: Notes for lectures that introduce students of physics to quantum field theory with applications to high energy physics, condensed matter and statistical physics.

Contents

0.1	Organizational issues	1
0.2	Literature	1
0.3	Typos	2
1	Introduction	2
2	Functional integral	2
2.1	Ising model in one dimension	2
2.2	Continuum functional integral	5
2.3	$O(N)$ models in classical statistical equilibrium	6
2.4	Non-linear σ models	10
3	Operators and transfer matrix	11
3.1	Transfer matrix for the Ising model	11
3.2	Non-commutativity in classical statistics	16
3.3	Classical Wave functions	17
4	Quantum Fields and Functional Integral	20
4.1	Phonons as quantum fields in one dimension	20
4.2	Functional integral for quantum fields	24
4.3	Thermodynamic equilibrium	27
4.4	Real time evolution	31
4.5	Expectation values of time ordered operators	33
5	Relativistic scalar fields and $O(N)$-models	38
5.1	Lorentz invariant action and antiparticles	38
5.2	Unified Scalar field theories	40
5.3	Propagator for free field	41
5.4	Magnetisation in classical statistics	43
6	Non-relativistic bosons	45
6.1	Functional integral for spinless atoms	45
6.2	Spontaneous symmetry breaking: Bose-Einstein condensation and superfluidity	47
7	Scattering	48
7.1	Scattering of non-relativistic bosons	49
7.2	The S-matrix	51
7.3	Perturbation theory for interacting scalar fields	53
7.4	From the S-matrix to a cross-section	56
8	Fermions	59
8.1	Non-relativistic fermions	59

9 Lorentz symmetry and the Dirac equation	70
9.1 Lorentz transformations and invariant tensors	70
9.2 Lorentz group	73
9.3 Generators and Lorentz Algebra	75
9.4 Representations of the Lorentz group	77
9.5 Transformation of Fields	79
9.6 Functional Integral, Correlation Functions	81
10 Quantum electrodynamics	85
10.1 Action and propagators	85
10.2 Feynman rules and Feynman diagrams	94
10.3 Elementary scattering processes	98
10.4 Relativistic scattering and decay kinematics	104
10.5 Higgs/Yukawa theory	107

0.1 Organizational issues

There is a webpage to accompany this lecture: <https://uebungen.physik.uni-heidelberg.de/vorlesung/20182/qft1>. Exercises will be proposed every week and discussed in tutorial classes. The registration goes via the webpage above.

0.2 Literature

There is a large amount of literature on different aspects of quantum field theory. Here is only a fine selection.

Statistical field theory / renormalization group

- John Cardy, *Scaling and renormalization in statistical physics* (1996)
- Giorgio Parisi, *Statistical field theory* (1998)
- Jean Zinn-Justin, *Quantum field theory and critical phenomena* (2002)
- Crispin Gardiner, *Handbook of stochastic methods* (1985)

Relativistic quantum field theory

- Mark Srednicki, *Quantum field theory* (2007)
- Michael Peskin & Daniel Schroeder, *An introduction to quantum field theory* (1995)
- Steven Weinberg, *The quantum theory of fields I & II* (1998)

Non-relativistic quantum field theory / condensed matter

- Alexander Altland & Ben Simons, *Condensed matter field theory* (2010)
- Lev Pitaevskii & Sandro Stringari, *Bose-Einstein condensation* (2003)
- Crispin Gardiner & Peter Zoller, *The quantum world of ultra-cold atoms and light* (2014)

Group theory

- Anthony Zee, *Group theory in a nutshell for physicists* (2016)

0.3 Typos

Please send any typos to a.sengupta@thphys.uni-heidelberg.de.

1 Introduction

What is quantum field theory? Historically, quantum field theory (QFT) has been developed as quantum mechanics for many (in fact infinitely many) degrees of freedom. For example, the quantum mechanical description for electromagnetic fields (light) and its excitations, the photons, leads to a quantum field theory. Quantum mechanics of photons, electrons and positrons is quantum electrodynamics (QED) and so one can go on.

In contrast to the transition from classical mechanics to quantum mechanics, the step from there to quantum field theory does not lead to a conceptually entirely new theoretical framework. Still, it was historically not an easy development and a lot of confusion was connected with notions like “second quantization” etc.

There are many new phenomena arising in a field theory setting. This includes collective effects of many degrees of freedom, e. g. spontaneous symmetry breaking. Particle number is not necessarily conserved and one can have particle creation and annihilation.

Historically, quantum field theory has been developed as a *relativistic* theory, which combines quantum mechanics with Lorentz symmetry. This was necessary for quantum electrodynamics. Until today, Lorentz symmetry remains to be a key ingredient for the quantum field theoretic description of elementary particle physics. It is not central for quantum field theory itself, however. Concepts of quantum field theory can also be used to describe the quantum theory of many atoms, for example ultra-cold quantum gases, or phonons in solids, or the spins composing magnets. These systems are treated by non-relativistic QFT.

Probabilistic fields. One may characterize much of the content of the following lectures by two main ingredients

- (i) *Fields* (degrees of freedom at every point x)
- (ii) *Probabilistic theory* (as every quantum theory is one)

In this sense, one may speak of quantum field theory as a probabilistic theory of fields. The reader may note that “quantum” is missing in the above characterization. Indeed, in modern developments, all probabilistic field theories, be they “quantum” or “classical”, are described with the same concepts and methods based on the *functional integral*. The key element here is the one of fluctuating fields as one has it in many situations. Something as tangible as the surface of an ocean is already an example. The concepts are useful in many areas, ranging from statistical mechanics to particle physics, quantum gravity, cosmology, biology, economics and so on. The common view on all these subjects, based on the functional integral, will be the guideline of these lectures.

PFT, *probabilistic field theory*, would be a more appropriate name. We will nevertheless use the traditional, historic name, QFT. Neither “quantum” nor “relativistic” are crucial conceptually. Relativistic quantum field theory is from this perspective an important “special case”, to which we will pay much attention.

2 Functional integral

2.1 Ising model in one dimension

Ising spin. An *Ising spin* has two possible values,

$$s = \pm 1. \tag{2.1}$$

One can also start somewhat more general with some two-level variable with possible values A_1 and A_2 and relate them to the Ising spins via a map,

$$A_1 \rightarrow s = +1, \quad A_2 \rightarrow s = -1. \quad (2.2)$$

For example, a state could be occupied, $n = 1$, or empty, $n = 0$. These states can be mapped to Ising spins via $s = 2n - 1$. From an information theoretic point of view, each Ising spin carries one *bit of information*.

Ising chain. Let us consider a chain of discrete points x and take them to be equidistant,

$$x \in \{x_{\text{in}}, x_{\text{in}} + \varepsilon, x_{\text{in}} + 2\varepsilon, \dots, x_{\text{f}} - \varepsilon, x_{\text{f}}\}. \quad (2.3)$$

Now let us pose one Ising spin at each point or lattice site x and denote its value by $s(x)$. For example,

$$\begin{array}{cccccc|c} 1 & 1 & -1 & -1 & -1 & 1 & -1 & s \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & n \\ \uparrow & \uparrow & \downarrow & \downarrow & \downarrow & \uparrow & \downarrow & \end{array}$$

In general, for P points, or lattice sites, there are $N = 2^P$ possible configurations. We can label them by an index $\tau = 1, \dots, N$. Let us write $\{s(x)\}$ for a configuration of spins on the Ising chain, which should be seen as an abbreviation for $\{s(x_{\text{in}}), s(x_{\text{in}} + \varepsilon), \dots, s(x_{\text{f}})\}$.

Action. We now introduce the concept of an *euclidean action* by assigning to each configuration a real number,

$$\{s(x)\} \rightarrow S(\{s(x)\}) \equiv S[s], \quad \text{where} \quad S \in \mathbb{R}. \quad (2.4)$$

For example, one may have a next neighbor interaction and the action corresponds to

$$S[s] = - \sum_x \beta s(x + \varepsilon) s(x), \quad (2.5)$$

where we use the following abbreviation for a sum over lattice sites

$$\sum_x = \sum_{x=x_{\text{in}}}^{x_{\text{f}}-\varepsilon}, \quad (2.6)$$

and β is a real parameter.

Partition function. One can define a partition function as a sum over all configurations, weighted by the exponential of minus the action,

$$Z = \sum_{\{s(x)\}} e^{-S[s]} = \sum_{\tau} e^{-S_{\tau}}. \quad (2.7)$$

Note that the partition function is here a real and positive number, $Z > 0$.

Probability distribution. Let us now assign to each configuration a probability, $\{s(x)\} \rightarrow p(\{s(x)\}) = p[s]$, or in another notation, $\tau \rightarrow p_{\tau}$. We will set

$$p[s] = \frac{1}{Z} e^{-S[s]}. \quad (2.8)$$

Note the following properties

- (i) positivity $p[s] \geq 0$ (and $p[s] \rightarrow 0$ for $S[s] \rightarrow \infty$),
- (ii) normalization $\sum_{\{s(x)\}} p[s] = \sum_{\tau} p_{\tau} = 1$.

Observables $A[s]$. We may construct an observable by assigning to every configuration τ a value $A_\tau = A[s]$,

$$\{s(x)\} \rightarrow A[s], \quad \tau \rightarrow A_\tau. \quad (2.9)$$

In other words, the observable A has the value A_τ in the configuration τ .

Expectation value. The expectation value of an observable is then given by

$$\langle A \rangle = \sum_{\tau} p_{\tau} A_{\tau} = \frac{1}{Z} \sum_{\{s(x)\}} e^{-S[s]} A[s]. \quad (2.10)$$

Two-point correlation. A correlation function of two observables is given by the expression

$$\langle AB \rangle = \sum_{\tau} p_{\tau} A_{\tau} B_{\tau} = \frac{1}{Z} \sum_{\{s(x)\}} e^{-S[s]} A[s] B[s]. \quad (2.11)$$

Local action. Oftentimes one can write the action as a sum of the form

$$S[s] = \sum_x \mathcal{L}(x), \quad (2.12)$$

with $\mathcal{L}(x)$ depending only on the spins in some neighborhood of x . For our example (2.5) with next neighbor interaction one would have

$$\mathcal{L}(x) = -\beta s(x + \varepsilon)s(x). \quad (2.13)$$

In fact, the simplest version of the traditional *Ising model* has $\beta = \frac{J}{k_B T}$ with interaction parameter J , temperature T and Boltzmann constant k_B . In this context, the Euclidean action corresponds in fact to the ratio $S = \frac{H}{k_B T}$ of Energy or Hamiltonian H and temperature as it appears in the Boltzmann weight factor $\exp(-\frac{H}{k_B T})$. The Hamiltonian is then obviously

$$H = - \sum_x J s(x + \varepsilon)s(x). \quad (2.14)$$

Boundary terms. One must pay some attention to the boundaries of the Ising chain. Let us denote by \mathcal{L}_{in} a term that depends only on $s(x_{\text{in}})$, the *initial spin* and similarly by \mathcal{L}_{f} a term that depends only on $s(x_{\text{f}})$, the *final spin*. We write the action as

$$S = \sum_t \mathcal{L}(t) + \mathcal{L}_{\text{in}} + \mathcal{L}_{\text{f}}. \quad (2.15)$$

By choosing \mathcal{L}_{in} and \mathcal{L}_{f} appropriately one can pose different boundary conditions, in general probabilistic, or also deterministic as an appropriate limit.

Typical problem. A typical problem one may encounter in the context of the Ising model in one dimension is: What is the expectation value $\langle s(x) \rangle$ or the two-point correlation function $\langle s(x_1)s(x_2) \rangle$ for given boundary conditions specified by \mathcal{L}_{in} and \mathcal{L}_{f} ?

Functional integral language. We now formulate the model in a language that is convenient for generalization. We write for expectation values

$$\langle A \rangle = \frac{1}{Z} \int Ds e^{-S[s]} A, \quad (2.16)$$

with the partition function

$$Z = \int Ds e^{-S[s]}. \quad (2.17)$$

The *functional measure* is here defined by

$$\int Ds = \sum_{\{s(x)\}} = \sum_{\tau} = \prod_x \sum_{s(x)=\pm 1}. \quad (2.18)$$

For a finite Ising chain, the functional integral is simply a finite sum over configurations.

2.2 Continuum functional integral

Lattice functional integral. Let us now take a real, continuous variable $\phi(x) \in \mathbb{R}$ instead of the discrete Ising spins $s(x) \in \{+1, -1\}$. The position variable x is for the time being still labeling discrete points or lattice sites. We then define the *functional measure*

$$\int D\phi = \prod_x \int_{-\infty}^{\infty} d\phi(x). \quad (2.19)$$

This is now the continuum version of a sum over configurations. Indeed it sums over all possible functions $\phi(x)$ of the (discrete) position x . To realize that indeed every function appears in $\int D\phi$ one may go back to a discrete variable, $\phi(x) \in \{\phi_1, \dots, \phi_M\}$ with M possible values and take $M \rightarrow \infty$.

Configuration. For every lattice site x we specify now a real number $\phi(x)$ which in total gives then one configuration. Obviously there are now infinitely many configurations even if the number of lattice sites is finite.

Path integral. At this point one can make the transition to a probabilistic *path integral*. To this end one would replace $x \rightarrow t$ and $\phi(x) \rightarrow \vec{x}(t)$, such that the sum over functions $\phi(x)$ becomes one over paths $\vec{x}(t)$. The functional measure would be $\int D\vec{x}$.

Action. The Euclidean action can be written as

$$S = \sum_x \mathcal{L}(x) + \mathcal{L}_{\text{in}} + \mathcal{L}_{\text{f}}, \quad (2.20)$$

where $\mathcal{L}(x)$ depends on $\phi(x')$ with x' in the vicinity of x . Similarly, \mathcal{L}_{in} depends on $\phi(x_{\text{in}}) = \phi_{\text{in}}$ and \mathcal{L}_{f} depends on $\phi(x_{\text{f}}) = \phi_{\text{f}}$.

Lattice ϕ^4 theory. Here we take the action local with

$$\mathcal{L}(x) = \frac{K}{8\varepsilon} [\phi(x + \varepsilon) - \phi(x - \varepsilon)]^2 + \varepsilon V(\phi(x)), \quad (2.21)$$

where the *potential* is given by

$$V(\phi(x)) = \frac{m^2}{2} \phi(x)^2 + \frac{\lambda}{8} \phi(x)^4. \quad (2.22)$$

The partition function is

$$Z = \int D\phi e^{-S[\phi]}, \quad (2.23)$$

and a field expectation value is given by

$$\langle \phi(x) \rangle = \frac{1}{Z} \int Ds e^{-S[\phi]} \phi(x). \quad (2.24)$$

The *functional integral* is here still a finite-dimensional integral where the dimension corresponds to the number of lattice points P . The action $S[\phi]$ is a function of P continuous variables $\phi(x)$.

Continuum limit. Let us now take the limit $\varepsilon \rightarrow 0$ for $x_{\text{f}} - x_{\text{in}}$ fixed. Of course, this means that the number of lattice points P needs to diverge. The “lattice derivative”

$$\partial_x \phi(x) = \frac{1}{2\varepsilon} (\phi(x + \varepsilon) - \phi(x - \varepsilon)) \quad (2.25)$$

becomes a standard derivative, at least for sufficiently smooth configurations, where it exists. One also has

$$\sum_x \varepsilon \rightarrow \int dx, \quad (2.26)$$

and the Euclidean action becomes

$$S = \int dx \mathcal{L}(x) + \mathcal{L}_{\text{in}} + \mathcal{L}_{\text{f}}, \quad (2.27)$$

where now

$$\mathcal{L}(x) = \frac{K}{2} [\partial_x \phi(x)]^2 + V(\phi(x)). \quad (2.28)$$

The first term is called the kinetic term, the second the potential. In the limit $\varepsilon \rightarrow 0$ the action is a *functional* of the functions $\phi(x)$.

Physical observables. As *physical observables* one takes those $A[\phi]$ for which the limit $\langle A \rangle$, $\langle AB \rangle$ and so on exists in the limit $\varepsilon \rightarrow 0$. It will not always be easy to decide whether a given $A[\phi]$ is a physical observable, but the definition is simple. For $\varepsilon \rightarrow 0$ the expression $A[\phi]$ is again a functional.

Functional integral. The functional integral in the continuum theory is now defined as the “continuum limit” of the lattice functional integral for $\varepsilon \rightarrow 0$. By definition, this is well defined for “physical observables”. One may ask: what are such physical observables? The answer to this question is not simple, in general. One should note here that also very rough functions $\phi(x)$ are included in the functional integral, although their contribution is suppressed. If the kinetic term in the Euclidean action $S_{\text{kin}} = \sum_x \frac{K}{8\varepsilon} [\phi(x+\varepsilon) - \phi(x-\varepsilon)]^2$ diverges for $\varepsilon \rightarrow 0$, i. e. $S \rightarrow \infty$, then one has $e^{-S} \rightarrow 0$ and the probability of such configuration vanishes. The corresponding limits may not be trivial, however, because very many rough configurations exist.

Additive rescaling of action. Let us consider a change $S \rightarrow S' = S + C$ or $\mathcal{L}(x) \rightarrow \mathcal{L}'(x) = \mathcal{L}(x) + \tilde{c}$ where $C = (x_{\text{f}} - x_{\text{in}})\tilde{c}$. The partition function changes then like $Z \rightarrow Z' = e^{-C}Z$. Similarly,

$$\int D\phi e^{-S} A[\phi] \rightarrow e^{-C} \int D\phi e^{-S} A[\phi]. \quad (2.29)$$

But this means that C drops out when one considers expectation values like $\langle A \rangle$! It can even happen that C diverges for $\varepsilon \rightarrow 0$ such that formally $Z \rightarrow 0$ or $Z \rightarrow \infty$. But this is not a problem because the absolute value of Z is irrelevant. The probability distribution $p[\phi] = \frac{1}{Z} e^{-S[\phi]}$ is unchanged.

2.3 $O(N)$ models in classical statistical equilibrium

Classical thermal fluctuations. For the time being we are concerned with static (equilibrium) aspects of field theory models at finite temperature. These field theories can arise for example from a lattice model such as the Ising model if the latter is probed on distances that are large against the typical microscopic scale or inter-particle distance ε . Formally one can then take the limit $\varepsilon \rightarrow 0$ as discussed in the previous subsection. It turns out (and will become more clear latter on), that in such a situation *classical* thermal fluctuations dominate over quantum fluctuations. We discuss here therefore *classical statistical field theories* in thermal equilibrium.

Such theories have a probabilistic description in terms of functional integrals with weight given by the Boltzmann factor $e^{-\beta H}$. Here $\beta = 1/T$ and we use now units where $k_B = 1$ such that temperature is measured in units of energy. In the following we will discuss possible forms of the field theory and in particular the Hamiltonian H .

Universality classes and models. In condensed matter physics, microscopic Hamiltonians are often not very well known and if they are, they are not easy to solve. However, in particular in the vicinity of second order phase transitions, there are some universal phenomena that are independent of the precise microscopic physics. This will be discussed in more detail later on, in the context of the *renormalization group*. Essentially, this arises as a consequence of thermal fluctuations and the fact that at a second order phase transition fluctuations are important on all scales. Roughly speaking, a theory changes in form when fluctuations are taken into account and can approach a largely universal *scaling form* for many different microscopic theories that happen to be in the same *universality class*.

In the following we will discuss a class of *model systems*. These are particularly simple field theories for which one can sometimes answer certain questions analytically, but one can also see them as representatives for their respective universality classes. In the context of quantum field theory, we will see that these field theory models gain a substantially deeper significance.

Scalar $O(N)$ models in d dimensions. Let us consider models of the form

$$\beta H[\phi] = S[\phi] = \int d^d x \left\{ \frac{1}{2} \partial_j \phi_n \partial_j \phi_n + \frac{1}{2} m^2 \phi_n \phi_n + \frac{1}{8} \lambda (\phi_n \phi_n)^2 \right\}. \quad (2.30)$$

Here, $\phi_n = \phi_n(x)$ with $n = 1, \dots, N$ are the fields. We use Einsteins summation convention which implies that indices that appear twice are summed over. We have formulated the theory in d spatial dimensions (where in practice $d = 3, 2, 1$ or even 0 for condensed matter systems). The index j is accordingly summed in the range $j = 1, \dots, d$. Although not very precise, one sometimes calls $S[\phi]$ the *Euclidean microscopic action*. The square brackets indicate here that the action depends on the fields in a functional way, which means it depends not on single numbers but on the entire set of *functions* of space $\phi_n(x)$, with $x \in \mathbb{R}^d$ and $n = 1, \dots, N$.

Fields as vectors. One can consider $\phi_n(x)$ as a vector in a vector space of infinite dimension where components are labeled by the spatial position x and the discrete index n . If in doubt, one can go back to a lattice model where x is discrete.

Applications. Models of the type (2.30) have many applications. For $N = 1$ they correspond in a certain sense to the continuum limit of the Ising model. For $N = 2$ the model can equivalently be described by complex scalar fields. It has then applications to Bose-Einstein condensates, for example. For $N = 3$ and $d = 3$ one can have situations where the rotation group and the internal symmetry group are coupled. This describes then vector fields, for example magnetization. Finally, for $N = 4$ and $d = 4$, the model essentially describes the Higgs field after a Wick rotation to Euclidean space.

Engineering dimensions. In equation (2.30) we have rescaled the fields such that the coefficient of the derivative term is $1/2$. This is always possible. It is useful to investigate the so-called *engineering scaling dimension* of the different terms appearing in (2.30). The combination βH or the action S must be dimensionless. Derivatives have dimension of inverse length $[\partial] = L^{-1}$ and the fields must accordingly have dimension $[\phi] = L^{-\frac{d}{2}+1}$. One also finds $[m] = L^{-1}$ and $[\lambda] = L^{d-4}$. Note in particular that λ is dimensionless in $d = 4$ dimensions.

Symmetries. It is useful to discuss the symmetries of the model (2.30). Symmetries are (almost) always very helpful in theoretical physics. In the context of statistical field theory, they are useful as a guiding principle in particular because they still survive (in a sense to be defined) when the effect of fluctuations is taken into account.

For the model (2.30) we have a space symmetry group consisting of rotations and translations, as well as a continuous, so-called internal symmetry group of global $O(N)$ transformations. We now discuss them step-by-step.

Rotations. Rotations in space are transformations of the form

$$x^j \rightarrow x'^j = R^{jk} x^k. \quad (2.31)$$

The matrices R fulfill the condition $R^T R = 1$ and we demand that they connect smoothly to the unit matrix $R = \mathbb{1}$. This fixes $\det(R) = 1$. Matrices of this type in d spatial dimensions form a group, the *special orthogonal group* $SO(d)$. Mathematically, this is a *Lie group* which implies that all group elements can be composed of many infinitesimal transformations. An infinitesimal transformation can be written as

$$R^{jk} = \delta^{jk} + \frac{i}{2} \delta\omega_{mn} J_{(mn)}^{jk}, \quad (2.32)$$

where $J_{(mn)}^{jk} = -i(\delta_{mj}\delta_{nk} - \delta_{mk}\delta_{nj})$ are the *generators of the Lie algebra* and $\delta\omega_{mn}$ are infinitesimal, anti-symmetric matrices. One may easily count that there are $d(d-1)/2$ independent components of an anti-symmetric matrix in d dimensions and as many generators. Finite group elements can be obtained as

$$R = \lim_{N \rightarrow \infty} \left(\mathbb{1} + \frac{i}{2} \frac{\omega_{mn}}{N} J_{(mn)} \right)^N = \exp \left(\frac{i}{2} \omega_{mn} J_{(mn)} \right). \quad (2.33)$$

Let us now work out how fields transform under rotations. We will implement them such that a field configuration with a maximum at some position x before the transformation will have a maximum at Rx afterwards. The field must transform as

$$\phi_n(x) \rightarrow \phi'_n(x) = \phi_n(R^{-1}x). \quad (2.34)$$

Note that derivatives transform as

$$\partial_j \phi_n(x) \rightarrow (R^{-1})_{kj} (\partial_k \phi_n)(R^{-1}x) = R_{jk} (\partial_k \phi_n)(R^{-1}x). \quad (2.35)$$

The brackets should denote that the derivatives are with respect to the full argument of ϕ_n and we have used the chain rule. The action in (2.30) is invariant under rotations acting on the fields, as one can confirm easily. Colloquially speaking, no direction in space is singled out.

Translations. Another useful symmetry transformations are translations $x \rightarrow x + a$. The fields get transformed as

$$\phi_n(x) \rightarrow \phi'_n(x) = \phi_n(x - a). \quad (2.36)$$

One easily confirms that the action (2.30) is also invariant under translations. Colloquially speaking, this implies that no point in space is singled out.

Global internal $O(N)$ symmetry. There is another useful symmetry of the action (2.30) given by rotations (and mirror reflections) in the “internal” space of fields,

$$\phi_n(x) \rightarrow O_{nm} \phi_m(x). \quad (2.37)$$

The matrices O_{nm} are here independent of the spatial position x (therefore this is a *global* and not a *local* transformation) and they satisfy $O^T O = \mathbb{1}$. Because we do not demand them to be smoothly connected to the unit matrix, they can have determinant $\det(O) = \pm 1$. These matrices are part of the *orthogonal group* $O(N)$ in N dimensions. It is an easy exercise to show that the action (2.30) is indeed invariant under these transformations.

Partition function. The partition function for the model (2.30) reads

$$Z[J] = \int D\phi e^{-S[\phi] + \int d^d x \{J_n(x)\phi_n(x)\}} \quad (2.38)$$

We have introduced here an external source term $\int d^d x \{J_n(x)\phi_n(x)\}$ which can be used to probe the theory in various ways. For example, one can take *functional derivatives* to calculate expectation values,

$$\langle \phi_n(x) \rangle = \frac{1}{Z[J]} \frac{\delta}{\delta J_n(x)} Z[J] \Big|_{J=0}, \quad (2.39)$$

and correlation functions, e. g.

$$\langle \phi_n(x)\phi_m(y) \rangle = \frac{1}{Z[J]} \frac{\delta^2}{\delta J_n(x)\delta J_m(y)} Z[J] \Big|_{J=0} = \frac{\int D\phi \phi_n(x)\phi_m(y) e^{-S[\phi]}}{\int D\phi e^{-S[\phi]}}. \quad (2.40)$$

Classical field equation. Note that in the the functional integral, field configurations $\phi(x)$ are suppressed, if the corresponding action $S[\phi]$ is large. In the partition function (2.38), large contributions come mainly from the region around the minima of $S[\phi] - \int_x J_n\phi_n$, which are determined by the equation

$$\frac{\delta}{\delta \phi_n(x)} \left(S[\phi] - \int d^d x \{J_n(x)\phi_n(x)\} \right) = \frac{\delta S[\phi]}{\delta \phi_n(x)} - J_n(x) = 0. \quad (2.41)$$

Note that this equation resembles the equation of motion of a classical field theory. For the model (2.30) one has concretely

$$\frac{\delta S[\phi]}{\delta \phi_n(x)} - J_n(x) = -\partial_j \partial_j \phi_n(x) + m^2 \phi_n(x) + \frac{1}{2} \lambda \phi_n(x) \phi_k(x) \phi_k(x) - J_n(x) = 0. \quad (2.42)$$

Note that this field equation is from a mathematical point of view a second order, semi-linear, partial differential equation. It contains non-linear terms in the fields ϕ_n , but the term involving derivatives is linear; therefore semi-linear. The equation involves the Euclidean Laplace operator $\Delta = \partial_j \partial_j$ and is therefore of elliptic type (as opposed to hyperbolic or parabolic).

The $O(N)$ symmetric potential. The model in (2.30) can be generalized somewhat to the action

$$S[\phi] = \int d^d x \left\{ \frac{1}{2} \partial_j \phi_n \partial_j \phi_n + V(\rho) \right\}, \quad (2.43)$$

where $\rho = \frac{1}{2} \phi_n \phi_n$ is an $O(N)$ symmetric combination of fields and $V(\rho)$ is the microscopic $O(N)$ symmetric potential. Of course, the previous case (2.30) can be recovered for $V(\rho) = m^2 \rho + \frac{1}{2} \lambda \rho^2$.

More general, $V(\rho)$ might be some function with a minimum at ρ_0 and a Taylor expansion around it,

$$V(\rho) = m^2(\rho - \rho_0) + \frac{1}{2} \lambda (\rho - \rho_0)^2 + \frac{1}{3!} \gamma (\rho - \rho_0)^3 + \dots \quad (2.44)$$

If the minimum is positive, $\rho_0 > 0$, the linear term vanishes of course, $m^2 = 0$. In contrast, if the minimum is at $\rho_0 = 0$ one has in general $m^2 > 0$. Note that it costs a certain amount of energy for the field to move away from the minimum. In particular, for large λ such configurations are suppressed.

Homogeneous solutions. It is instructive to discuss homogeneous solutions of the field equation, i.e. solutions that are independent of the space variable x . For vanishing source $J_n(x) = 0$, and the model (2.43) we need to solve

$$\frac{\partial}{\partial \phi_n} V(\rho) = \phi_n \frac{\partial}{\partial \rho} V(\rho) = 0. \quad (2.45)$$

This has always a solution $\phi_n = 0$ and for $\rho_0 = 0$ and positive m^2 this is indeed a minimum of the action $S[\phi]$. For positive ρ_0 the situation is more interesting, however. In that case, $\phi_n = 0$ is actually typically a maximum while the minimum is at $\phi_k \phi_k = 2\rho_0$, i. e. at a non-zero field value. One possibility is $\phi_1 = \sqrt{2\rho_0}$ with $\phi_2 = \dots = \phi_n = 0$, but there are of course many more. But such a solution breaks the $O(N)$ symmetry! One says that the $O(N)$ symmetry is here *spontaneously broken on the microscopic level* which technically means that the action $S[\phi]$ is invariant, but the solution to the field equation (i. e. the minimum of $S[\phi]$) breaks the symmetry. It is an interesting and non-trivial question whether the symmetry breaking survives the effect of fluctuations. One has proper *macroscopic* spontaneous symmetry breaking if the field expectation value $\langle \phi_n \rangle$ is non-vanishing and singles out a direction in field space. An example for spontaneous symmetry breaking is the magnetization field in a ferromagnet.

2.4 Non-linear σ models

Constrained fields. It is also interesting to consider models where $\rho = \rho_0$ is fixed. In fact, they arise naturally in the low energy limit of the models described above when the fields do not have enough energy to climb up the effective potential. Technically, this corresponds here to the limit $\lambda \rightarrow \infty$ with fixed ρ_0 and can be implemented as a constraint

$$\phi_n(x)\phi_n(x) = 2\rho_0. \quad (2.46)$$

Note that with this constraint, the field is now living on a manifold corresponding to the surface of an N -dimensional sphere, denoted by S_{N-1} . One can parametrize the field as

$$\phi_1 = \sigma, \quad \phi_2 = \pi_1, \quad \dots \quad \phi_N = \pi_{N-1}, \quad (2.47)$$

where only the fields π_n are independent while σ is related to them via the non-linear constraint

$$\sigma = \sqrt{2\rho_0 - \vec{\pi}^2}. \quad (2.48)$$

Linear and non-linear symmetries. The symmetry group $O(N)$ falls now into two parts. The first consists of transformations $O(N-1)$ which only act on the fields π_n but do not change the field σ . Such transformations are realized in the standard, linear way

$$\pi_n \rightarrow O_{nm}^{(N-1)} \pi_m, \quad \sigma \rightarrow \sigma. \quad (2.49)$$

In addition to this, there are transformations in the complement part of the group (rotations that also involve the first component σ). They act infinitesimally on the independent fields like

$$\delta \pi_n = \delta \alpha_n \sigma = \delta \alpha_n \sqrt{2\rho_0 - \vec{\pi}^2}, \quad \delta \sigma = -\delta \alpha_n \pi_n, \quad (2.50)$$

where $\delta \alpha_n$ are infinitesimal parameters (independent of the fields). Note that this is now a non-linearly realized symmetry in the internal space of fields. This explains also the name *non-linear sigma model*.

Action. Let us now write an action for the non-linear sigma model. Because of the constraint (2.46), the effective potential term in (2.43) becomes irrelevant and only the kinetic term remains,

$$S[\pi] = \int d^d x \left\{ \frac{1}{2} \partial_j \phi_n \partial_j \phi_n \right\} = \int d^d x \left\{ \frac{1}{2} G_{mn}(\vec{\pi}) \partial_j \pi_m \partial_j \pi_n \right\}. \quad (2.51)$$

In the last equation we rewrote the action in terms of the independent fields π_n and introduced the *metric in the field manifold*

$$G_{mn}(\vec{\pi}) = \delta_{mn} + \frac{\pi_m \pi_n}{2\rho_0 - \vec{\pi}^2}. \quad (2.52)$$

The second term originates from

$$\partial_j \sigma = \partial_j \sqrt{2\rho_0 - \vec{\pi}^2} = \frac{1}{\sqrt{2\rho_0 - \vec{\pi}^2}} \pi_m \partial_j \pi_m. \quad (2.53)$$

Functional integral. Note that also the functional integral is now more complicated. It must involve the determinant of the metric G_{mn} to be $O(N)$ invariant. For a single space point x one has

$$\int \prod_n d\phi_n \rightarrow \int \prod_n d\phi_n \delta(\phi_n^2 - 2\rho_0) = \text{const} \times \int \sqrt{\det(G(\vec{\pi}))} \prod_n d\pi_n. \quad (2.54)$$

Accordingly, the functional integral must be adapted.

Ising model. Everything becomes rather simple again for $N = 1$. The constraint $\phi(x)^2 = 2\rho_0$ allows only the field values $\phi(x) = \pm\sqrt{2\rho_0}$. On a discrete set of space points (a lattice), this leads us back to the Ising model.

3 Operators and transfer matrix

Our lecture will be based on the discussion of functional integrals. These are a generalization of ordinary, multi-dimensional integrals to the limit of infinitely many degrees of freedom, i. e. infinite dimensional integrals. For bosons, the variables or fields all commute. (For fermions we will later use the anti-commuting Grassmann variables). One has learned that non-commuting operators play a crucial role in quantum mechanics. These non-commuting structures are not directly visible in the bosonic functional integral which only contains commuting quantities. One may wonder how such integrals can describe the non-commutative properties of quantum mechanics. The next two lectures are devoted to reveal the structural relation between the operator formalism, known from quantum mechanics and the functional integral.

3.1 Transfer matrix for the Ising model

Boundary problem for Ising chain. Let us consider the one-dimensional Ising model

$$S = \sum_x \mathcal{L}(x) + \mathcal{L}_{\text{in}} + \mathcal{L}_{\text{f}}, \quad (3.1)$$

with

$$\mathcal{L}(x) = -\beta s(x + \varepsilon) s(x). \quad (3.2)$$

We choose boundary conditions such that $s(x_{\text{in}}) = 1$, $s(x_{\text{f}}) = 1$. This can be implemented by

$$e^{-\mathcal{L}_{\text{in}}} = \delta(s(x_{\text{in}}) - 1), \quad e^{-\mathcal{L}_{\text{f}}} = \delta(s(x_{\text{f}}) - 1), \quad (3.3)$$

which in turn can be implemented by limits like

$$\mathcal{L}_{\text{in}} = - \lim_{\kappa \rightarrow \infty} \kappa (s(x_{\text{in}}) - 1). \quad (3.4)$$

Question: What is the expectation value $\langle s(x) \rangle$ for x in the bulk, i. e. between the endpoints x_{in} and x_{f} ? The single configuration with minimal action has all spins aligned, $s(x) = 1$. There are, however, many more configurations where some of the spins take negative values. Even though the particular probability for one such configuration is smaller, this is outweighed by the number of configurations. Qualitatively one expects something like in figure 1.

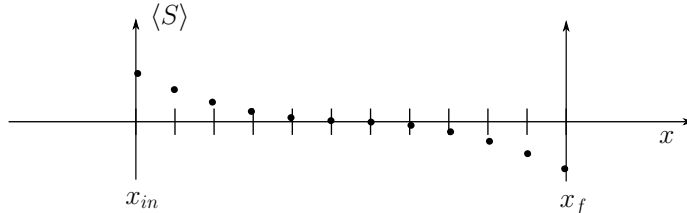


Figure 1. Ising chain with spins at the endpoints fixed to $s(x_{\text{in}}) = 1$ and $s(x_{\text{f}}) = 1$. What is $\langle s(x) \rangle$ for x between the endpoints?

In the bulk, far away from the boundaries, the average spin may vanish to a good approximation. We look for a formalism to compute this behaviour as a function of the parameter β .

Product form of probability distribution. We can write e^{-S} in product form

$$e^{-S} = e^{-\mathcal{L}_{\text{f}} + \sum_x \mathcal{L}(x) + \mathcal{L}_{\text{in}}} = \bar{f}_{\text{f}} \left[\prod_x e^{-\mathcal{L}(x)} \right] f_{\text{in}} = \bar{f}_{\text{f}} \left[\prod_x \mathcal{K}(x) \right] f_{\text{in}} \quad (3.5)$$

with boundary terms $\bar{f}_{\text{f}} = e^{-\mathcal{L}_{\text{f}}}$ and $f_{\text{in}} = e^{-\mathcal{L}_{\text{in}}}$. Here $\mathcal{K}(x)$ depends on the two spins $s(x)$ and $s(x + \varepsilon)$, while f_{in} depends on $s(x_{\text{in}})$ and \bar{f}_{f} depends on $s(x_{\text{f}})$.

Occupation number basis. Any function $f(s(x))$ that depends only on the spin $s(x)$ can be expanded in terms of two basis functions $h_{\tau}(s(x))$ where $\tau = 1, 2$,

$$f(s(x)) = q_1(x) h_1(s(x)) + q_2(x) h_2(s(x)). \quad (3.6)$$

We choose the occupation number basis with

$$h_1(s) = \frac{1+s}{2} = n, \quad h_2(s) = \frac{1-s}{2} = (1-n). \quad (3.7)$$

This is easily seen by noting that the occupation number n has only the values 1 (for $s = 1$) and 0 (for $s = -1$), such that

$$n^2 = n. \quad (3.8)$$

Any polynomial in s can be written as $an + b$, such that any $f(s)$ can indeed be expressed in terms of the two basis functions.

We note some properties of the basis functions. The relation

$$h_{\tau}(s)h_{\rho}(s) = \delta_{\tau\rho}h_{\rho}(s) \quad (3.9)$$

is simply verified by $h_{\tau}^2(s) = h_{\tau}(s)$ and $h_1(s)h_2(s) = 0$. Other useful relations are

$$\sum_{s=\pm 1} h_{\tau}(s) = h_{\tau}(s=1) + h_{\tau}(s=-1) = 1, \quad (3.10)$$

$$\sum_{\tau} h_{\tau}(s) = h_1(s) + h_2(s) = 1, \quad (3.11)$$

and finally by combination

$$\sum_{s=\pm 1} h_{\tau}(s)h_{\rho}(s) = \delta_{\tau\rho}. \quad (3.12)$$

Transfer matrix. Let us now expand $\mathcal{K}(x)$ in terms of the basis functions $h_\tau(s(x+\varepsilon))$ and $h_\rho(s(x))$,

$$\mathcal{K}(x) = \hat{T}_{\tau\rho}(x) h_\tau(s(x+\varepsilon)) h_\rho(s(x)). \quad (3.13)$$

We use here the Einstein summation convention which implies summation over the indices τ and ρ . The expansion coefficients $\hat{T}_{\tau\rho}(x)$ are the elements of the transfer matrix \hat{T} . This is a 2×2 matrix. Indeed using shorthands $\bar{n} = n(t+\varepsilon)$, $n = n(t)$ and similar for \bar{h}_τ , h_τ , an arbitrary $\mathcal{K}(x)$ can be written as

$$\begin{aligned} \mathcal{K} &= a\bar{n}n + b\bar{n} + cn + d \\ &= \hat{T}_{11}\bar{h}_1h_1 + \hat{T}_{12}\bar{h}_1h_2 + \hat{T}_{21}\bar{h}_2h_1 + \hat{T}_{22}\bar{h}_2h_2. \end{aligned} \quad (3.14)$$

Matrix product for transfer matrix. Consider now the product of two neighbouring factors $\mathcal{K}(s+\varepsilon)$ and $\mathcal{K}(x)$, summed over the common spin $s(x+\varepsilon)$

$$\begin{aligned} \sum_{s(x+\varepsilon)} \mathcal{K}(x+\varepsilon)\mathcal{K}(x) &= \sum_{s(x+\varepsilon)} h_\tau(s(x+2\varepsilon))\hat{T}_{\tau\rho}(x+\varepsilon)h_\rho(s(x+\varepsilon))h_\alpha(s(x+\varepsilon))\hat{T}_{\alpha\beta}(x)h_\beta(s(x)) \\ &= \sum_\rho \sum_{s(x+\varepsilon)} h_\tau(s(x+2\varepsilon))\hat{T}_{\tau\rho}(x+\varepsilon)\hat{T}_{\rho\beta}(x)h_\rho(s(x+\varepsilon))h_\beta(s(x)) \\ &= \sum_\rho h_\tau(s(x+2\varepsilon))\hat{T}_{\tau\rho}(x+\varepsilon)\hat{T}_{\rho\beta}(x)h_\beta(s(x)) \\ &= h_\tau(s(x+2\varepsilon)) \left[\hat{T}(x+\varepsilon)\hat{T}(x) \right]_{\tau\beta} h_\beta(s(x)). \end{aligned} \quad (3.15)$$

The second line uses $h_\tau h_\rho = \delta_{\tau\rho} h_\rho$ and the third line $\sum_s h_\rho = 1$. We observe that the matrix product of transfer matrices appears in this product. For the Ising model we have that $\mathcal{K}(x)$ is the same for all x (except for different spins being involved), and therefore \hat{T} is independent of x . One simply finds

$$\sum_{s(x+\varepsilon)} \mathcal{K}(x+\varepsilon)\mathcal{K}(x) = h_\tau(s(x+2\varepsilon)) \left[\hat{T}^2 \right]_{\tau\rho} h_\rho(s(x)). \quad (3.16)$$

Doing one more similar step yields

$$\sum_{s(x+2\varepsilon)} \sum_{s(x+\varepsilon)} \mathcal{K}(x+2\varepsilon)\mathcal{K}(x+\varepsilon)\mathcal{K}(x) = h_\tau(s(x+3\varepsilon)) \left[\hat{T}(x+2\varepsilon)\hat{T}(x+\varepsilon)\hat{T}(x) \right]_{\tau\rho} h_\rho(s(x)), \quad (3.17)$$

and so one can go on.

Partition function as product of transfer matrices. One can write the partition function as

$$\begin{aligned} Z &= \left[\prod_{x=x_{\text{in}}}^{x_{\text{f}}} \sum_{s(x)} \right] \bar{f}_{\text{f}}(s(x_{\text{f}})) \left[\prod_{x=x_{\text{in}}}^{(x_{\text{f}}-\varepsilon)} \mathcal{K}(x) \right] f_{\text{in}}(s(x_{\text{in}})) \\ &= \sum_{s(x_{\text{f}})} \sum_{s(x_{\text{in}})} \bar{f}_{\text{f}}(s(x_{\text{f}})) h_\tau(s(x_{\text{f}})) \left[\hat{T}(x_{\text{f}}-\varepsilon) \cdots \hat{T}(x_{\text{in}}) \right]_{\tau\rho} h_\rho(s(x_{\text{in}})) f_{\text{in}}(s(x_{\text{in}})) \\ &= \sum_{s(x_{\text{f}})} \sum_{s(x_{\text{in}})} \bar{q}_\beta(x_{\text{f}}) h_\beta(s(x_{\text{f}})) h_\tau(s(x_{\text{f}})) \left[\hat{T} \cdots \hat{T} \right]_{\tau\rho} h_\rho(s(x_{\text{in}})) \tilde{q}_\alpha(s(x_{\text{in}})) h_\alpha(s(x_{\text{in}})). \end{aligned} \quad (3.18)$$

Here we have expanded \bar{f}_{f} and f_{in} in terms of the basis functions,

$$\begin{aligned} \bar{f}_{\text{f}}(s(x_{\text{f}})) &= \bar{q}_\beta(x_{\text{f}}) h_\beta(s(x_{\text{f}})), \\ f_{\text{in}}(s(x_{\text{in}})) &= \tilde{q}_\alpha(x_{\text{in}}) h_\alpha(s(x_{\text{in}})). \end{aligned}$$

Performing the sums over the initial and final spins leads to

$$Z = \bar{q}_\tau(x_f) \left[\hat{T}(x_f - \varepsilon) \cdots \hat{T}(x_{\text{in}}) \right]_{\tau\rho} \tilde{q}_\rho(x_{\text{in}}). \quad (3.19)$$

This has the structure of an initial vector (or wave function) $\tilde{q}(x_{\text{in}})$ multiplied by a matrix, and then contracted with a final vector (or conjugate wave function) $\bar{q}(x_f)$. We can use the bracket notation familiar from quantum mechanics,

$$Z = \langle \bar{q}(x_f) | \hat{T}(x_f - \varepsilon) \cdots \hat{T}(x_{\text{in}}) | \tilde{q}(x_{\text{in}}) \rangle. \quad (3.20)$$

This product formulae resembles quantum mechanics if one associates the transfer matrix with the infinitesimal evolution operator $U(t)$

$$\psi(t + \varepsilon) = U(t)\psi(t), \quad (3.21)$$

where

$$U(t) = e^{i\varepsilon H(t)}. \quad (3.22)$$

With

$$\psi(t_f) = U(t_f - \varepsilon) \cdots U(t_{\text{in}})\psi(t_{\text{in}}), \quad (3.23)$$

one can write the amplitude in the form

$$\langle \phi(t_f) | \psi(t_f) \rangle = \langle \phi(t_f) | U(t_f - \varepsilon) \cdots U(t_{\text{in}}) | \psi(t_{\text{in}}) \rangle. \quad (3.24)$$

Formally, the map between quantum mechanics and the classical statistics of the Ising model is

QM	CS
U	\hat{T}
t	x
ψ	\tilde{q}
$\bar{\phi}$	\bar{q}

A main difference to quantum mechanics is that \hat{T} does not preserve the norm of the wave function.

Computation of transfer matrix. We employ the defining relation of the transfer matrix,

$$e^{\beta \bar{s}s} = \hat{T}_{\tau\rho} h_\tau(\bar{s}) h_\rho(s), \quad (3.25)$$

where we use the shorthand notation

$$\bar{s} = s(x + \varepsilon), \quad s = s(x). \quad (3.26)$$

Using the decomposition

$$s = h_1 - h_2 = n - (1 - n) = 2n - 1, \quad (3.27)$$

and

$$\beta \bar{s}s = \beta(\bar{h}_1 - \bar{h}_2)(h_1 - h_2) = \beta(\bar{h}_1 h_1 + \bar{h}_2 h_2 - \bar{h}_1 h_2 - \bar{h}_2 h_1), \quad (3.28)$$

one obtains by analyzing the possible cases,

$$e^{\beta \bar{s}s} = e^\beta (\bar{h}_1 h_1 + \bar{h}_2 h_2) + e^{-\beta} (\bar{h}_1 h_2 + \bar{h}_2 h_1). \quad (3.29)$$

From this one can read off the transfer matrix

$$\hat{T} = \begin{pmatrix} e^\beta & e^{-\beta} \\ e^{-\beta} & e^\beta \end{pmatrix}. \quad (3.30)$$

Note that, in general, the transfer matrix \hat{T} is not a unitary matrix as for quantum mechanics. For the Ising model $\hat{T}(x)$ does not depend on x so that one obtains

$$Z = \bar{q}_\tau(x_f) \left[\hat{T}^{P-1} \right]_{\tau\rho} \tilde{q}_\rho(x_{\text{in}}). \quad (3.31)$$

Periodic Boundary Condition. Replace $\mathcal{L}_f + \mathcal{L}_{\text{in}}$ by $-\beta s(x_f)s(s_{\text{in}})$. This closes the circle by defining x_f and x_{in} as next neighbours. The partition function becomes

$$Z = \text{Tr} \left\{ \hat{T}^P \right\}. \quad (3.32)$$

Diagonalising \hat{T} solves the Ising model in a simple way,

$$Z = \lambda_+^P + \lambda_-^P, \quad (3.33)$$

with λ_{\pm} the two eigenvalues of the transfer matrix,

$$\lambda_+ = 2 \cosh(\beta), \quad \lambda_- = 2 \sinh(\beta). \quad (3.34)$$

In the limit $P \rightarrow \infty$ only the largest eigenvalue λ_+ contributes.

Generalisations. The transfer matrix can be generalised to an arbitrary number of Ising spins $s_{\gamma}(x)$. For M spins, $\gamma = 1, \dots, M$, the transfer matrix \hat{T} is an $N \times N$ matrix, $N = 2^M$, $\tau = 1, \dots, N$.

For example, if $M = 2$, \hat{T} is a 4×4 matrix. The basis functions in the occupation number basis are taken as

$$\begin{aligned} h_1 &= n_1 n_2, & h_2 &= (1 - n_1) n_2, \\ h_3 &= n_1 (1 - n_2), & h_4 &= (1 - n_1) (1 - n_2). \end{aligned} \quad (3.35)$$

This structure can be extended to arbitrary M . The basis functions obey the same rules as discussed for $M = 1$. In particular, γ may denote a second coordinate y such that,

$$s_{\gamma}(x) \rightarrow s(x, y). \quad (3.36)$$

In this way one can define formally the transfer matrix for the two-dimensional Ising model. The coordinate x denotes now lines in a two-dimensional plane, see fig. 2.

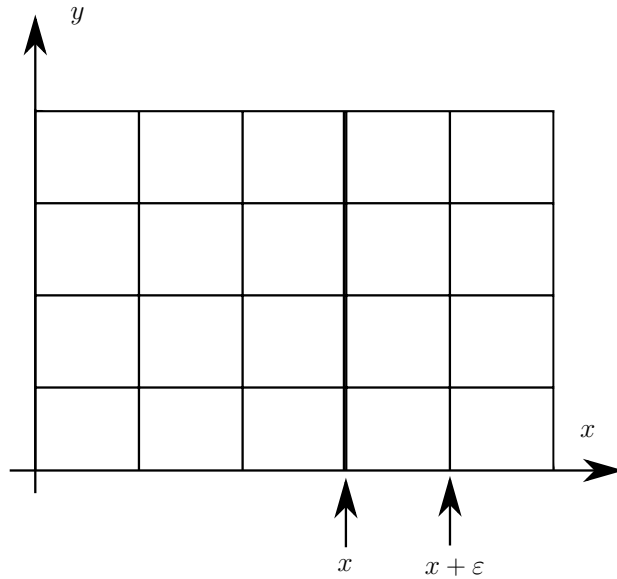


Figure 2. Illustration of the two dimensional Ising model.

More generally, in d dimensions, x denotes the partition of a particular $d - 1$ dimensional hypersurface. The transfer matrix contains the information of what happens if one goes from one hypersurface to the next one.

3.2 Non-commutativity in classical statistics

Local observables and operators. A local observable $A(x)$ depends only on the local spin $s(x)$. We want to find an expression for its expectation value in terms of the transfer matrix. For this purpose we consider the expression

$$\sum_{s(x)} \mathcal{K}(x) A(x) \mathcal{K}(x - \varepsilon) = \sum_{s(x)} h_\tau(x + \varepsilon) \hat{T}_{\tau\rho}(x) h_\rho(x) A_\gamma(x) h_\gamma(x) h_\alpha(x) \hat{T}_{\alpha\beta}(x - \varepsilon) h_\beta(x - \varepsilon), \quad (3.37)$$

where we use the shorthand

$$h_\tau(x) = h_\tau(s(x)), \quad (3.38)$$

and the expansion

$$A(x) = A_\gamma(x) h_\gamma(s(x)). \quad (3.39)$$

We employ

$$A_\gamma(x) \sum_{s(x)} h_\rho(x) h_\gamma(x) h_\alpha(x) = \sum_\gamma A_\gamma(x) \delta_{\rho\gamma} \delta_{\gamma\alpha}, \quad (3.40)$$

and introduce the diagonal operator

$$(\hat{A}(x))_{\rho\alpha} = \sum_\gamma A_\gamma(x) \delta_{\rho\gamma} \delta_{\gamma\alpha} = \begin{pmatrix} A_1(x) & 0 \\ 0 & A_2(x) \end{pmatrix}. \quad (3.41)$$

In terms of this operator we can write

$$\sum_{s(x)} \mathcal{K}(x) A(x) \mathcal{K}(x - \varepsilon) = h_\tau(x + \varepsilon) \hat{T}_{\tau\rho}(x) \hat{A}_{\rho\alpha}(x) \hat{T}_{\alpha\beta}(x - \varepsilon) h_\beta(x - \varepsilon). \quad (3.42)$$

The expectation value of $A(x)$ obtains by an insertion of the operator $\hat{A}(x)$,

$$\begin{aligned} \langle A(x) \rangle &= \frac{1}{Z} \int Ds e^{-S} A(x) \\ &= \frac{1}{Z} \bar{q}_\tau(x_f) [\hat{T}(x_f - \varepsilon) \cdots \hat{T}(x) \hat{A}(x) \hat{T}(x - \varepsilon) \cdots \hat{T}(x_{\text{in}})]_{\tau\rho} \tilde{q}_\rho(x_{\text{in}}) \end{aligned} \quad (3.43)$$

The operators $\hat{T}(x)$ and $\hat{A}(x)$ do in general not commute,

$$[\hat{T}(x), \hat{A}(x)] \neq 0. \quad (3.44)$$

Non-commutativity is present in classical statistics if one asks questions related to hypersurfaces! Let us concentrate on observables that are represented by operators \hat{A} which are independent of x . As an example we take the local occupation number $n(x) = 2s(x) - 1$. The associated operator is

$$\hat{N} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (3.45)$$

If we want to obtain the expectation value at x , we need to compute

$$\langle n(x) \rangle = \frac{1}{Z} \langle \bar{q}_f | \hat{T}(x_f - \varepsilon) \cdots \hat{T}(x) \hat{N} \hat{T}(x - \varepsilon) \cdots \hat{T}(x_{\text{in}}) | \tilde{q}_{\text{in}} \rangle, \quad (3.46)$$

where we employ a notation familiar from quantum mechanics,

$$\langle \bar{q}_f | \hat{M} | \tilde{q}_{\text{in}} \rangle = (\bar{q}_f(x_f))_\tau \hat{M}_{\tau\rho}(q_{\text{in}}(x_{\text{in}}))_\rho. \quad (3.47)$$

We may now consider the operator

$$\hat{N}_+ = \hat{T}(x)^{-1} \hat{N} \hat{T}(x), \quad (3.48)$$

and compute

$$\langle \tilde{q}_f | \hat{T}(x_f - \varepsilon) \cdots \hat{T}(x) \hat{N}_+ \hat{T}(x - \varepsilon) \cdots \hat{T}(x_{\text{in}}) | \tilde{q}_{\text{in}} \rangle = \langle n(x + \varepsilon) \rangle. \quad (3.49)$$

When we use the same prescription (with x singled out as a reference point) the operator \hat{N} corresponds to the observable $n(x)$, while \hat{N}_+ is associated to the observable $n(x + \varepsilon)$. The operator \hat{N}_+ is not diagonal and does not commute with \hat{N} ,

$$[\hat{N}_+, \hat{N}] \neq 0. \quad (3.50)$$

We conclude that non-commuting operators do not only appear in quantum mechanics. The appearance of non-commuting structures is an issue of what questions are asked and which formalism is appropriate for the answer to these questions. One can actually devise a Heisenberg picture for classical statistical systems in close analogy to quantum mechanics. The Heisenberg operators depend on x and do not commute for different x .

3.3 Classical Wave functions

Local Probabilities. The probability distribution is given by

$$p[s] = \frac{1}{Z} e^{-S[s]}, \quad Z = \int Ds e^{-S[s]}. \quad (3.51)$$

A *local* probability distribution at x , which involves only the spin $s(x)$, can be obtained by summing over all spins at $x' \neq x$,

$$p_l(s(x)) = \frac{1}{Z} \left[\prod_{x' \neq x} \sum_{s(x') = \pm 1} \right] e^{-S} \equiv p_l(x). \quad (3.52)$$

It is properly normalized,

$$\sum_{s(x) = \pm 1} p_l(s(x)) = 1. \quad (3.53)$$

The expectation value of the spin $s(x)$ can be computed from $p_l(s(x))$,

$$\langle s(x) \rangle = \sum_{s(x) = \pm 1} p_l(s(x)) s(x). \quad (3.54)$$

If there would be a simple evolution law how to compute $p_l(x + \varepsilon)$ from $p_l(x)$, the problem could be solved iteratively. Unfortunately, such a simple evolution law does not exist for the local probabilities. We will see next that it exists for local wave functions or probability amplitudes.

Wave Functions. Define for a given x the actions S_- and S_+ by

$$\begin{aligned} S_- &= \mathcal{L}_{\text{in}} + \sum_{x'=x_{\text{in}}}^{x-\varepsilon} \mathcal{L}(x'), \\ S_+ &= \sum_{x'=x}^{x_f-\varepsilon} \mathcal{L}(x') + \mathcal{L}_f, \\ S &= S_- + S_+. \end{aligned} \quad (3.55)$$

Here S_- depends only on the Ising spins $s(x')$ with $x' \leq x$, and S_+ depends on spins $s(x')$ with $x' \geq x$.

The wave function $f(x)$ is defined by

$$f(x) = \left[\prod_{x'=x_{\text{in}}}^{x-\varepsilon} \sum_{s(x') = \pm 1} \right] e^{-S_-}. \quad (3.56)$$

Because we sum over all $s(x')$ with $x' < x$, and S_- depends only on those $s(x')$ and on $s(x)$, the wave function $f(x)$ depends only on the single spin $s(x)$. Similarly, we define the conjugate wave function

$$\bar{f}(x) = \left[\prod_{x'=x+\varepsilon}^{x_f} \sum_{s(x')=\pm 1} \right] e^{-S_+}, \quad (3.57)$$

which also depends only on $s(x)$.

Wave functions and local probability distribution. The product

$$\bar{f}(x)f(x) = \left[\prod_{x' \neq x} \sum_{s(x')=\pm 1} \right] e^{-S} = Z p_l(x), \quad (3.58)$$

is closely related to the local probability distribution $p_l(x)$. One has

$$\sum_{s(x)=\pm 1} \bar{f}(x)f(x) = Z. \quad (3.59)$$

At this point we could employ the possibility of an additive renormalisation $S \rightarrow S + C$ in order to normalise the partition function to $Z = 1$. The wave functions \bar{f} and f are then a type of probability amplitudes, similar as in quantum mechanics. We have, however, two distinct types of probability amplitudes, f and \bar{f} .

Quantum rule for expectations values of local observables. The expectation value of $A(x)$ can be written in terms of a bilinear in the wave functions.

$$\begin{aligned} \langle A(x) \rangle &= \sum_{s(x)=\pm 1} A(x) p_l(x) \\ &= \frac{1}{Z} \sum_{s(x)=\pm 1} \bar{f}(x) A(x) f(x). \end{aligned} \quad (3.60)$$

We expand again in the occupation number bases

$$\begin{aligned} f(x) &= \tilde{q}_\rho(x) h_\rho(x), \\ \bar{f}(x) &= \bar{q}_\tau(x) h_\tau(x), \\ A(x) &= A_\sigma(x) h_\sigma(x). \end{aligned} \quad (3.61)$$

Here $\tilde{q}_\rho(x)$ are the components of the wave function in the occupation number basis at x , and $\bar{q}_\tau(x)$ are the components of the conjugate wave function. This yields for the expectation values

$$\langle A(x) \rangle = \frac{1}{Z} \bar{q}_\tau(x) A_\sigma(x) \tilde{q}_\rho(x) \sum_{s(x)=\pm 1} h_\tau(x) h_\sigma(x) h_\rho(x). \quad (3.62)$$

Using again the product properties of the bases functions one finds the ‘‘quantum rule’’ for the expectation value as a bilinear in the wave functions,

$$\begin{aligned} \langle A(x) \rangle &= \frac{1}{Z} \langle \bar{q}(x) | \hat{A}(x) | \tilde{q}(x) \rangle \\ &= \frac{1}{Z} \sum_{\sigma} \bar{q}_\tau(x) A_\sigma(x) \delta_{\tau\sigma} \delta_{\sigma\rho} \tilde{q}_\rho(x). \end{aligned} \quad (3.63)$$

Knowledge of the wave function at x is therefore sufficient for the computation of $\langle A(x) \rangle$.

Evolution equation for the wave function. In contrast to the local probability distribution, the x -dependence of the wave functions is a simple linear evolution law. This makes the wave function the appropriate object for the discussion of boundary value problems and beyond. From the definition of the wave function $f(x)$ one infers immediately

$$f(x + \varepsilon) = \sum_{s(x)=\pm 1} \mathcal{K}(x) f(x). \quad (3.64)$$

As it should be, $f(x + \varepsilon)$ depends on the spin $s(x + \varepsilon)$. The expansion in the occupation number basis yields

$$\begin{aligned} f(x + \varepsilon) &= \tilde{q}_\tau(x + \varepsilon) h_\tau(x + \varepsilon) \\ &= \sum_{s(x)=\pm 1} h_\tau(x + \varepsilon) \hat{T}_{\tau\rho}(x) h_\rho(x) \tilde{q}_\sigma(x) h_\sigma(x) \\ &= \hat{T}_{\tau\rho}(x) \tilde{q}_\rho(x) h_\tau(x + \varepsilon). \end{aligned} \quad (3.65)$$

The linear evolution operator for the wave function is the transfer matrix.

$$\tilde{q}_\tau(x + \varepsilon) = \hat{T}_{\tau\rho}(x) \tilde{q}_\rho(x), \quad (3.66)$$

or, in a vector matrix notation

$$\tilde{q}(x + \varepsilon) = \hat{T}(x) \tilde{q}(x). \quad (3.67)$$

By the same type of argument one obtains for the conjugate wave function (as a row vector)

$$\bar{q}(x) = \bar{q}(x + \varepsilon) \hat{T}(x), \quad (3.68)$$

or, written as a column vector,

$$\bar{q}(x) = \hat{T}^T(x) \bar{q}(x + \varepsilon), \quad (3.69)$$

and

$$\bar{q}(x + \varepsilon) = (\hat{T}^T(x))^{-1} \bar{q}(x). \quad (3.70)$$

In cases where \hat{T} is orthogonal, $\hat{T}^{-1} = \hat{T}^T$, both \bar{q} and \tilde{q} obey the same evolution law. The evolution law is linear. The superposition law familiar from quantum mechanics follows. If $\tilde{q}_1(x)$ and $\tilde{q}_2(x)$ are two solutions of the evolution equation, this also holds for linear combinations $\alpha \tilde{q}_1(x) + \beta \tilde{q}_2(x)$.

Continuous evolution. For a sufficiently smooth wave function $\tilde{q}(x)$ one defines the derivative

$$\begin{aligned} \frac{\partial \tilde{q}}{\partial x} &= \frac{1}{2\varepsilon} (\tilde{q}(x + \varepsilon) - \tilde{q}(x - \varepsilon)) \\ &= \frac{1}{2\varepsilon} (\hat{T}(x) - \hat{T}^{-1}(x - \varepsilon)) \tilde{q}(x). \end{aligned} \quad (3.71)$$

This yields the generalised Schrödinger equation

$$\begin{aligned} \partial_x \tilde{q} &= \frac{\partial}{\partial x} \tilde{q} = W \tilde{q}, \\ W(x) &= \frac{1}{2\varepsilon} [\hat{T}(x) - \hat{T}^{-1}(x - \varepsilon)]. \end{aligned} \quad (3.72)$$

For the same \mathcal{L} at every x , both \hat{T} and W are independent of x ,

$$W = \frac{1}{2\varepsilon} [\hat{T} - \hat{T}^{-1}]. \quad (3.73)$$

Step evolution operator. The additive renormalization of the action results in a multiplicative renormalization of the transfer matrix. The step evolution operator is the transfer matrix normalized such that the absolute value of the largest eigenvalue equals unity. For the Ising model, the step evolution operator is given by

$$\hat{T} = \frac{1}{2 \cosh(\beta)} \begin{pmatrix} e^\beta & e^{-\beta} \\ e^{-\beta} & e^\beta \end{pmatrix}. \quad (3.74)$$

Equilibrium state. If only one eigenvalue of the step evolution operator equals unity in absolute magnitude, the eigenstate to this eigenvalue is the unique equilibrium state \tilde{q}_* . For the Ising model the equilibrium wave function is

$$\tilde{q}_* \sim \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (3.75)$$

The equilibrium state is invariant under the evolution.

Boundary value problem. For given boundary conditions $\tilde{q}(x_{\text{in}})$ and $\bar{q}(x_{\text{f}})$ are fixed. One can use the evolution equation to compute both $\tilde{q}(x)$ and $\bar{q}(x)$. The value of a local observable $A(x)$, with associated operator $\hat{A}(x)$, follows from

$$\langle A(x) \rangle = \frac{1}{Z} \langle \bar{q}(x) | \hat{A}(x) | \tilde{q}(x) \rangle. \quad (3.76)$$

Choose for $\tilde{q}(x_{\text{in}})$ a decomposition into eigenfunctions of the transfer matrix \hat{T} , e. g. with eigenvalues λ_+ and λ_- ,

$$\tilde{q}(x_{\text{in}}) = c_+(x_{\text{in}})\tilde{q}_+ + c_-(x_{\text{in}})\tilde{q}_-, \quad (3.77)$$

such that

$$\tilde{q}(x) = \tilde{q}(x_{\text{in}} + N\varepsilon) = c_+(x_{\text{in}})(\lambda_+)^N \tilde{q}_+ + c_-(x_{\text{in}})(\lambda_-)^N \tilde{q}_-. \quad (3.78)$$

For $\lambda_+ = 1$, the corresponding eigenfunction is the equilibrium wave function,

$$\lambda_+^N \tilde{q}_+ = \tilde{q}_+. \quad (3.79)$$

For $\lambda_- \leq 1$ the contribution $\sim (\lambda_-)^N \tilde{q}_-$ vanishes for large N . This describes the approach to equilibrium. The correlation length is directly related to λ_- .

4 Quantum Fields and Functional Integral

In this lecture we will start from quantum mechanics and construct the functional integral. In the last lecture we did *functional integral* \rightarrow *operators*. In this lecture we will do *operators* \rightarrow *functional integral*. The aim of the lecture is once more to show the equivalence of the functional integral and the operator formalism. We will do this already for quantum fields, establishing in this way also the basic notions of quantum field theory in the operator formalism.

4.1 Phonons as quantum fields in one dimension

One-dimensional crystal. Consider a one-dimensional crystal of atoms with lattice sites $x_j = j\varepsilon$ and lattice distance ε . Denote the displacement from the equilibrium position at x_j by Q_j and the momentum of the atoms by P_j . The Hamiltonian for small displacements can be taken quadratic in Q_j , and we decompose $H = H_0 + H_{\text{nn}}$ with

$$H_0 = \sum_j \left(\frac{P_j^2}{2M} + \frac{D}{2} Q_j^2 \right), \quad H_{\text{nn}} = -\frac{B}{2} \sum_j Q_{j+1} Q_j. \quad (4.1)$$

Here Q_j and P_j are quantum operators with the usual commutation relations

$$[Q_j, P_k] = i\delta_{jk}, \quad [Q_j, Q_k] = 0, \quad [P_i, P_j] = 0. \quad (4.2)$$

We use units where $\hbar = 1$. The displacements are a *quantum field*,

$$Q_j = Q(x). \quad (4.3)$$

This is an *operator field*. For each x one has an operator $Q(x)$. Similarly $P(x) = P_j$ is a quantum field. One may consider the pairs $\{Q_j, P_j\}$ as a common (two-component) quantum field.

Occupation number basis. At each site j we define annihilation and creation operators a_j and a_j^\dagger . The annihilation operators are

$$a_j = \frac{1}{\sqrt{2}} \left((DM)^{\frac{1}{4}} Q_j + i(DM)^{-\frac{1}{4}} P_j \right). \quad (4.4)$$

The creation operators are

$$a_j^\dagger = \frac{1}{\sqrt{2}} \left((DM)^{\frac{1}{4}} Q_j - i(DM)^{-\frac{1}{4}} P_j \right). \quad (4.5)$$

Note that they are formally hermitian conjugates, $a_j^\dagger = (a_j)^\dagger$. The commutation relations are

$$[a_j, a_k^\dagger] = \delta_{jk}, \quad [a_j, a_k] = 0, \quad [a_j^\dagger, a_k^\dagger] = 0. \quad (4.6)$$

Both $a(x) = a_j$ and $a^\dagger(x) = a_j^\dagger$ are operator fields. Inserting

$$Q(x) = Q_j = \frac{1}{\sqrt{2}} (DM)^{-\frac{1}{4}} (a_j + a_j^\dagger), \quad (4.7)$$

and similar for P_j , we express the Hamiltonian in terms of a and a^\dagger ,

$$H_0 = \omega_0 \sum_j \left(a_j^\dagger a_j + \frac{1}{2} \right) = \omega_0 \sum_j \left(\hat{n}_j + \frac{1}{2} \right), \quad (4.8)$$

with the frequency $w_0 = \sqrt{D/M}$. Occupation numbers at positions x_j are expressed in terms of the operator $\hat{n}_j = a_j^\dagger a_j$. It has the eigenvalues $n_j = (0, 1, 2, \dots)$. At each site j there are a number n_j of ‘‘phonons’’. For $B = 0$ the system describes uncoupled harmonic oscillators, one at each lattice site. We next need the next-neighbour interaction which involves products of a_j, a_{j+1} etc.,

$$\begin{aligned} H_{\text{nn}} &= -\frac{B}{2} \sum_j Q_{j+1} Q_j \\ &= -\frac{B}{2} \frac{(DM)^{-\frac{1}{2}}}{2} \sum_j (a_{j+1} + a_{j+1}^\dagger) (a_j + a_j^\dagger). \end{aligned} \quad (4.9)$$

Momentum Space. It is possible to diagonalize H by Fourier transform. To this end, we write

$$a_j = \frac{1}{\sqrt{\mathcal{N}}} \sum_q e^{i\varepsilon q j} a_q, \quad a_j^\dagger = \frac{1}{\sqrt{\mathcal{N}}} \sum_q e^{-i\varepsilon q j} a_q^\dagger. \quad (4.10)$$

Here the sum is periodic in q ,

$$\sum_q = \sum_{|q| \leq \frac{\pi}{\varepsilon}}, \quad (4.11)$$

and $\mathcal{N} = \sum_j$ is a normalization factor corresponding to the number of lattice sites. This yields

$$\begin{aligned} Q_j &= \frac{1}{\sqrt{2\mathcal{N}}} (DM)^{-\frac{1}{4}} \sum_q (e^{i\varepsilon q j} a_q + e^{-i\varepsilon q j} a_q^\dagger) \\ &= \frac{1}{\sqrt{2\mathcal{N}}} (DM)^{-\frac{1}{4}} \sum_q e^{i\varepsilon q j} (a_q + a_{-q}^\dagger), \end{aligned} \quad (4.12)$$

and therefore

$$H_{\text{nn}} = -\frac{B}{4\mathcal{N}} (DM)^{-\frac{1}{2}} \sum_j \sum_q \sum_{q'} e^{i\varepsilon q' j} e^{i\varepsilon q(j+1)} (a_q + a_{-q}^\dagger) (a_{q'}^\dagger + a_{-q'}). \quad (4.13)$$

Use now the following identity for discrete Fourier transforms,

$$\sum_j e^{i\varepsilon(q+q')j} = \mathcal{N} \delta_{q,-q'}, \quad (4.14)$$

which corresponds to the familiar continuum expression

$$\int dx e^{i(q+q')x} = 2\pi \delta(q+q'). \quad (4.15)$$

One obtains

$$\begin{aligned} H_{\text{nn}} &= -b \sum_q e^{i\varepsilon q} (a_q + a_{-q}^\dagger) (a_{-q} + a_q^\dagger) \\ &= -b \sum_q \cos(\varepsilon q) (a_q + a_q^\dagger) (a_{-q} + a_{-q}^\dagger), \end{aligned} \quad (4.16)$$

with $b = \frac{B}{4} (DM)^{-\frac{1}{2}}$. Similarly, one has

$$H_0 = \omega_0 \sum_q \left(a_q^\dagger a_q + \frac{1}{2} \right). \quad (4.17)$$

At this stage, the Hamiltonian H involves separate q -blocks,

$$H = \sum_q H_q, \quad (4.18)$$

with

$$H_q = \omega_q \left(a_q^\dagger a_q + \frac{1}{2} \right) - b \cos(\varepsilon q) (a_q + a_{-q}^\dagger) (a_{-q} + a_q^\dagger). \quad (4.19)$$

Each block involves q and $-q$. What remains is the diagonalization of the q -blocks, done by the *Bogoliubov transformation*,

$$a_q = \alpha(q) A_q + \beta(q) A_{-q}^\dagger, \quad a_q^\dagger = \alpha(q) A_q^\dagger + \beta(q) A_{-q}, \quad (4.20)$$

where the commutation relations

$$[a_q, a_q^\dagger] = 1, \quad [A_q, A_q^\dagger] = 1, \quad (4.21)$$

require

$$\alpha(q)^2 - \beta(q)^2 = 1. \quad (4.22)$$

One finds after some simple algebra

$$H = \sum_q \omega_q \left(A_q^\dagger A_q + \frac{1}{2} \right), \quad (4.23)$$

with

$$\omega_q^2 = \frac{D}{M} \left(1 - \frac{B}{D} \cos(\varepsilon q) \right). \quad (4.24)$$

Phonons can be described as uncoupled harmonic oscillators, one for every momentum q . They are a *free quantum field*, which means that they do not interact with themselves.

The vacuum obeys, as usual $A_q|0\rangle = 0$. This is not the same as for $B = 0$, where one has $a_q|0\rangle = 0$. The vacuum can be a complicated object. The excitations, *quasiparticles* or simply *particles* depend on the vacuum, e. g. the dispersion relation depends on B .

Dispersion relation. The equation

$$\omega(q) = \omega_q = \sqrt{\frac{D - B \cos(\varepsilon q)}{M}}, \quad (4.25)$$

is called the dispersion relation. Consider the limit of small εq , where one can expand, $\cos(\varepsilon q) = 1 - \frac{1}{2}\varepsilon^2 q^2$, such that

$$\omega^2(q) = \frac{D - B}{M} + \frac{\varepsilon^2 B}{2M} q^2. \quad (4.26)$$

The dispersion relation corresponds to the energy momentum relation of the *phonon-quasi-particles*. The sound velocity is given here by

$$v = \left| \frac{d\omega}{dq} \right| = \frac{\varepsilon^2 B q}{2M\omega(q)}. \quad (4.27)$$

For $D > B$ the occupation relation has a gap, one needs positive energy even for a phonon with $q = 0$. For many cases the interaction between atoms is of the form $(Q_j - Q_{j-1})^2$, involving only the distance between two neighbouring atoms. Then $D = B$, phonons are gapless and the dispersion relation becomes linear for small εq .

Generalisations. In three dimensions $d = 3$ one has $q \rightarrow \vec{q}$ and the dispersion relation becomes an equation for $\omega(\vec{q})$.

Continuum limit. This corresponds to the limit $\varepsilon \rightarrow 0$.

Photons. For photons the dispersion relation is (in units where the velocity of light is unity, $c = 1$),

$$\omega(\vec{q}) = |\vec{q}|. \quad (4.28)$$

There are two photon helicities.

Quantum fields for photons. For photons, the quantum fields would have to be the electric field $\vec{E}(\vec{q})$ in momentum space or $\vec{E}(\vec{x})$ in position space and the magnetic field $\vec{B}(\vec{q})$ or $\vec{B}(\vec{x})$, respectively. In other words, the electric field \vec{E} and the magnetic field \vec{B} are quantum operators! One at each \vec{x} or for each \vec{q} .

Bosonic atoms without interaction. For free, non-relativistic atoms, the dispersion relation is given by

$$\omega(\vec{q}) = \frac{\vec{q}^2}{2M}. \quad (4.29)$$

For the grand-canonical ensemble, one includes a chemical potential, multiplying the total particle number. This shifts effectively

$$\omega(\vec{q}) \rightarrow \varepsilon(\vec{q}) = \frac{\vec{q}^2}{2M} - \mu. \quad (4.30)$$

We will not distinguish $\omega(\vec{q})$ and $\varepsilon(\vec{q})$ unless stated otherwise.

4.2 Functional integral for quantum fields

Free quantum boson gas in thermal equilibrium. For the Hamiltonian

$$H = \sum_q \omega(q) \left(a_q^\dagger a_q + \frac{1}{2} \right), \quad (4.31)$$

the partition function in thermal equilibrium is given by the trace

$$Z = \text{Tr} e^{-\beta H}, \quad (4.32)$$

with $\beta = \frac{1}{k_B T} = \frac{1}{T}$ (we use units where $k_B = 1$). It decays into independent factors

$$Z = \prod_q \text{Tr} e^{-\beta \omega_q \left(a_q^\dagger a_q + \frac{1}{2} \right)} = \prod_q Z_q. \quad (4.33)$$

We can compute the individual Z_q ,

$$Z = \text{Tr} e^{-\tilde{\beta} \left(a^\dagger a + \frac{1}{2} \right)}, \quad (4.34)$$

with $\tilde{\beta} = \beta \omega_q$ (we omit the index q). As an example, for a free gas of bosonic atoms one has

$$\omega(q) = \frac{\vec{q}^2}{2M} - \mu, \quad (4.35)$$

with chemical potential μ . From $Z(\beta, \mu)$ one can derive all thermodynamics of the quantum boson gas. This will be done in lecture 6 including interactions. In this lecture we will derive a functional integral representation of the partition function

$$Z = \text{Tr} e^{-\beta H} = \int D\phi e^{-S[\phi]}, \quad (4.36)$$

with Euclidean action

$$S = \int_{-\frac{\beta}{2}}^{\frac{\beta}{2}} d\tau \sum_q \phi^*(\tau, q) \left(\frac{\partial}{\partial \tau} + \omega(q) \right) \phi(\tau, q), \quad (4.37)$$

and complex fields $\phi(\tau, q)$.

Partition function with boundary conditions. Let us consider the expression

$$\tilde{Z} = \text{Tr} \left\{ b e^{-\tilde{\beta} \left(a^\dagger a + \frac{1}{2} \right)} \right\}. \quad (4.38)$$

For $b = 1$ one has $\tilde{Z} = Z$ for thermodynamic equilibrium if $\tilde{\beta} = \beta \omega$ is real. More, generally, b is a matrix in Hilbert space reflecting boundary conditions. For example, in the occupation number basis one has

$$\tilde{Z} = b_{nm} \left(e^{-\tilde{\beta} \left(a^\dagger a + \frac{1}{2} \right)} \right)_{mn}. \quad (4.39)$$

We may take the “boundary term” b as a product of wave functions,

$$b_{nm} = (\psi_{\text{in}})_n (\phi_{\text{f}})_m, \quad (4.40)$$

such that

$$\begin{aligned} \tilde{Z} &= (\phi_{\text{f}})_m \left(e^{-\tilde{\beta} \left(a^\dagger a + \frac{1}{2} \right)} \right)_{mn} (\psi_{\text{in}})_n \\ &= \left\langle \phi_{\text{f}} \left| e^{-\tilde{\beta} \left(a^\dagger a + \frac{1}{2} \right)} \right| \psi_{\text{in}} \right\rangle. \end{aligned} \quad (4.41)$$

Extension to complex formulation. Take imaginary $\tilde{\beta}$,

$$\tilde{\beta} = i\omega\Delta t, \quad (4.42)$$

and admit complex ϕ_f and ψ_{in} , defining $\langle\phi_f|$ by involving complex conjugation as in quantum mechanics, e. g. $\langle\phi_f|m = (\phi_f^*)_m$. In general, \tilde{Z} will now be a complex number. Everything remains well defined.

Transition amplitude. With this setting \tilde{Z} is the transition amplitude

$$\begin{aligned} \tilde{Z} &= \langle\phi_f|e^{-i\Delta t\omega\left(a^\dagger a + \frac{1}{2}\right)}|\psi_{\text{in}}\rangle \\ &= \langle\phi_f|e^{-i\Delta tH}|\psi_{\text{in}}\rangle. \end{aligned} \quad (4.43)$$

Here $H = \omega\left(a^\dagger a + \frac{1}{2}\right)$ stands for H_q . If we take

$$\psi_{\text{in}} = \psi(t_{\text{in}}), \quad \phi_f = \phi(t_f), \quad (4.44)$$

the quantity \tilde{Z} denotes the transition amplitude between ψ and ϕ at the common time t_f ,

$$\tilde{Z} = \langle\phi(t_f)|\psi(t_f)\rangle, \quad \Delta t = t_f - t_{\text{in}}, \quad (4.45)$$

where

$$\psi(t_f) = e^{-i(t_f - t_{\text{in}})H}\psi(t_{\text{in}}). \quad (4.46)$$

The square $|\tilde{Z}|^2$ measures the probability that a given $\psi(t_{\text{in}})$ coincides at t_f with $\phi(t_f)$. The transition amplitude is a key element for the S-matrix for scattering to be discussed in coming lectures.

Split into factors. The idea is now to split $\tilde{\beta}$ into small steps by writing $\tilde{\beta} = (2N + 1)\delta$, where $|\delta| \ll 1$ and assuming N to be even. One has then

$$\exp\left\{-\tilde{\beta}\left[a^\dagger a + \frac{1}{2}\right]\right\} = \prod_{j=-N}^N \exp\left\{-\delta\left[a^\dagger a + \frac{1}{2}\right]\right\}. \quad (4.47)$$

For small δ , the exponential simplifies. This would not be necessary for the present very simple case, but is very useful for more complicated Hamiltonians which involves pieces that do not commute with each other. The split will be used to define a functional integral. Indeed, the expression (4.47) looks already like a product of transfer matrices. At the end $N \rightarrow \infty$ is possible. Define now the operators

$$\hat{x} = \frac{1}{\sqrt{2}}(a^\dagger + a), \quad \hat{p} = \frac{i}{\sqrt{2}}(a^\dagger - a), \quad (4.48)$$

with commutation relation

$$[\hat{x}, \hat{p}] = i. \quad (4.49)$$

Note that the operators \hat{x} and \hat{p} have similar properties as position and momentum operators. In our context they are abstract operators, since for phonons or photons already $a^\dagger a$ stands for $a_q^\dagger a_q$ in momentum space. Thus \hat{x} and \hat{p} have nothing to do with position and momentum of phonons or photons. One has

$$\hat{H} = a^\dagger a + \frac{1}{2} = \frac{\hat{p}^2}{2} + V(\hat{x}), \quad V(\hat{x}) = \frac{\hat{x}^2}{2}. \quad (4.50)$$

This yields the expression

$$\exp\left\{-\tilde{\beta}\left[a^\dagger a + \frac{1}{2}\right]\right\} = \prod_{j=-N}^N \exp\left\{-\delta\left[\frac{\hat{p}^2}{2} + V(\hat{x})\right]\right\}, \quad (4.51)$$

where

$$\tilde{H} = \frac{\hat{p}^2}{2} + V(\hat{x}). \quad (4.52)$$

Eigenfunctions of \hat{x} and \hat{p} . We next define eigenfunctions of the operators \hat{x} and \hat{p} ,

$$|x\rangle \quad \text{such that} \quad \hat{x}|x\rangle = x|x\rangle, \quad (4.53)$$

and

$$|p\rangle \quad \text{such that} \quad \hat{p}|p\rangle = p|p\rangle. \quad (4.54)$$

We can choose a normalization such that

$$\langle x'|x\rangle = \delta(x' - x), \quad \langle p'|p\rangle = 2\pi\delta(p' - p), \quad (4.55)$$

and

$$\int dx |x\rangle\langle x| = \mathbb{1}, \quad \int \frac{dp}{2\pi} |p\rangle\langle p| = \mathbb{1}. \quad (4.56)$$

We next insert complete systems of functions between each of the factors,

$$\prod_{j=-N}^N e^{-\delta\tilde{H}} = \left[\prod_{j=-N}^{N+1} dx_j \right] |x_{N+1}\rangle\langle x_{N+1}| e^{-\delta\tilde{H}} |x_N\rangle\langle x_N| \cdots |x_{1-N}\rangle\langle x_{1-N}| e^{-\delta\tilde{H}} |x_{-N}\rangle\langle x_{-N}|. \quad (4.57)$$

Evaluation of factors. The factors $\langle x_{j+1}|e^{-\delta\tilde{H}}|x_j\rangle$ are complex numbers, no longer operators. For their computation it is convenient to insert a complete set of \hat{p} -eigenstates,

$$\langle x_{j+1}|e^{-\delta\tilde{H}}|x_j\rangle = \int \frac{dp_j}{2\pi} \langle x_{j+1}|p_j\rangle\langle p_j|e^{-\delta\tilde{H}}|x_j\rangle. \quad (4.58)$$

We next use for $\delta \rightarrow 0$ the expansion

$$\exp\left\{-\delta\left[\frac{\hat{p}^2}{2} + V(\hat{x})\right]\right\} = \exp\left\{-\delta\frac{\hat{p}^2}{2}\right\} \exp\{-\delta V(\hat{x})\} + \mathcal{O}(\delta^2), \quad (4.59)$$

where the term $\sim \mathcal{O}(\delta^2)$ arises from the commutator of \hat{x} and \hat{p} . Corrections $\sim \delta^2$ can be neglected for $\delta \rightarrow 0$ such that

$$\langle x_{j+1}|e^{-\delta\tilde{H}}|x_j\rangle = \int \frac{dp_j}{2\pi} e^{-\delta\frac{p_j^2}{2}} e^{-\delta V(x_j)} \langle x_{j+1}|p_j\rangle\langle p_j|x_j\rangle. \quad (4.60)$$

No operators appear anymore in this expression and we only need

$$\langle p_j|x_j\rangle = e^{-ip_j x_j}, \quad \langle x_{j+1}|p_j\rangle\langle p_j|x_j\rangle = e^{ip_j(x_{j+1}-x_j)}. \quad (4.61)$$

This yields the expression

$$\langle x_{j+1}|e^{-\delta\tilde{H}}|x_j\rangle = \int \frac{dp_j}{2\pi} \exp\left\{ip_j(x_{j+1}-x_j) - \delta\left[\frac{p_j^2}{2} + V(x_j)\right]\right\}. \quad (4.62)$$

Functional integral. Insertion of these factors yields

$$e^{-\tilde{\beta}\tilde{H}} = \int dx_{-N} \int dx_{N+1} |x_{N+1}\rangle F |x_{-N}\rangle, \quad (4.63)$$

with

$$F = \int D\phi' \exp\left\{\sum_{j=-N}^N \left[ip_j(x_{j+1}-x_j) - \delta\frac{p_j^2}{2} + \delta V(x_j)\right]\right\}, \quad (4.64)$$

and functional measure

$$\int D\phi' = \left[\prod_{j=-N+1}^N \int_{-\infty}^{\infty} dx_j \right] \left[\prod_{j=-N}^N \int_{-\infty}^{\infty} \frac{dp_j}{2\pi} \right]. \quad (4.65)$$

With boundary terms one obtains

$$\langle \phi_f | e^{-\tilde{\beta}\tilde{H}} | \psi_{\text{in}} \rangle = \int dx_{-N} \int dx_{N+1} \langle \phi_f | x_{N+1} \rangle F |x_{-N}\rangle \langle \psi_{\text{in}}|. \quad (4.66)$$

4.3 Thermodynamic equilibrium

For thermodynamic equilibrium, $Z = \text{Tr} e^{-\tilde{\beta}\tilde{H}}$, one identifies x_{N+1} with x_{-N} and includes no integration over x_{N+1} . The variable j is periodic, reflecting in

$$x_{N+1} = x_{-N}, \quad p_{N+1} = p_{-N}. \quad (4.67)$$

One has for any given q -mode

$$\tilde{Z} = \text{Tr} e^{\tilde{\beta}\tilde{H}} = \int D\phi e^{-S}, \quad (4.68)$$

with

$$S = - \sum_{j=N}^N \left\{ ip_j(x_{j+1} - x_j) - \delta \left[\frac{p_j^2}{2} + V(x_j) \right] \right\}, \quad (4.69)$$

and

$$\int D\phi = \left[\prod_j \int dx_j \int \frac{dp_j}{2\pi} \right]. \quad (4.70)$$

For $2N + 1$ factors one has (periodic boundary conditions) $\delta = \frac{\tilde{\beta}}{2N+1}$.

Matsubara modes. We can diagonalize the action S by a type of Fourier transform

$$x_j = \sum_{n=-N}^N \exp\left(\frac{2\pi inj}{2N+1}\right) \tilde{x}_n, \quad \tilde{x}_{-n} = \tilde{x}_n^*, \quad (4.71)$$

$$p_j = \sum_{n=-N}^N \exp\left(\frac{2\pi in(j + \frac{1}{2})}{2N+1}\right) \tilde{p}_n, \quad \tilde{p}_{-n} = \tilde{p}_n^*, \quad (4.72)$$

such that

$$- \sum_{j=-N}^N [ip_j(x_{j+1} - x_j)] = \sum_{n=-N}^N [(2N+1)\sin\left(\frac{\pi n}{2N+1}\right) (\tilde{p}_n^* \tilde{x}_n - \tilde{p}_n \tilde{x}_n^*)]. \quad (4.73)$$

Here we use the identity ($j = -N$ and $j = N + 1$ identified)

$$\sum_{j=-N}^N \exp\left(\frac{2\pi i(m-n)j}{2N+1}\right) = (2N+1)\delta_{m,n}. \quad (4.74)$$

Similarly, with $V(x_j) = x_j^2/2$, one has

$$\frac{\delta}{2} \sum_{j=-N}^N (x_j^2 + p_j^2) = \frac{(2N+1)\delta}{2} \sum_{n=-N}^N (\tilde{x}_n^* \tilde{x}_n + \tilde{p}_n^* \tilde{p}_n) = \frac{\tilde{\beta}}{2} \sum_{n=-N}^N (\tilde{x}_n^* \tilde{x}_n + \tilde{p}_n^* \tilde{p}_n). \quad (4.75)$$

We next introduce complex number ϕ_n by

$$\tilde{x}_n = \frac{1}{\sqrt{2}}(\phi_n + \phi_{-n}^*), \quad \tilde{p}_n = -\frac{i}{\sqrt{2}}(\phi_n - \phi_{-n}^*), \quad (4.76)$$

With

$$\tilde{p}_n^* \tilde{x}_n - \tilde{x}_n^* \tilde{p}_n = i(\phi_n^* \phi_n - \phi_{-n}^* \phi_{-n}), \quad (4.77)$$

and

$$\tilde{x}_n^* \tilde{x}_n + \tilde{p}_n^* \tilde{p}_n = \phi_n^* \phi_n + \phi_{-n}^* \phi_{-n}. \quad (4.78)$$

we finally obtain for the action

$$S = \sum_{n=-N}^N \left[2(2N+1) i \sin \left(\frac{\pi n}{(2N+1)} \right) + \tilde{\beta} \right] \phi_n^* \phi_n. \quad (4.79)$$

At the end we take the limit $N \rightarrow \infty$. In this limit the neglected terms (from commutators of \hat{x} and \hat{p}) vanish. This yields the central functional integral equation for thermodynamic equilibrium,

$$\text{Tr}\{e^{-\beta H}\} = \int D\phi e^{-S}. \quad (4.80)$$

For $H = \omega(a^\dagger a + \frac{1}{2})$ one has

$$S = \sum_{n=-\infty}^{\infty} (2\pi i n + \beta\omega) \phi_n^* \phi_n. \quad (4.81)$$

(Recall that $\tilde{H} = a^\dagger a + \frac{1}{2}$ and $\tilde{\beta} = \beta\omega$.) The modes ϕ_n are called Matsubara modes, and the sum over n is the Matsubara sum.

One can also translate the integration measure for the variables x_j and p_j to ϕ_n . With

$$\phi_n = \phi_{nR} + i\phi_{nI}, \quad (4.82)$$

one has

$$\int D\phi = \prod_n \left(\int_{-\infty}^{\infty} d\phi_{nR} \int_{-\infty}^{\infty} d\phi_{nI} \right). \quad (4.83)$$

All variable transformations have been linear transformations and there is therefore no non-trivial Jacobian. Recall that an overall constant factor of Z or additive constant in S is irrelevant.

Action for free quantum fields. Since Z factorises, $Z = \prod_q Z_q$, the action for all momentum modes is simply the sum of actions for individual momentum modes, $S = \sum_q S_q$. For a given momentum mode one has $\tilde{\beta} = \beta\omega_q$. Thus for

$$H = \sum_q \omega(q) \left[a_q^\dagger a_q + \frac{1}{2} \right], \quad (4.84)$$

one obtains

$$\begin{aligned} S &= \sum_n \sum_q [2\pi i n + \beta\omega(q)] \phi_n^*(q) \phi_n(q) \\ &= \sum_n \sum_q \beta [i\tilde{\omega}_n + \omega(q)] \phi_n^*(q) \phi_n(q). \end{aligned}$$

One often denotes the *dispersion relation* by $\omega(q)$ or by $\varepsilon(q)$. The quantities

$$\tilde{\omega}_n = \frac{2\pi n}{\beta} = 2\pi n T \quad (4.85)$$

are called *Matsubara frequencies*. At this point we have formulated the thermodynamics of phonons or photons as a functional integral. It is Gaussian and can easily be solved explicitly.

Euclidean time. We can consider the Matsubara modes ϕ_n as the modes of a discrete Fourier transformation. Indeed, making a Fourier transformations of functions on a circle yields discrete modes. Consider a function $\phi(\tau)$, with τ parameterizing a circle with circumference β . Equivalently, we can take τ to be a periodic variable with period β

$$\tau + \beta \equiv \tau. \quad (4.86)$$

The Fourier expansion reads

$$\phi(\tau) = \sum_n \exp\left(\frac{2\pi i n \tau}{\beta}\right) \phi_n, \quad (4.87)$$

with integer n . With

$$\begin{aligned} \partial_\tau \phi(\tau) &= \sum_n \left(\frac{2\pi i n}{\beta}\right) \exp\left(\frac{2\pi i n \tau}{\beta}\right) \phi_n \\ &= \sum_n i\tilde{\omega}_n \exp\left(\frac{2\pi i n \tau}{\beta}\right) \phi_n, \end{aligned}$$

one has

$$\int_{-\frac{\beta}{2}}^{\frac{\beta}{2}} d\tau \{\phi^*(\tau) \partial_\tau \phi(\tau)\} = \sum_n i\tilde{\omega}_n \phi_n^* \phi_n, \quad (4.88)$$

using

$$\int_{-\frac{\beta}{2}}^{\frac{\beta}{2}} d\tau \exp\left(\frac{2\pi i(n-m)\tau}{\beta}\right) = \beta \delta_{m,n}. \quad (4.89)$$

In this basis the action reads

$$S = \int_{-\frac{\beta}{2}}^{\frac{\beta}{2}} d\tau \sum_q [\phi^*(\tau, q) \partial_\tau \phi(\tau, q) + \omega(q) \phi^*(\tau, q) \phi(\tau, q)]. \quad (4.90)$$

One calls τ the *Euclidean time*.

Local action. This action is a *local action* in the sense of lectures 2 and 3. Discretizing τ on a lattice with distance ε , and with $\tau = j\varepsilon$, $j = -N \cdots N$ periodic, $\varepsilon = \frac{\beta}{2N+1}$,

$$\partial_\tau \phi(\tau) = \frac{1}{\varepsilon} [\phi(\tau + \varepsilon) - \phi(\tau)], \quad (4.91)$$

One can write (with $\sum_\tau \equiv \sum_j$)

$$S = \sum_\tau \mathcal{L}(\tau), \quad (4.92)$$

with

$$\mathcal{L}(\tau) = \frac{1}{2} \sum_q \{\phi(\tau + \varepsilon) \phi^*(\tau) - \phi^*(\tau + \varepsilon) \phi(\tau) + \varepsilon \omega(q) [\phi^*(\tau + \varepsilon) \phi(\tau) + \phi(\tau + \varepsilon) \phi^*(\tau)]\}. \quad (4.93)$$

Note that $\mathcal{L}(\tau)$ is a complex function of complex variables $\phi(\tau)$ and $\phi(\tau + \varepsilon)$. The action involves *next neighbour interactions*, similar to the Ising model. We could go the inverse way and compute the transfer matrix. We know already the answer in the bosonic occupation number basis

$$\hat{T} = \exp \left[-\frac{\beta}{2N+1} \sum_q \omega(q) \left(a_q^\dagger a_q + \frac{1}{2} \right) \right], \quad (4.94)$$

with $2N+1$ the number of *time points*. This is compatible with

$$Z = \text{Tr} \left\{ \hat{T}^{2N+1} \right\}. \quad (4.95)$$

Quantum gas of bosonic atoms. For free bosonic atoms (without internal degrees of freedom) the dispersion relation is

$$\omega(q) = \frac{\vec{q}^2}{2M} - \mu, \quad (4.96)$$

with μ the chemical potential. We can make a Fourier-transform to three-dimensional position space,

$$S = \int_{-\frac{\beta}{2}}^{\frac{\beta}{2}} d\tau \int d^3x \{ \phi^*(\tau, \vec{x}) \partial_\tau \phi(\tau, x) + \frac{1}{2M} \vec{\nabla} \phi^*(\tau, \vec{x}) \vec{\nabla} \phi(\tau, \vec{x}) - \mu \phi^*(x) \phi(x) \} \quad (4.97)$$

This is the action of a *classical field theory* (in Euclidean time).

Quantum field theory: the action defines the weight factor in a functional integral. Extremum of action yield classical field equation. For QFT the fluctuations matter!

Interactions So far we have discussed models that represent quantum fields without interactions. This is a very good approximation for photons if the energy is not too high. Free quantum field theories can be represented in momentum space as uncoupled harmonic oscillators. For them the description is simple both in the functional integral formalism (gaussian integration) and in the operator formalism. The situation changes in the presence of interactions.

Consider a particle interaction between bosonic atoms.

$$H = H_0 + H_{int} \quad (4.98)$$

$$H_0 = \sum_q \omega(q) \left(a_q^\dagger a_q + \frac{1}{2} \right) \quad (4.99)$$

$$H_{int} = \frac{\lambda}{2} \sum_{q_1, q_2, q_3, q_4} a_{q_4}^\dagger a_{q_3}^\dagger a_{q_2} a_{q_1} \delta(q_1 + q_2 - q_3 - q_4). \quad (4.100)$$

Two atoms with momentum q_1 and q_2 are annihilated, two atoms with momenta q_3 and q_4 are created. Momentum conservation is guaranteed by the δ -function.

For the functional integral this adds to the action a piece

$$S_{int} = \frac{\lambda}{2} \int d\tau \int d^3x [(\phi^*(\tau, \vec{x}) \phi(\tau, \vec{x}))^2 - 2\delta \phi^*(\tau, \vec{x}) \phi(\tau, \vec{x})] \quad (4.101)$$

with $\delta \sim \lambda$ a *counterterm* that corrects μ . A systematic treatment of interactions is rather hard in the operator formalism. For the functional integral formulation powerful methods are available.

Zero temperature limit For $T \rightarrow 0$ one has $\beta \rightarrow \infty$. The circumference of the circle goes to infinity. Instead of discrete Matsubara modes one has continuous modes with frequency $\tilde{\omega} = q_0$ and therefore a continuous four-dimensional momentum integral. The momenta q_0 and \vec{q} appear, however differently in the action. The same holds for the dependence of S on τ and \vec{x} . There is a first derivative with respect to τ , but a squared first derivative or second derivative with respect to \vec{x} . This difference will go away for relativistic particles. For bosonic atoms with a pointlike interaction one finds for the $T \rightarrow 0$ limit of the thermal equilibrium state

$$S = \int_q \left[\phi^*(q) \left(i\tilde{\omega} + \frac{\vec{q}^2}{2M} - \mu + \lambda\delta \right) \phi(q) + \frac{\lambda}{2} \int_{q_1} \int_{q_2} \int_{q_3} \int_{q_4} \phi^*(q_4) \phi^*(q_3) \phi(q_2) \phi(q_1) \delta(q_4 + q_3 - q_2 - q_1) \right], \quad (4.102)$$

where we have chosen an appropriate *continuum normalization* of $\phi(q)$, with

$$\phi(q) \equiv \phi(\tilde{\omega}, \vec{q}) \quad (4.103)$$

$$\int_q = \frac{1}{(2\pi)^4} \int d\tilde{\omega} d^3\vec{q} \quad (4.104)$$

$$\delta(q) = (2\pi)^4 \delta(\omega) \delta(q_1) \delta(q_2) \delta(q_3). \quad (4.105)$$

The δ function expresses conservation of the euclidean four momentum q . It reflects translation symmetry in space and euclidean time τ . The limit $T \rightarrow 0$ may be associated in some sense with the vacuum, for μ chosen such that the mean particle number vanishes.

4.4 Real time evolution

Recall the transition amplitude for the quantum mechanical time evolution

$$\langle \phi(t_f) | \psi(t_f) \rangle = \langle \phi(t_f) | U(t_f - t_{\text{in}}) | \psi(t_{\text{in}}) \rangle = \langle \phi(t_f) | e^{-i(t_f - t_{\text{in}})H} | \psi(t_{\text{in}}) \rangle. \quad (4.106)$$

Up to boundary terms this is the same expression as for thermal equilibrium, with a replacement

$$\beta \rightarrow i(t_f - t_{\text{in}}). \quad (4.107)$$

The split into infinitesimal pieces, Fourier-transforms etc can be done for complex β .

For $\beta \rightarrow \infty$ ($T \rightarrow 0$), $t_f - t_{\text{in}} \rightarrow \infty$ one finds

$$\langle \phi(t_f) | \psi(t_f) \rangle = B(t_f, t_{\text{in}}) Z_M \quad (4.108)$$

$$Z_M = \int D\phi \exp(-S). \quad (4.109)$$

In the action we have to multiply the terms $\sim \beta$ by i , before taking the limit $\beta \rightarrow \infty$. This results (for $\mu = 0$) in

$$S = \int_q \left[\phi^*(q) \left[i\tilde{\omega} + i \left(\frac{\vec{q}^2}{2M} + \lambda\delta \right) \right] \phi(q) + i \frac{\lambda}{2} \int_{q_1} \int_{q_2} \int_{q_3} \int_{q_4} \phi^*(q_4) \phi^*(q_3) \phi(q_2) \phi(q_1) \delta(q_3 + q_4 - q_1 - q_2) \right] \quad (4.110)$$

After a Fourier-transform in $\tilde{\omega}$ and \vec{q} one finds, with time labeled now by t

$$S = \int_x [\phi^*(x) \partial_t \phi(x) + \frac{i}{2M} (\vec{\nabla} \phi^*(x)) (\vec{\nabla} \phi(x)) + \frac{i\lambda}{2} (\phi^*(x) \phi(x))^2 + i\lambda \delta \phi^*(x) \phi(x)] \quad (4.111)$$

where

$$x = (t, \vec{x}), \quad \int_x = \int_{-\infty}^{\infty} dt \int d^3\vec{x}. \quad (4.112)$$

The transfer matrix for this functional integral is now

$$T_M = \exp\left\{ -\frac{i(t_f - t_{\text{in}})}{(2N + 1)} H \right\} \quad (4.113)$$

instead of

$$\hat{T} = \exp\left[-\frac{\beta}{(2N + 1)} H \right]. \quad (4.114)$$

The matrix \hat{T}_M is a unitary matrix (for $H^\dagger = H$).

Local Physics For observations and experiments done in some time involved around t the details of boundary conditions at t_f and t_{in} play no role for large $|t_f - t|$ and $|t - t_{\text{in}}|$. Doing physics now is not much influenced by what happened precisely to the dinosaurs or what will happen in the year 10000. For many purposes the boundary term $B(t_f, t_{\text{in}})$ is just an irrelevant multiplicative factor in Z which drops out from the expectation values of interest. One can then simply omit it and work directly with Z_M .

Minkowski action Define the Minkowski action S_M by multiplying the euclidean action S with a factor i

$$S_M = iS, \quad e^{-S} = e^{iS_M}. \quad (4.115)$$

The Minkowski action reads

$$S_M = - \int_x \phi^* (-i\partial_t - \frac{\Delta}{2M}) \phi + \dots \quad (4.116)$$

Variation of S_M or S with respect to ϕ^* yields for $\lambda = 0$ the free Schrodinger equation for a simple particle

$$(-i\partial_t - \frac{\Delta}{2M})\phi = 0 \quad (4.117)$$

$$i\partial_t\phi = H\phi = -\frac{\Delta}{2M}\phi \quad (4.118)$$

There is a reason for that, but the connection needs a few steps, concentrating on simple particle states. Recall that the functional integral describes arbitrary particle numbers. For $\lambda \neq 0$ the *classical field equation* $\frac{\delta S}{\delta \phi^*(x)} = 0$ is a non-linear equation, called Gross-Pitaevskii equation

$$i\partial_t\phi = -\frac{\Delta}{2M}\phi + \lambda(\phi^*\phi)\phi + \lambda\delta\varphi \quad (4.119)$$

This is not a linear Schrodinger equation for a quantum wave function, but has a different interpretation.

Analytic continuation Replacing in the action (4.111)

$$t = -i\tau \quad (4.120)$$

$$\int_x = -i \int d\tau d^3\vec{x} \quad (4.121)$$

$$\partial_t\phi = i\partial_\tau\phi \quad (4.122)$$

we get

$$S = \int d\tau d^3x [\phi^* \left(\partial_\tau - \frac{\Delta}{2M} \right) \phi + \frac{\lambda}{2} (\phi^*\phi)^2 + \lambda\delta\varphi^*\varphi]$$

This is precisely the action for the $T \rightarrow 0$ limit ($\mu = 0$) for thermal equilibrium! Thus the (euclidean) action S for two models, one for the real time evolution, the other for the $T = 0$ limit of thermal equilibrium, are related by analytic continuation. Note that S_M is not the analytic continuation of S , but rather related to S by a fixed definition. The sign of S_M is of historical origin. For Fourier transformation in momentum space

$$\tilde{\omega}\tau = \omega_M t = -i\omega_M \tau \quad (4.123)$$

$$\omega_M = i\tilde{\omega} \equiv q_0 \quad (4.124)$$

Analytic Continuation: Compute quantities first in euclidean space. ($T \rightarrow 0$ limit of thermal equilibrium). Obtain correlation functions in momentum space. Continue the correlation functions analytically to Minkowski space.

$$\tilde{\omega} \rightarrow -iq_0 \quad (4.125)$$

$$\tilde{\omega}^2 \rightarrow -q_0^2 = q_0 q_0 \eta^{00} = q^0 q_0 \quad (4.126)$$

relativistic theory

$$\tilde{\omega}^2 + \vec{q}^2 \rightarrow q^0 q_0 + q^i q_i = q^\mu q_\mu = q^2 \quad (4.127)$$

$$q_E^2 \rightarrow q_M^2 \quad (4.128)$$

Big advantage : euclidean functional integral well defined! Numerical simulations etc. possible.

4.5 Expectation values of time ordered operators

Heisenberg picture in quantum mechanics. $\hat{A}_H(t)$: Heisenberg operators, depend on time.

$$\hat{A}_H(t) = U^\dagger(t, t_{\text{in}}) \hat{A}_s U(t, t_{\text{in}}), \quad \hat{A}_s = \text{operator in Schrodinger picture} \quad (4.129)$$

Consider for $t_2 \geq t_1$

$$\hat{A}_H(t_2) \hat{B}_H(t_1) = U^\dagger(t_2, t_{\text{in}}) \hat{A}_s U(t_2, t_{\text{in}}) U^\dagger(t_1, t_{\text{in}}) \hat{B}_s U(t_1, t_{\text{in}}) \quad (4.130)$$

and use

$$U^\dagger(t_1, t_2) = U(t_2, t_1) \quad (4.131)$$

$$U(t_3, t_2) U(t_2, t_1) = U(t_3, t_1). \quad (4.132)$$

With

$$U(t_2, t_{\text{in}}) U^\dagger(t_1, t_{\text{in}}) = U(t_2, t_1) U(t_1, t_{\text{in}}). U^\dagger(t_1, t_{\text{in}}) = U(t_2, t_1) \quad (4.133)$$

one has

$$\hat{A}_H(t_2) \hat{B}_H(t_1) = U^\dagger(t_2, t_{\text{in}}) \hat{A}_s U(t_2, t_1) \hat{B}_s U(t_1, t_{\text{in}}) \quad (4.134)$$

. In the Heisenberg picture, one keeps fixed $|\psi\rangle = |\psi(t_{\text{in}})\rangle$ and describes the time evolution by the t-dependence of the Heisenberg operators. The transition amplitude

$$\langle \phi(t_{\text{in}}) | \hat{A}_H(t_2) \hat{B}_H(t_1) | \psi(t_{\text{in}}) \rangle = \langle A(t_2) B(t_1) \rangle_{\phi\psi} \quad (4.135)$$

reads in the Schrodinger picture

$$\begin{aligned} \langle A(t_2) B(t_1) \rangle_{\phi\psi} &= \langle \phi(t_{\text{in}}) | U^\dagger(t_2, t_{\text{in}}) \hat{A}_s U(t_2, t_1) \hat{B}_s U(t_1, t_{\text{in}}) | \psi(t_{\text{in}}) \rangle \\ &= \langle \phi(t_2) | \hat{A}_s U(t_2, t_1) \hat{B}_s | \psi(t_1) \rangle. \end{aligned} \quad (4.136)$$

We may insert a complete set of states

$$\int d\chi(t_1) |\chi(t_1)\rangle \langle \chi(t_1)| = 1, \quad (4.137)$$

in order to obtain

$$\begin{aligned} \langle A(t_2) B(t_1) \rangle_{\phi\psi} &= \int d\chi(t_1) \langle \phi(t_2) | \hat{A}_s U(t_2, t_1) | \chi(t_1) \rangle \langle \chi(t_1) | \hat{B}_s | \psi(t_1) \rangle \\ &= \int d\chi(t_1) \langle \phi(t_2) | \hat{A}_s | \chi(t_2) \rangle \langle \chi(t_1) | \hat{B}_s | \psi(t_1) \rangle \end{aligned} \quad (4.138)$$

This has an intuitive interpretation: The transition amplitudes are evaluated for B at time t_1 between $\psi(t_1)$ and arbitrary intermediate states $\chi(t_1)$. Then $\chi(t_1)$ propagates in time to $\chi(t_2)$, and one evaluates the transition amplitude at t_2 of A between $\chi(t_2)$ and $\phi(t_2)$. One finally sums over intermediate states.

Propagator Consider an initial vacuum state $|0\rangle$ for φ and ψ ,

$$|\psi(t_{\text{in}})\rangle = |0\rangle, \quad |\varphi(t_{\text{in}})\rangle = |0\rangle \quad (4.139)$$

Take for \hat{B}_s the creation operator $a^\dagger(\vec{x})$ which creates a particle at position \vec{x} , and for \hat{A}_s the annihilation operator $a(\vec{x}')$ for a particle at \vec{x}' . The state

$$a^\dagger(\vec{x})U(t_1, t_{\text{in}})|0\rangle = |(\vec{x}, t_1); t_1\rangle \quad (4.140)$$

is a one particle state, where the particle sits at \vec{x} at the time t_1 . For $t > t_1$ the particle will move. The wave function changes in the Schrodinger picture.

$$|(\vec{x}, t_1)t\rangle = U(t, t_1)|(\vec{x}, t_1); t_1\rangle \quad (4.141)$$

Note that for $|(\vec{x}, t_1)t\rangle$ the time argument t_1 is a label (together with \vec{x}) specifying which state is meant. It is not the time argument in the Schrodinger evolution of this wave function. The latter is given by t .

The transition amplitude at a given time t with a one particle state $|(\vec{y}, t_2); t\rangle$ determines the probability to find a particle that was at \vec{x} at time t_1 to be a particle that is at \vec{y} at time t_2 . It reads

$$G(\vec{y}, t_2; \vec{x}, t_1) = \langle(\vec{y}, t_2); t|(\vec{x}, t_1); t\rangle. \quad (4.142)$$

Take $t = t_2$:

$$G(\vec{y}, t_2; \vec{x}, t_1) = \langle 0|U^\dagger(t_2, t_{\text{in}})a(\vec{y})U(t_2, t_1)a^\dagger(\vec{x})U(t_1, t_{\text{in}})|0\rangle. \quad (4.143)$$

In the Heisenberg picture this reads

$$G(\vec{y}, t_2; \vec{x}, t_1) = \langle 0|a_H(\vec{y}, t_2)a_H^\dagger(\vec{x}, t_1)|0\rangle. \quad (4.144)$$

The transition amplitude G is called the propagator or Green's function. It is a central quantity in quantum field theory.

One particle wave function A one particle wave function at time t is a superposition

$$\psi_1(t) = \int_{\vec{x}} \varphi(\vec{x}, t)|(\vec{x}, t); t\rangle. \quad (4.145)$$

The position representation of the one-particle wave function $\varphi(\vec{x}, t)$ is defined by

$$\varphi(\vec{x}, t) = \langle(\vec{x}, t); t|\psi_1(t)\rangle \quad (4.146)$$

Proof

$$\begin{aligned} \langle(\vec{x}, t); t|\psi_1(t)\rangle &= \int_{\vec{y}} \langle(\vec{x}, t); t|\varphi(\vec{y}, t)|(\vec{y}, t); t\rangle \\ &= \int_{\vec{y}} \varphi(\vec{y}, t)\langle(\vec{x}, t); t|(\vec{y}, t); t\rangle \\ &= \int_{\vec{y}} \varphi(\vec{y}, t)\delta(\vec{x} - \vec{y}) \\ &= \varphi(\vec{x}, t) \end{aligned} \quad (4.147)$$

Evolution The time evolution of a one particle wave function can be found from the time evolution of $|(\vec{x}, t_1); t\rangle$

$$\begin{aligned}\varphi(\vec{y}, t_2) &= \langle(\vec{y}, t_2); t_2|\psi(t_2)\rangle \\ &= \int_{\vec{x}} \varphi(\vec{x}, t_1) \langle(\vec{y}, t_2); t_2|(\vec{x}, t_1); t_2\rangle \\ &= \int_{\vec{x}} G(\vec{y}, t_2; \vec{x}, t_1) \varphi(\vec{x}, t_1)\end{aligned}\quad (4.148)$$

The propagator G allows one to compute the one-particle wave function at t_2 from an *initial wave function* at t_1 . This is Huygens' principle for the propagation of waves.

Propagator from functional integral We employ the functional integral expression for the evolution operator in the expression

$$\begin{aligned}G(\vec{y}, t_2; \vec{x}, t_1) &= \langle 0|U^\dagger(t_2, t_{\text{in}})a(\vec{y})U(t_2, t_1)a^\dagger(x)U(t_1, t_{\text{in}})|0\rangle \\ &= \langle 0|_f U(t_f, t_2)a(\vec{y})U(t_2, t_1)a^\dagger(x)U(t_1, t_{\text{in}})|0\rangle\end{aligned}\quad (4.149)$$

where

$$\langle 0|_f = \langle 0|U^\dagger(t_f, t_{\text{in}}), \quad (4.150)$$

using

$$U^\dagger(t_f, t_{\text{in}})U(t_f, t_2) = U^\dagger(t_2, t_{\text{in}}). \quad (4.151)$$

One often calls $|0\rangle = |0\rangle_{\text{in}}$ the initial vacuum at t_{in} , and $|0\rangle_f = U(t_f, t_{\text{in}})|0\rangle_{\text{in}}$ the *final vacuum* at t_f . (For a time-translation invariant vacuum one has $|0\rangle_f = |0\rangle_{\text{in}}$.) We have derived before the functional integral expression for the evolution operator

$$U(t_2, t_1) = \int dx(t_2) \int dx(t_1) |x(t_2)\rangle F(t_2, t_1) \langle x(t_1)| \quad (4.152)$$

with

$$F(t_2, t_1) = \int D\varphi(t_1 < t' < t_2) \exp\left\{-\int_{t_1}^{t_2} dt \mathcal{L}(t)\right\} \quad (4.153)$$

The integrals over $x(t_2)$ and $x(t_1)$ are not yet included in $\int D\varphi(t_1 < t' < t_2)$. Recall

$$\hat{x} = \frac{1}{\sqrt{2}}(a^\dagger + a), \quad \hat{p} = \frac{i}{\sqrt{2}}(a^\dagger - a) \quad (4.154)$$

$$\hat{x}|x(t)\rangle = x(t)|x(t)\rangle, \quad \hat{p}|p(t)\rangle = p(t)|p(t)\rangle \quad (4.155)$$

$$a = \frac{1}{\sqrt{2}}(\hat{x} + i\hat{p}), \quad a^\dagger = \frac{1}{\sqrt{2}}(\hat{x} - i\hat{p}) \quad (4.156)$$

For the expression

$$\begin{aligned}U(t_3, t_2)\hat{A}U(t_2, t_1) &= \int dx(t_3) dx'(t_2) dx(t_2) dx(t_1) |x(t_3)\rangle \\ &\quad \cdot F(t_3, t_2) \langle x'(t_2)|\hat{A}|x(t_2)\rangle F(t_2, t_1) \langle x(t_1)|\end{aligned}\quad (4.157)$$

we need the matrix element

$$\langle x'(t_2)|\hat{A}|x(t_2)\rangle = \int \frac{dp}{2\pi} \langle x'(t_2)|p(t_2)\rangle \langle p(t_2)|\hat{A}|x(t_2)\rangle. \quad (4.158)$$

For \hat{A} depending on a^\dagger and a replace

$$a \rightarrow \frac{1}{\sqrt{2}}(x(t_2) + ip(t_2)) \quad (4.159)$$

$$a^\dagger \rightarrow \frac{1}{\sqrt{2}}(x(t_2) - ip(t_2)) \quad (4.160)$$

(If necessary, the ordering of operators has to be performed conveniently.)

$$\begin{aligned} \Rightarrow U(t_3, t_2) \hat{A} U(t_2, t_1) &= \int dx(t_3) dx(t_1) |x(t_3)\rangle \int D\varphi(t_1 < t' < t_3) \\ &\cdot \exp \left\{ - \int_{t_1}^{t_3} dt' \mathcal{L}(t') \right\} A(x(t_2), p(t_2)) \langle x(t_1) | \end{aligned} \quad (4.161)$$

The operator \hat{A} at t_2 leads to the insertion of a function $A(t_2)$ into functional integral. Recall the inverse: an observable $A(t)$ in the functional integral results in the insertion of an operator \hat{A} in the chain of transfer matrices.

We have been here a bit vague with the precise choice of integrations. In a precise discrete formulation one replaces

$$\langle x_{j+1} | e^{-i\Delta t \hat{H}} | x_j \rangle \quad \text{by} \quad \langle x_{j+1} | e^{-i\Delta t \hat{H}} \hat{A} | x_j \rangle \quad (4.162)$$

at the appropriate place in the chain.

We can now follow $A(x(t_2), p(t_2))$ through the chain of variable transformations:

$$x_j \rightarrow \tilde{x}_n \rightarrow \frac{1}{\sqrt{2}}(\varphi_n + \varphi_{-n}^*) \rightarrow \frac{1}{\sqrt{2}}(\varphi(t) + \varphi^*(t)) \quad (4.163)$$

$$p_j \rightarrow \tilde{p}_n \rightarrow -\frac{i}{\sqrt{2}}(\varphi_n - \varphi_{-n}^*) \rightarrow -\frac{i}{\sqrt{2}}(\varphi(t) - \varphi^*(t)), \quad (4.164)$$

resulting in the simple replacement rule

$$a \rightarrow \varphi(t) \quad , \quad a^\dagger \rightarrow \varphi^*(t). \quad (4.165)$$

This yields for the correlation function

$$G(\vec{y}, t_2, \vec{x}, t_1) = Z^{-1} \int D\varphi e^{-S[\varphi]} \varphi(\vec{y}, t_2) \varphi^*(\vec{x}, t_1) \equiv \langle \varphi(\vec{y}, t_2) \varphi^*(\vec{x}, t_1) \rangle. \quad (4.166)$$

For complex functional integrals in Minkowski space we define expectation values similar to classical statistical physics

$$\langle A \rangle = Z^{-1} \int D\varphi e^{-S[\varphi]} A[\varphi] \quad (4.167)$$

$$Z = \int D\varphi e^{-S[\varphi]}. \quad (4.168)$$

Remarks:

-Origin of the normalization factor Z . We have not paid much attention to the normalization of the wave function, the additive normalization of the action, and the formal boundary terms. All this is accounted for by Z^{-1} .

- Since $A[\varphi]$ is a function(functional) of φ , variable transformations are straightforward. (No complications with commutator relations as for a, a^\dagger . The Fourier transform of correlation function reads

$$G(\vec{q}, t_2; \vec{p}, t_1) = \int_y \int_x e^{-i\vec{q}\vec{y}} e^{i\vec{p}\vec{x}} G(\vec{y}, t_2; \vec{x}, t_1) \quad (4.169)$$

Translation symmetry implies

$$G \sim \delta(\vec{q} - \vec{p}). \quad (4.170)$$

- So far we have assumed implicitly that the vacuum is trivial. In general $\langle \varphi(\vec{x}, t) \rangle$ may be different from zero. A more general definition of the (connected) correlation function is given by

$$G(\vec{y}, t_2; \vec{x}, t_1) = \langle \delta\varphi(\vec{y}, t_2) \delta\varphi(\vec{x}, t_1) \rangle, \quad \delta\varphi = \varphi - \langle \varphi \rangle \quad (4.171)$$

Definition of quantum field theory A quantum field theory is defined by

- (1) Choice of fields φ
- (2) Action as functional of fields $S[\varphi]$
- (3) Measure $\int D\varphi$

Correlation function is defined by

$$G_{\alpha\beta} = \langle \varphi_\alpha \varphi_\beta^* \rangle - \langle \varphi_\alpha \rangle \langle \varphi_\beta^* \rangle, \quad (4.172)$$

with α, β collective indices, e.g. $\alpha = (\vec{x}, t)$ or (\vec{p}, t) .

No need of knowledge of vacuum. This is important, since the precise properties of the vacuum for interacting theories are not known.

Chains of operators Consider for $t_n > t_{n-1} > \dots > t_2 > t_1$ a chain of Heisenberg operators

$$\tilde{G} = \langle 0 | \hat{A}_H^{(n)}(t_n) \hat{A}_H^{(n-1)}(t_{n-1}) \dots \hat{A}_H^{(2)}(t_2) \hat{A}_H^{(1)}(t_1) | 0 \rangle \quad (4.173)$$

The Green's function is a special case

$$G = \langle 0 | a_H(t_2) a_H^\dagger(t_1) | 0 \rangle. \quad (4.174)$$

In complete analogy one finds the functional integral expression

$$\tilde{G} = Z^{-1} \int D\varphi e^{-S} \bar{A} = \langle A \rangle \quad (4.175)$$

for the observable

$$\bar{A} = A(t_n) A(t_{n-1}) \dots A(t_2) A(t_1) \quad (4.176)$$

with

$$A(t_n) = A(\varphi^*(t_n), \varphi(t_n)). \quad (4.177)$$

Time ordering The product $A(t')A(t) = A(t)A(t')$ is commutative. The product $\hat{A}_H(t')\hat{A}_H(t)$ in general not. What happens to commutation relations?

Define the time order operator T by putting in a product of two Heisenberg operators the one with larger time argument to the left. e.g. for $t_2 > t_1$

$$\begin{aligned} T(\hat{A}_H^{(2)}(t_2) \hat{A}_H^{(1)}(t_1)) &= \hat{A}_H^{(2)}(t_2) \hat{A}_H^{(1)}(t_1) \\ T(\hat{A}_H^{(1)}(t_1) \hat{A}_H^{(2)}(t_2)) &= \hat{A}_H^{(2)}(t_2) \hat{A}_H^{(1)}(t_1). \end{aligned} \quad (4.178)$$

The time ordered operator product is commutative. Generalize to products with several factors.

$$\langle 0 | T(\hat{A}_H) | 0 \rangle = \langle A \rangle \quad (4.179)$$

On the left one has an operator expression, and on the right functional integral expression.

Transition amplitude for multiparticle states Consider two particles at t_1 with momenta \vec{p}_1 and \vec{p}_2 , and compute the transition amplitude to a two particle state at $t_2 > t_1$ with momenta \vec{p}_3 and \vec{p}_4 .

$$\begin{aligned} \tilde{G}_{2,2} &= \langle 0 | a_H(\vec{p}_4, t_2) a_H(\vec{p}_3, t_2) a_H^\dagger(\vec{p}_2, t_1) a_H^\dagger(\vec{p}_1, t_1) | 0 \rangle \\ &= \langle \varphi(\vec{p}_4, t_2) \varphi(\vec{p}_3, t_2) \varphi^*(\vec{p}_2, t_1) \varphi^*(\vec{p}_1, t_1) \rangle \end{aligned} \quad (4.180)$$

This is a four-point function. It is a basic element of scattering theory.

5 Relativistic scalar fields and O(N)-models

5.1 Lorentz invariant action and antiparticles

Neutral relativistic scalar fields are the neutral pion π^0 in QCD, or the inflaton or cosmon. A scalar field is a real function $\chi(\vec{x}, t)$. In principle, its expectation value can be measured, similar to the electric or magnetic field. Complex scalar fields are the charged pions and the Kaons, represented by a complex scalar field $\chi(\vec{x}, t)$. An important field is the Higgs-doublet, represented by a two-component complex scalar field $\chi_i(t)$, $i=1,2$.

Action The action has to respect the symmetries of the model. For a fundamental theory of elementary particles they always include the Lorentz-symmetry and translations in space and time (Poincare-symmetry). The functional measure is the canonical measure for real or complex functions in four dimensional space. $\chi(x) = \chi(x^\mu)$, $x^\mu = (t, \vec{x})$.

We consider local actions of the form

$$S = \int_x \mathcal{L}(x), \quad \int_x = \int dt d^3\vec{x}. \quad (5.1)$$

$$\mathcal{L}(x) = \mathcal{L}_{kin} + iV + \dots \quad (5.2)$$

Kinetic term The *kinetic term* \mathcal{L}_{kin} involves derivatives of fields. For non-relativistic free atoms we have found

$$\mathcal{L}_{kin} = \chi^*(x)\partial_t\chi(x) + \frac{i}{2M}\partial_i\chi^*(x)\partial_i\chi(x), \quad \partial_i = \frac{\partial}{\partial x^i} = \vec{\nabla}_i \quad (5.3)$$

The two space derivatives are needed for rotation symmetry. Lorentz-symmetry needs again two derivatives,

$$\mathcal{L}_{kin} = i\partial^\mu\chi^*(x)\partial_\mu\chi(x), \quad (5.4)$$

with

$$\partial_\mu = \left(\frac{\partial}{\partial t}, \vec{\nabla}\right) = (\partial_0, \partial_i), \quad (5.5)$$

$$\partial^\mu = \eta^{\mu\nu}\partial_\nu, \quad \eta_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}. \quad (5.6)$$

The scalar product of two four-vectors is invariant. We conclude that relativistic theories of scalars involves two time derivatives. The kinetic term can be formulated for real fields in the same way. Writing a complex field as two real fields ($\chi = \frac{1}{\sqrt{2}}(\chi_1 + i\chi_2)$) one has

$$\mathcal{L}_{kin} = \frac{i}{2} \sum_{a=1}^N \partial^\mu\chi_a(x)\partial_\mu\chi_a(x). \quad (5.7)$$

Here $N = 1$ for a real scalar, $N = 2$ for a complex scalar and $N = 4$ for the Higgs doublet.

Potential The potential V involves no derivatives. It is a function of the fields

$$V(x) = V(\chi(x)) = V(\chi). \quad (5.8)$$

Internal symmetries yield further restrictions. Charge conservation corresponds to the symmetry

$$\chi \rightarrow e^{i\alpha}\chi. \quad (5.9)$$

The potential can only depend on

$$\rho = \chi^*\chi = \frac{1}{2}(\chi_1^2 + \chi_2^2). \quad (5.10)$$

For the Higgs doublet, the symmetry is SU(2) such that

$$\rho = \chi^\dagger \chi = \frac{1}{2} \sum_{a=1}^4 \chi_a^2. \quad (5.11)$$

Often one can expand

$$V(\rho) = \mu^2 \rho + \frac{1}{2} \lambda \rho^2 + \dots \quad (5.12)$$

One infers that

$$\mathcal{L}_{kin} + V(\rho) \quad (5.13)$$

has SO(N)- symmetry.

Two fields with one time-derivative Let us recall that differential equations with two derivatives = two differential equations with one derivative. **One field with two time-derivatives are equivalent to two fields with one time derivative.**

- We consider a free relativistic scalar field (complex):

$$\mathcal{L} = i(\partial^\mu \chi^* \partial_\mu \chi + M^2 \chi^* \chi) \quad (5.14)$$

In momentum space, $\partial_t = \partial_0 = -\partial^0$, one has

$$\mathcal{L}_p = -i\partial_t \chi^* \partial_t \chi + i(p^2 + M^2) \chi^* \chi \quad (5.15)$$

$$Z = \int D\chi e^{-\int dt \int_{\vec{p}} \mathcal{L}_p(t)} \quad (5.16)$$

We treat every \vec{p} mode separately. Let us insert a unit factor

$$\int D\pi \exp\{-i(\partial_t \chi^* - \pi^*)(\partial_t \chi - \pi)\} = \text{const} = \mathbf{1} \quad (5.17)$$

such that

$$Z = \int D\chi D\pi \exp \left[- \int_t \left\{ -i\partial_t \chi^* \partial_t \chi + i(p^2 + M^2) \chi^* \chi \right. \right. \\ \left. \left. + i\partial_t \chi^* \partial_t \chi - i\partial_t \chi^* \pi - i\pi^* \partial_t \chi + i\pi^* \pi \right\} \right]. \quad (5.18)$$

This eliminates the term with two derivatives. What remains are two complex fields χ and π with one time derivative,

$$Z = \int D\chi D\pi e^{-\int_t \mathcal{L}}, \quad (5.19)$$

where, after doing a partial integration.

$$\mathcal{L} = i\chi^* \partial_t \pi - i\pi^* \partial_t \chi + i(p^2 + M^2) \chi^* \chi + i\pi^* \pi \quad (5.20)$$

Perform next a variable transformation

$$\chi(t) = \frac{1}{\sqrt{2}} (p^2 + M^2)^{-\frac{1}{4}} (\varphi_1(t) + \varphi_2(-t)), \\ \pi(t) = -\frac{i}{\sqrt{2}} (p^2 + M^2)^{\frac{1}{4}} (\varphi_1(t) - \varphi_2(-t)). \quad (5.21)$$

This yields

$$\begin{aligned}
(p^2 + M^2)\chi^*(t)\chi(t) &= \frac{1}{2}(p^2 + M^2)^{\frac{1}{2}} [\varphi_1^*(t)\varphi_1(t) + \varphi_2^*(-t)\varphi_2(-t) \\
&\quad + \varphi_1^*(t)\varphi_2(-t) + \varphi_2^*(t)\varphi_1(t)], \\
\pi^*(t)\pi(t) &= \frac{1}{2}(p^2 + M^2)^{\frac{1}{2}} [\varphi_1^*(t)\varphi_1(t) + \varphi_2^*(-t)\varphi_2(-t) \\
&\quad - \varphi_1^*(t)\varphi_2(-t) - \varphi_2^*(t)\varphi_1(t)], \\
i((p^2 + M^2)\chi^*\chi + \pi^*\pi) &= i(p^2 + M^2)^{\frac{1}{2}} [\varphi_1^*(t)\varphi_1(t) + \varphi_2^*(-t)\varphi_2(-t)],
\end{aligned} \tag{5.22}$$

and

$$\begin{aligned}
i(\chi^*\partial_t\pi - \pi^*\partial_t\chi) &= \frac{1}{2} \{(\varphi_1^*(t) + \varphi_2^*(-t))\partial_t(\varphi_1(t) - \varphi_2(-t)) \\
&\quad + (\varphi_1^*(t) - \varphi_2^*(-t))\partial_t(\varphi_1(t) + \varphi_2(-t))\} \\
&= \varphi_1^*(t)\partial_t\varphi_1(t) - \varphi_2^*(-t)\partial_t\varphi_2(-t).
\end{aligned} \tag{5.23}$$

Under the t-integral one can replace $-\varphi_2^*(-t)\partial_t\varphi_2(-t) \rightarrow \varphi_2^*(t)\partial_t\varphi_2(t)$.

Taking the terms together we find the action for two particles with dispersion relation $E = \sqrt{p^2 + M^2}$

$$S = \int dt \left\{ \varphi_1^*\partial_t\varphi_1 + \varphi_2^*\partial_t\varphi_2 - i\sqrt{p^2 + M^2}(\varphi_1^*\varphi_1 + \varphi_2^*\varphi_2) \right\} \tag{5.24}$$

where $\varphi_i = \varphi_i(t)$.

Antiparticles The field χ with two time-derivatives describes a pair of fields φ_1, φ_2 with one time-derivative. One field is the antiparticle of the other. Take χ to be a charged field with coupling to the electromagnetic field.

$$\partial_\mu \rightarrow D_\mu\chi = (\partial_\mu - ieA_\mu)\chi. \tag{5.25}$$

Here A_μ is for this purpose an *external field*. Take $A_i = 0$ and constant electric potential A_0 . This adds to \mathcal{L} an additional term

$$\Delta\mathcal{L} = eA_0 [\chi^*(t)\pi(t) - \pi^*(t)\chi(t)] \tag{5.26}$$

One obtains after transforming to φ_1 and φ_2

$$\begin{aligned}
\Delta\mathcal{L} &= eA_0 \left[-\frac{i}{2}(\varphi_1^*(t) + \varphi_2^*(-t))(\varphi_1(t) - \varphi_2(-t)) - \frac{i}{2}(\varphi_1^*(t) - \varphi_2^*(t))(\varphi_1(t) + \varphi_2(-t)) \right] \\
&= -ieA_0(\varphi_1^*(t)\varphi_1(t) - \varphi_2^*(-t)\varphi_2(-t))
\end{aligned} \tag{5.27}$$

or

$$S = \int dt \{ \varphi_1^*(\partial_t - ieA_0)\varphi_1 + \varphi_2^*(\partial_t + ieA_0)\varphi_2 + \dots \} \tag{5.28}$$

We conclude that φ_1 and φ_2 have opposite electric charge. Same mass, opposite charge means on antiparticle

We conclude that a complex scalar field with two derivatives describes a particle and an antiparticle.

5.2 Unified Scalar field theories

Euclidean space Analytic continuation yields

$$\eta^{\mu\nu}\partial_\mu\partial_\nu \rightarrow \delta^{\mu\nu}\partial_\mu\partial_\nu \tag{5.29}$$

Another factor arises from $dt = -id\tau$. In euclidean space the action therefore reads

$$\Rightarrow S = \int_x \frac{1}{2} \sum_a \partial^\mu\chi_a\partial_\mu\chi_a + V(\rho). \tag{5.30}$$

where now $\partial^\mu = \delta^{\mu\nu} \partial_\nu$ and $\int_x = \int dt \int d^3\vec{x}$. This is the four-dimensional O(N)-model introduced in lecture 2. The euclidean action is also the one that appears for the $T \rightarrow 0$ limit of thermal equilibrium, while for $T > 0$ the τ -integration becomes periodic with period τ .

In euclidean space, the Lorentz-symmetry SO(1,3) gets replaced by the four dimensional rotations SO(4). This symmetry is broken for $T > 0$ since space and time are no longer treated equally. One should distinguish two different symmetries: SO(N) : internal symmetry, SO(d): space symmetry.

Unified description of scalar theories The euclidean O(N)-models in arbitrary dimension d, admit a classical statistical probability distribution, with real action

$$p = Z^{-1} e^{-S}, \quad Z = \int D\varphi e^{-S}. \quad (5.31)$$

They can be simulated on a computer.

$d = 1, 2, 3$	models of classical statistical systems in d-dimensions
$N = 3$	magnets, $\langle \chi_a(x) \rangle$ is magnetisation (order parameter).
$N = 1$	Ising type models.
$N = 2, d = 2$	Two dimensional x-y model with Kosterlitz-Thouless phase transition.
$d = 4$	relativistic scalar theories at $T = 0$.

If the euclidean model is solved, the n-point functions can be analytically continued to Minkowski space

$$q_{0E} = q_E^0 = -iq_{0M} = iq_M^0. \quad (5.32)$$

n-point functions The task is the computation of n-point functions

$$G_{ab\dots f}^{(n)}(x_1 \dots x_n) = \langle \chi_a(x_1) \chi_b(x_2) \dots \chi_f(x_n) \rangle, \quad x \equiv x^\mu \quad (5.33)$$

or in Fourier space

$$G^{(n)}(p_1 \dots p_n), \quad p \equiv p_\mu \quad (5.34)$$

Example of two point function

$$G_{ab}(p_1, p_2) = \langle \chi_a(p_1) \chi_b(p_2) \rangle - \langle \chi_a(p_1) \rangle \langle \chi_b(p_2) \rangle = G(p_1) \delta(p_1 + p_2) \delta_{ab} \quad (5.35)$$

It can only depend on one momentum by virtue of d-dimensional translation symmetry. SO(d)-rotations imply that G can only depend on

$$p^2 = p_\mu p_\nu \delta^{\mu\nu}, \quad i.e. \quad G(p^\mu) = G(p^2) \quad (5.36)$$

Analytic continuation does not change $G(p^2)$, one only has to switch to $p^2 = p_\mu p^\nu \eta^{\mu\nu}$ in momentum space.

5.3 Propagator for free field

$$S = \int_x \left\{ \frac{1}{2} \partial^\mu \chi_a \partial_\mu \chi_a + \frac{1}{2} M^2 \chi_a \chi_a \right\} \quad (5.37)$$

Sum of independent pieces, each particle can be treated separately. Consider for simplicity one complex field

$$S = \int_x \partial^\mu \chi^* \partial_\mu \chi + M^2 \chi^* \chi, \quad (5.38)$$

and transform to Fourier space

$$S = \int_q (q^2 + M^2) \chi^*(q) \chi(q), \quad \int_q = \int \frac{d^d q}{(2\pi)^d}. \quad (5.39)$$

The propagator is defined as

$$G(p, q) = \langle \chi(p) \chi^*(q) \rangle - \langle \chi(p) \rangle \langle \chi^*(q) \rangle. \quad (5.40)$$

We use a torus with discrete modes and take the volume to infinity at the end. For

$$S = \sum_q (q^2 + M^2) \chi^*(q) \chi(q) \quad (5.41)$$

the expectation value obeys

$$\langle \chi(p) \rangle = Z^{-1} \int D\chi \exp(-S) \chi(p) = 0. \quad (5.42)$$

For $p \neq q$, one finds

$$\langle \chi(p) \chi^*(q) \rangle = Z^{-1} \int D\chi e^{-S} \chi(p) \chi^*(q) = 0. \quad (5.43)$$

Only for equal momenta $p = q$ the two point function differs from zero,

$$\begin{aligned} \langle \chi(q) \chi^*(q) \rangle &= Z^{-1} \int D\chi e^{-S} \chi(q) \chi^*(q) \\ &= \frac{\int d\chi(q) e^{-(q^2+M^2)\chi^*(q)\chi(q)} \chi^*(q)\chi(q)}{\int d\chi(q) e^{-(q^2+M^2)\chi^*(q)\chi(q)}} \end{aligned}$$

We can first compute the Gaussian integral

$$Z(M^2) = \int d\chi(q) e^{-(q^2+M^2)\chi^*(q)\chi(q)}, \quad (5.44)$$

and then take the derivative with respect to M^2 ,

$$\langle \chi(q) \chi^*(q) \rangle = -\frac{\partial}{\partial M^2} \ln Z(M^2). \quad (5.45)$$

The Gaussian integral has the solution

$$Z(M^2) = \frac{\pi}{q^2 + M^2}, \quad (5.46)$$

$$-\ln Z = \ln(q^2 + M^2) - \ln \pi, \quad (5.47)$$

$$-\frac{\partial}{\partial M^2} \ln Z = \frac{1}{q^2 + M^2}. \quad (5.48)$$

We can summarise for the free propagator

$$G(q, p) = \frac{1}{q^2 + M^2} \delta(q - p). \quad (5.49)$$

Propagator in Minkowski space The analytic continuation of the free euclidean propagator is straightforward in momentum space:

$$\begin{aligned} G(p, q) &= \frac{1}{(q^2 + M^2)} \delta(p - q) \\ &= \frac{1}{q^\mu q_\mu + M^2} \delta(p - q) \\ &= \frac{1}{-q_0^2 + \vec{q}^2 + M^2} \delta(p - q). \end{aligned}$$

This propagator has poles at

$$q_0 = \pm \sqrt{\vec{q}^2 + M^2}. \quad (5.50)$$

This corresponds to a particle and its antiparticle.

The solutions of the free field equations are

$$\chi_+ = e^{-i\sqrt{q^2+M^2}t} \quad (5.51)$$

$$\chi_- = e^{+i\sqrt{q^2+M^2}t} = e^{-i\sqrt{q^2+M^2}\tilde{t}}, \quad \tilde{t} = -t \quad (5.52)$$

Antiparticles appear as particles *propagating backwards in time*

5.4 Magnetisation in classical statistics

Action $\sigma_a(x)$: magnets at every point x can be viewed as elementary magnets averaged over small volumes. The Hamiltonian with next neighbour interaction reads in the continuum limit

$$H = \int_x K \partial_i \sigma_a(x) \partial_i \sigma_a(x) + c \sigma_a(x) \sigma_a(x) + d (\sigma_a(x) \sigma_a(x))^2 - B \sigma_a(x). \quad (5.53)$$

We take $K > 0$. This tends to align magnets at neighbouring points. The magnetic field B breaks the $O(N)$ -symmetry.

Symmetric magnets $N=3, d=3$

Asymmetric magnets $N=2$ or $N=1$

We have the classical partition function with

$$Z = \int D\sigma e^{-\beta H} = \int D\sigma e^{-S} \quad (5.54)$$

where the classical action is

$$S = \beta H \quad (5.55)$$

Choose fields

$$\sigma_a(x) = \sqrt{\frac{1}{\beta K}} \chi_a(x). \quad (5.56)$$

With this normalisation the action reads

$$S = \int_x \partial_i \chi_a(x) \partial_i \chi_a(x) + \frac{c}{K} \chi_a(x) \chi_a(x) + \frac{d}{\beta K^2} (\chi_a(x) \chi_a(x))^2 - \frac{B\sqrt{\beta}}{K} \chi_a(x) \quad (5.57)$$

or

$$S \equiv \int_x \partial_i \chi_a(x) \partial_i \chi_a(x) + \frac{m^2}{2} \chi_a(x) \chi_a(x) + \frac{\lambda}{8} (\chi_a(x) \chi_a(x))^2 - J \chi_a(x) \quad (5.58)$$

The parameter m^2 can be positive or negative. The name is purely historical, in analogy to the mass term for a relativistic particle.

Magnetisation For $m^2 > 0$ the microscopic magnets have for $J = 0$ a preferred value $\chi_a = 0$.

For $m^2 < 0$ the preferred value differs from zero for $J=0$,

The minimum of the potential

$$V_0(\rho) = m^2 \rho + \frac{\lambda}{2} \rho^2, \quad \rho = \frac{1}{2} \varphi_a \varphi_a, \quad (5.59)$$

obeys $\frac{\partial V_0}{\partial \rho} = m^2 + \lambda \rho = 0$. For $m^2 < 0$ it occurs at $\rho_0 = -\frac{m^2}{\lambda}$. A nonvanishing magnetic field J prefers a certain direction. The minimum of $V = m^2 \rho + \frac{\lambda}{2} \rho^2 - \varphi_a J_a$ defines the microscopic magnetisation.

Question: What is the macroscopic magnetisation $\langle \chi(x) \rangle$ in function of the magnetic field J ?

Fluctuations play a role! We consider $m^2 < 0$ where things are most interesting. The factor e^{-S} is maximal if S is minimal. One first looks for the minimum of S and expands around it. The minimum of S is given by the microscopic magnetisation. Take $J = (J_1, 0, 0)$ The configuration with constant χ , $\chi_a(x) = \chi_0$ minimises the kinetic term. Look for minimum of V ; it occurs in the direction χ_1

$$V = \frac{1}{2} m^2 \chi_1^2 + \frac{\lambda}{8} \chi_1^4 - J \chi_1. \quad (5.60)$$

Minimum of V

$$\frac{\partial V}{\partial \chi} = m^2 \chi_1 + \frac{\lambda}{2} \chi_1^3 - J = 0. \quad (5.61)$$

If we take $J > 0$ a positive χ_{10} is preferred, being the minimum of V .

For small $J > 0$ one has

$$\frac{\lambda}{2}\chi_{10}^2 = -m^2, \quad \chi_{10} = \sqrt{-\frac{2m^2}{\lambda}}. \quad (5.62)$$

Fluctuations tend to wash out the microscopic magnetisation. How strong is this effect?

Compute $Z(J)$.

Then

$$\frac{\partial \ln Z}{\partial J} = \left\langle \int_x \chi_1 \right\rangle = \Omega \langle \chi_1 \rangle = \tilde{M} \quad (5.63)$$

Here \tilde{M} is the magnetisation in appropriate units and Ω the volume. We are interested here in small $J \rightarrow 0$.

Free Energy:

$$\Rightarrow F = -T \ln Z = -\frac{1}{\beta} \ln Z. \quad (5.64)$$

If $\tilde{M}(J \rightarrow 0) \neq 0$, one has **spontaneous symmetry breaking**.

Magnetisation in absence of magnetic field ($J = 0$):

$$\begin{aligned} V &= \frac{\lambda}{2}(\rho - \rho_0)^2, \quad \rho_0 = -\frac{m^2}{\lambda} \\ \rho &= \chi^* \chi = \frac{1}{2}(\chi_1^2 + \chi_2^2) \\ \text{expand around } \chi_{10}, \quad \rho_0 &= \frac{1}{2}\chi_{10}^2 \\ \chi_1 &= \chi_{10} + \delta\chi_1 \\ \frac{1}{2}\chi_1^2 &= \rho_0 + \chi_{10}\delta\chi_1 + \frac{1}{2}\delta\chi_1^2 \\ \rho - \rho_0 &= \chi_{10}\delta\chi_1 + \frac{1}{2}\delta\chi_1^2 + \frac{1}{2}\chi_2^2 \end{aligned} \quad (5.65)$$

Keep only terms quadratic in the fields

$$\frac{\lambda}{2}(\rho - \rho_0)^2 = \frac{\lambda}{2}\chi_{10}^2\delta\chi_1^2 = \lambda\rho_0\delta\chi_1^2$$

$$\delta\chi_1 \text{ behaves as massive field, with } M^2 = 2\lambda\rho_0 \quad (5.66)$$

and propagator

$$G = \frac{1}{q^2 + 2\lambda\rho_0}$$

$$\chi_2 \text{ behaves as massless field (only kinetic term)} \quad (5.67)$$

with propagator

$$G = \frac{1}{q^2} \quad (5.68)$$

The massless field is called a Goldstone boson.

Add small J

$$\begin{aligned} V &= \frac{\lambda}{2}(\rho - \rho_0)^2 - J\chi_1 \\ &= \lambda\rho_0\delta\chi_1^2 - J\chi_{10} - J\delta\chi_1 \end{aligned} \quad (5.69)$$

The action takes the form

$$\begin{aligned} S &= S_0 + \Delta S \\ S_0 &= -\Omega J\chi_{10} \end{aligned} \quad (5.70)$$

$$\Delta S = \int_x \frac{1}{2}\delta\chi_1(x)(-\Delta + 2\lambda\rho_0)\delta\chi_1(x) - J\delta\chi_1(x) + \frac{1}{2}\chi_2(x)(-\Delta)\chi_2(x)$$

Correspondingly, one obtains for the partition function in lowest order

$$\begin{aligned} Z_0 &= e^{-S_0} = \exp(\Omega J\chi_{10}) \\ \ln Z_0 &= \Omega J\chi_{10} \\ M &= \frac{\partial \ln Z_0}{\partial J} = \Omega\chi_{10} \end{aligned} \quad (5.71)$$

What are the corrections from ΔS ? In later lectures.

6 Non-relativistic bosons

6.1 Functional integral for spinless atoms

From relativistic to non-relativistic scalar fields. In this section we go from a relativistic quantum field theory back to non-relativistic physics but in a quantum field theoretic formalism. This non-relativistic QFT is in the few-body limit equivalent to quantum mechanics for a few particles but also has interesting applications to condensed matter physics (many body quantum theory) and it is interesting conceptually. We start from the action of a complex, relativistic scalar field in Minkowski space

$$S = \int dt d^3x \left\{ -\partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi - \frac{\lambda}{2} (\phi^* \phi)^2 \right\} \quad (6.1)$$

The quadratic part can be written in Fourier space with $(px = -p^0 x^0 + \vec{p}\vec{x})$

$$\phi(x) = \int \frac{d^4p}{(2\pi)^4} e^{ipx} \phi(p), \quad \phi^*(x) = \int \frac{d^4p}{(2\pi)^4} e^{-ipx} \phi^*(p), \quad (6.2)$$

as

$$\begin{aligned} S_2 &= - \int \frac{d^4p}{(2\pi)^4} \left\{ \phi^*(p) [-(p^0)^2 + \vec{p}^2 + m^2] \phi(p) \right\} \\ &= - \int \frac{d^4p}{(2\pi)^4} \left\{ \phi^*(p) [-(p^0 + \sqrt{\vec{p}^2 + m^2})(p^0 + \sqrt{\vec{p}^2 + m^2})] \phi(p) \right\}. \end{aligned} \quad (6.3)$$

One observes that the so-called inverse propagator has two zero-crossings, one at $p^0 = \sqrt{\vec{p}^2 + m^2}$ and one at $p^0 = -\sqrt{\vec{p}^2 + m^2}$. At this points the quadratic part of the action become stationary in the sense $\frac{\delta}{\delta \phi^*(p)} S_2 = 0$. The zero-crossings also correspond to poles of the propagator. These so-called on-shell relations give the relation between frequency and momentum for propagating, particle-type excitations of the theory. In fact, $p^0 = \sqrt{\vec{p}^2 + m^2}$ gives the one for particles, $p^0 = -\sqrt{\vec{p}^2 + m^2}$ the one of anti-particles. In the non-relativistic theory, anti-particle excitations are absent. Intuitively, one assumes that the fields are close to fulfilling the dispersion relation for particles, $p^0 = \sqrt{\vec{p}^2 + m^2}$ which is for large m^2 rather far from the frequency of anti-particles. One can therefore replace in a first step

$$p^0 + \sqrt{\vec{p}^2 + m^2} \rightarrow 2\sqrt{\vec{p}^2 + m^2} \approx 2m. \quad (6.4)$$

Moreover, one can expand the dispersion relation for particles for $m^2 \gg \vec{p}^2$,

$$p^0 = \sqrt{\vec{p}^2 + m^2} = m + \frac{\vec{p}^2}{2m} + \dots \quad (6.5)$$

This leads us to a quadratic action of the form

$$S_2 = - \int \frac{d^d p}{(2\pi)^4} \left\{ \phi^*(p) \left(-p^0 + m + \frac{\vec{p}^2}{2m} \right) 2m \phi(p) \right\}, \quad (6.6)$$

or for the full action in position space

$$S = \int dt d^3x \left\{ \phi^* \left(i\partial_t - m + \frac{\vec{\nabla}^2}{2m} \right) 2m \phi - \frac{\lambda}{2} (\phi^* \phi)^2 \right\}. \quad (6.7)$$

It is now convenient to introduce rescaled fields by setting

$$\phi(t, \vec{x}) = \frac{1}{\sqrt{2m}} e^{-i(m-V_0)t} \varphi(t, \vec{x}). \quad (6.8)$$

The action becomes then

$$S = \int dt d^3x \left\{ \varphi^* \left(i\partial_t - V_0 + \frac{\vec{\nabla}^2}{2m} \right) \varphi - \frac{\lambda}{8m^2} (\varphi^* \varphi)^2 \right\}. \quad (6.9)$$

The dispersion relation is now with

$$\varphi(t, \vec{x}) = \int \frac{d\omega}{2\pi} \frac{d^3p}{(2\pi)^3} e^{-i\omega t + i\vec{p}\vec{x}} \varphi(\omega, \vec{p}), \quad (6.10)$$

given by

$$\omega = V_0 + \frac{\vec{p}^2}{2m}. \quad (6.11)$$

This corresponds to the energy of a non-relativistic particle where V_0 is an arbitrary normalization constant corresponding to the offset of an external potential. The action in equation (6.9) describes a non-relativistic field theory for a complex scalar field. As we will see, one can obtain quantum mechanics from there but it is also the starting point for a description of superfluidity.

Symmetries of non-relativistic theory. The non-relativistic action in equation (6.9) has a number of symmetries that are interesting to discuss. First we have translations in space and time as well as rotations in space as in the relativistic case. There is also a global U(1) internal symmetry,

$$\varphi(x) \rightarrow e^{i\alpha} \varphi(x), \quad \varphi^*(x) \rightarrow e^{-i\alpha} \varphi^*(x). \quad (6.12)$$

By Noether's theorem this symmetry is related to particle number conservation (exercise).

Time-dependent U(1) symmetry. There is also an interesting extension of the global U(1) symmetry. One can in fact make it time-dependent according to

$$\varphi(x) \rightarrow e^{i\alpha + \beta t} \varphi(x), \quad \varphi^*(x) \rightarrow e^{-i\alpha + \beta t} \varphi^*(x). \quad (6.13)$$

All terms in the action are invariant except for

$$\varphi^* i\partial_t \varphi \rightarrow \varphi^* e^{-i(\alpha + \beta t)} i\partial_t e^{i(\alpha + \beta t)} \varphi(x) = \varphi^* (i\partial_t - \beta) \varphi. \quad (6.14)$$

However, if we also change $V_0 \rightarrow V_0 - \beta$ we have for the combination

$$\varphi^* (i\partial_t - V_0) \varphi \rightarrow \varphi^* (i\partial_t - \beta - V_0 + \beta) \varphi = \varphi^* (i\partial_t - V_0) \varphi. \quad (6.15)$$

This shows that

$$\varphi(x) \rightarrow e^{i(\alpha + \beta t)} \varphi, \quad \varphi^* \rightarrow e^{-i(\alpha + \beta t)} \varphi^*, \quad V_0 \rightarrow V_0 - \beta, \quad (6.16)$$

is in fact another symmetry of the action in equation eq:nonrelativisticactionScalar. One can say here that $(i\partial_t - V_0)$ acts like a *covariant derivative*. This says that $(i\partial_t - V_0)\varphi$ transforms in the same (covariant) way as φ itself. The physical meaning of this transformation is a change in the absolute energy scale, which is possible in non-relativistic physics.

Galilei transformation. Note that the action in equation (6.9) is not invariant under Lorentz transformations any more. This is directly clear because derivatives with respect to time and space do not enter in an equal way. However, non-relativistic physics is invariant under another kind of space-time transformations, namely Galilei boosts,

$$t \rightarrow t, \quad (6.17)$$

$$\vec{x} \rightarrow \vec{x} + \vec{v}t. \quad (6.18)$$

One can go to another reference frame that moves relative to the original one with a constant velocity. How is this transformation realized in the non-relativistic field theory described by equation (6.9)? This is a little bit complicated and we directly give the transformation law,

$$\varphi(t, \vec{x}) \rightarrow \varphi'(t, \vec{x}) = e^{i(m\vec{v}\cdot\vec{x} - \frac{1}{2}m\vec{v}^2t)} \varphi(t, \vec{x} - \vec{v}t). \quad (6.19)$$

Indeed one can confirm that

$$\left(i\partial_t + \frac{\vec{\nabla}^2}{2m}\right) \varphi(t, \vec{x}) \rightarrow e^{i(m\vec{v}\cdot\vec{x} - \frac{1}{2}m\vec{v}^2t)} \left[i\partial_t + \frac{\vec{\nabla}^2}{2m}\right] \varphi(t, \vec{x} - \vec{v}t), \quad (6.20)$$

so that the action (6.9) is invariant under Galilei transformations.

6.2 Spontaneous symmetry breaking: Bose-Einstein condensation and superfluidity

Effective potential. One can write the action in (6.9) also as

$$S = \int dt d^3x \left\{ \varphi^* \left(i\partial_t + \frac{\vec{\nabla}^2}{2m}\right) \varphi - V(\varphi^* \varphi) \right\}, \quad (6.21)$$

with microscopic potential as a function of $\rho = \varphi^* \varphi$,

$$V(\rho) = V_0 \rho + \frac{\lambda}{2} \rho^2 = -\mu \rho + \frac{\lambda}{2} \rho^2. \quad (6.22)$$

At non-vanishing density one has $V_0 = -\mu$, where μ is the chemical potential. For, $\mu > 0$ the minimum of the effective potential is at $\rho_0 > 0$. In a classical approximation where the effect of fluctuation is neglected, one has the equation of motion following from $\delta S = 0$.

Bose-Einstein condensate. If the solution $\varphi(x) = \phi_0$ is homogeneous (constant in space and time), it must correspond to a minimum of the effective potential. Without loss of generality we can assume $\phi_0 \in \mathbb{R}$ and

$$V'(\rho_0) = -\mu + \lambda \rho_0 = 0, \quad (6.23)$$

leads to

$$\phi_0 = \sqrt{\rho_0} = \sqrt{\frac{\mu}{\lambda}}. \quad (6.24)$$

Assuming that it survives the effect of quantum fluctuations, such a field expectation value breaks the global U(1) symmetry spontaneously, similar to magnetization. This phenomenon is known as Bose-Einstein condensation. One can see this as a macroscopic manifestation of quantum physics. The mode with vanishing momentum $\vec{p} = 0$ has a macroscopically large occupation number, which is possible for bosonic particles. On the other side, it arises here in a classical approximation to the quantum field theory described by the action in eq. (6.9). In this sense, a Bose-Einstein condensate can also be seen as a *classical* field, similar to the electro-magnetic field, for example.

Bogoliulov excitations. It is also interesting to study small perturbations around the homogeneous field value ϕ_0 . Let us write

$$\varphi(x) = \phi_0 + \frac{1}{\sqrt{2}} [\phi_1(x) + i \phi_2(x)], \quad (6.25)$$

with real fields $\phi_1(x)$ and $\phi_2(x)$. The action in eq. (6.21) becomes (up to total derivatives)

$$S = \int dt d^3x \left\{ \phi_2 \partial_t \phi_1 + \frac{1}{2} \sum_{j=1}^2 \phi_j \frac{\vec{\nabla}^2}{2m} \phi_j - V \left(\phi_0^2 + \sqrt{2} \phi_0 \phi_1 + \frac{1}{2} \phi_1^2 + \frac{1}{2} \phi_2^2 \right) \right\}. \quad (6.26)$$

It is instructive to expand to quadratic order in the deviations from a homogeneous field ϕ_1 and ϕ_2 . The quadratic part of the action reads

$$S_2 = \int dt d^3x \left\{ -\frac{1}{2}(\phi_1, \phi_2) \begin{pmatrix} -\frac{\nabla^2}{2m} + 2\lambda\phi_0^2 & \partial_t \\ -\partial_t & -\frac{\nabla^2}{2m} \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \right\}. \quad (6.27)$$

In momentum space, the matrix between the fields becomes

$$G^{-1}(\omega, \vec{p}) = \begin{pmatrix} \frac{\vec{p}^2}{2m} + 2\lambda\phi_0^2 & -i\omega \\ i\omega & \frac{\vec{p}^2}{2m} \end{pmatrix}. \quad (6.28)$$

In cases where the inverse propagator is a matrix, this holds also for the propagator. When the determinant of the inverse propagator has a zero-crossing, the propagator has a pole. This defines the dispersion relation for quasi-particle excitations,

$$\det G^{-1}(\omega, \vec{p}) = 0. \quad (6.29)$$

Here this leads to

$$-\omega^2 + \left(\frac{\vec{p}^2}{2m} + 2\lambda\phi_0^2 \right) \frac{\vec{p}^2}{2m} = 0, \quad (6.30)$$

or

$$\omega = \sqrt{\left(\frac{\vec{p}^2}{2m} + 2\lambda\phi_0^2 \right) \frac{\vec{p}^2}{2m}}. \quad (6.31)$$

This is known as Bogoliubov dispersion relation.

For small momenta, such that

$$\frac{\vec{p}^2}{2m} \ll 2\lambda\phi_0^2, \quad (6.32)$$

one finds

$$\omega \approx \sqrt{\frac{\lambda\phi_0^2}{m}} |\vec{p}|. \quad (6.33)$$

In contrast, for

$$\frac{\vec{p}^2}{2m} \gg 2\lambda\phi_0^2, \quad (6.34)$$

one recovers the usual dispersion relation for non-relativistic particles

$$\omega \approx \frac{\vec{p}^2}{2m}. \quad (6.35)$$

The low-momentum region describes phonons (quasi-particles of sound excitations), while the large-momentum region describes normal particles.

Superfluidity. The fact that the dispersion relation is linear for small momenta is also responsible for another interesting phenomenon, namely superfluidity, a fluid motion without viscosity.

7 Scattering

In this section we will discuss a rather useful concept in quantum field theory – the S-matrix. It describes situations where the incoming state is a perturbation of a symmetric (homogeneous and isotropic) vacuum state in terms of particle excitations and the outgoing state similarly. We are interested in calculating the transition amplitude, and subsequently transition probability, between such few-particle states. An important example is the scattering of two particles with a certain center-of-mass energy. This is an experimental situation in many high energy laboratories, for

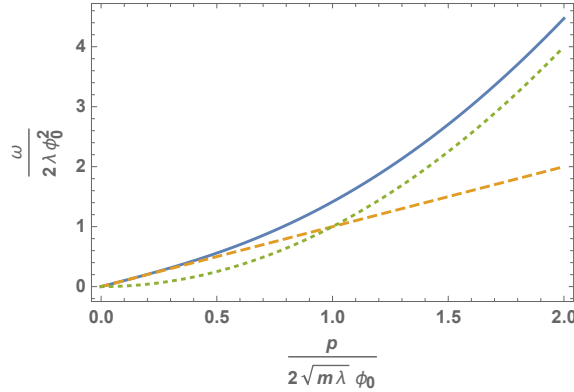


Figure 3. Bogoliubov dispersion relation as in eq. (6.31) (solid line). Also shown is the low momentum approximation (6.33) (dashed line) and the large-momentum approximation (6.35) (dotted line).

example at CERN. The final states consists again of a few particles (although “few” might be rather many if the collision energy is high). Another interesting example is a single incoming particle, or resonance, that can be unstable and decay into other particles. For example $\pi^+ \rightarrow \mu^+ + \nu_\mu$. As we will discuss later on in more detail, particles as excitations of quantum fields are actually closely connected with symmetries of space-time, in particular translations in space and time as well as Lorentz transformations including rotations. (In the non-relativistic limit, Lorentz transformations are replaced by Galilei transformations). The standard application of the S-matrix concept assumes therefore that the vacuum state has these symmetries. The S-matrix is closely connected to the functional integral. Technically, this connection is somewhat simpler to establish for non-relativistic quantum field theories. This will be discussed in the following. The relativistic case will be discussed in full glory in the second part of the lecture course.

7.1 Scattering of non-relativistic bosons

Mode function expansion. Let us recall that one can expand fields in the operator picture as follows

$$\varphi(t, \vec{x}) = \int_{\vec{p}} v_{\vec{p}}(t, \vec{x}) a_{\vec{p}}, \quad \varphi^\dagger(t, \vec{x}) = \int_{\vec{p}} v_{\vec{p}}^*(t, \vec{x}) a_{\vec{p}}^\dagger, \quad (7.1)$$

with $\int_{\vec{p}} = \int \frac{d^3 p}{(2\pi)^3}$, annihilation operators $a_{\vec{p}}$, creation operators $a_{\vec{p}}^\dagger$, and the *mode functions*

$$v_{\vec{p}}(t, \vec{x}) = e^{-i\omega_{\vec{p}}t + i\vec{p}\vec{x}}. \quad (7.2)$$

The dispersion relation in the non-relativistic limit is

$$\omega_{\vec{p}} = \frac{\vec{p}^2}{2m} + V_0. \quad (7.3)$$

Note that in contrast to the relativistic case, the expansion of $\varphi(t, \vec{x})$ contains no creation operator and the one of $\varphi^*(t, \vec{x})$ no annihilation operator. This is a consequence of the absence of anti-particles.

For the following discussion, it is useful to introduce a scalar product between two functions of space and time $f(t, \vec{x})$ and $g(t, \vec{x})$,

$$(f, g)_t = \int d^3 x \{f^*(t, \vec{x})g(t, \vec{x})\}. \quad (7.4)$$

The integer goes over the spatial coordinates at fixed time t . Note that if f and g were solutions of the non-relativistic, single-particle Schrödinger equation, the above scalar product were actually independent of time t as a consequence of unitarity in non-relativistic quantum mechanics.

The mode functions are normalized with respect to this scalar product as

$$(v_{\vec{p}}, v_{\vec{p}'})_t = (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{p}'). \quad (7.5)$$

One can write

$$\begin{aligned} a_{\vec{p}} &= (v_{\vec{p}}, \varphi)_t = \int d^3x e^{i\omega_{\vec{p}}t - i\vec{p}\vec{x}} \varphi(t, \vec{x}), \\ a_{\vec{p}}^\dagger &= (v_{\vec{p}}^*, \varphi^*)_t = \int d^3x e^{i\omega_{\vec{p}}t - i\vec{p}\vec{x}} \varphi^*(t, \vec{x}). \end{aligned} \quad (7.6)$$

The right hand side depends on time t and it is instructive to take the time derivative,

$$\begin{aligned} \partial_t a_{\vec{p}}(t) &= \int d^3x e^{i\omega_{\vec{p}}t - i\vec{p}\vec{x}} [\partial_t + i\omega_{\vec{p}}] \varphi(t, \vec{x}) \\ &= \int d^3x e^{i\omega_{\vec{p}}t - i\vec{p}\vec{x}} \left[\partial_t + i \left(\frac{\vec{p}^2}{2m} + V_0 \right) \right] \varphi(t, \vec{x}) \\ &= \int d^3x e^{i\omega_{\vec{p}}t - i\vec{p}\vec{x}} \left[\partial_t + i \left(-\frac{\vec{\nabla}^2}{2m} + V_0 \right) \right] \varphi(t, \vec{x}). \end{aligned} \quad (7.7)$$

We used here first the dispersion relation and expressed them \vec{p}^2 as a derivative acting on the mode function (it acts to the left). In a final step one can use partial integration to make the derivative operator act to the right,

$$\partial_t a_{\vec{p}}(t) = i \int d^3x e^{i\omega_{\vec{p}}t - i\vec{p}\vec{x}} \left[-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right] \varphi(t, \vec{x}). \quad (7.8)$$

This expression confirms that $a_{\vec{p}}$ were time-independent if $\varphi(t, \vec{x})$ were a solution of the one-particle Schrödinger equation. More general, it is a time-dependent, however. In a similar way one finds (exercise)

$$\partial_t a_{\vec{p}}^\dagger(t) = -i \int d^3x e^{-i\omega_{\vec{p}}t + i\vec{p}\vec{x}} \left[i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right] \varphi^*(t, \vec{x}). \quad (7.9)$$

Incoming states. To construct the S-matrix, we first need incoming and out-going states. Incoming states can be constructed by the creation operator

$$a_{\vec{p}}^\dagger(-\infty) = \lim_{t \rightarrow -\infty} a_{\vec{p}}^\dagger(t). \quad (7.10)$$

For example, an incoming two-particle state would be

$$|\vec{p}_1, \vec{p}_2; \text{in}\rangle = a_{\vec{p}_1}^\dagger(-\infty) a_{\vec{p}_2}^\dagger(-\infty) |0\rangle. \quad (7.11)$$

Bosonic exchange symmetry. We note as an aside point that these state automatically obey bosonic exchange symmetry

$$|\vec{p}_1, \vec{p}_2; \text{in}\rangle = |\vec{p}_2, \vec{p}_1; \text{in}\rangle, \quad (7.12)$$

as a consequence of

$$a_{\vec{p}_1}^\dagger(-\infty) a_{\vec{p}_2}^\dagger(-\infty) = a_{\vec{p}_2}^\dagger(-\infty) a_{\vec{p}_1}^\dagger(-\infty). \quad (7.13)$$

Fock space. We note also general states of few particles can be constructed as

$$|\psi; \text{in}\rangle = C_0|0\rangle + \int_{\vec{p}} C_1(\vec{p}) |\vec{p}; \text{in}\rangle + \int_{\vec{p}_1, \vec{p}_2} C_2(\vec{p}_1, \vec{p}_2) |\vec{p}_1, \vec{p}_2; \text{in}\rangle + \dots \quad (7.14)$$

This is a superposition of vacuum (0 particles), 1-particle states, 2-particle states and so on. The space of such states is known as *Fock space*. In the following we will sometimes use an abstract index α to label all the states in Fock space, i. e. $|\alpha; \text{in}\rangle$ is a general incoming state. These states are complete in the sense such that

$$\sum_{\alpha} |\alpha; \text{in}\rangle \langle \alpha; \text{in}| = \mathbb{1}, \quad (7.15)$$

and normalized such that $\langle \alpha; \text{in} | \beta; \text{in} \rangle = \delta_{\alpha\beta}$.

Outgoing states. In a similar way to incoming states one can construct outgoing states with the operators

$$a_{\vec{p}}^{\dagger}(\infty) = \lim_{t \rightarrow \infty} a_{\vec{p}}^{\dagger}(t). \quad (7.16)$$

For example

$$|\vec{p}_1, \vec{p}_2; \text{out}\rangle = a_{\vec{p}_1}^{\dagger}(\infty) a_{\vec{p}_2}^{\dagger}(\infty) |0\rangle. \quad (7.17)$$

7.2 The S-matrix

The S-matrix denotes now simply the transition amplitude between incoming and out-going general states $|\alpha; \text{in}\rangle$ and $|\beta; \text{out}\rangle$,

$$S_{\beta\alpha} = \langle \beta; \text{out} | \alpha; \text{in} \rangle. \quad (7.18)$$

Because α labels all states in Fock space, the S-matrix is a rather general and powerful object. It contains the vacuum-to-vacuum transition amplitude as well as transition amplitudes between all particle-like excited states.

Unitarity of the S-matrix. Let us first prove that the scattering matrix is unitary,

$$\begin{aligned} (S^{\dagger} S)_{\alpha\beta} &= \sum_{\gamma} (S^{\dagger})_{\alpha\gamma} S_{\gamma\beta} \\ &= \sum_j \langle \gamma; \text{out} | \alpha; \text{in} \rangle^* \langle \gamma; \text{out} | \beta; \text{in} \rangle \\ &= \sum_j \langle \alpha; \text{in} | \gamma; \text{out} \rangle \langle \gamma; \text{out} | \beta; \text{in} \rangle \\ &= \langle \alpha; \text{in} | \beta; \text{in} \rangle \\ &= \delta_{\alpha\beta}. \end{aligned} \quad (7.19)$$

We have used here the completeness of the out states

$$\sum_j |\gamma; \text{out}\rangle \langle \gamma; \text{out}| = \mathbb{1}. \quad (7.20)$$

Conservation laws. The S-matrix respects a number of conservation laws such as for energy and momentum. There can also be conservation laws for particle numbers, in particular also in the non-relativistic domain. One distinguishes between elastic collisions where particle numbers do not change, e.g. $2 \rightarrow 2$, and inelastic collisions, such as $2 \rightarrow 4$. In a non-relativistic theory, such inelastic processes can occur for bound states, for example two H_2 - molecules can scatter into their constituents

$$H_2 + H_2 \rightarrow 4H. \quad (7.21)$$

Connection between outgoing and incoming states. What is the connection between incoming and outgoing states? Let us write

$$\begin{aligned} a_{\vec{p}}(\infty) - a_{\vec{p}}(-\infty) &= \int_{-\infty}^{\infty} \partial_t a_{\vec{p}}(t) \\ &= i \int_{-\infty}^{\infty} dt \int d^3x e^{i\omega_{\vec{p}}t - i\vec{p}\vec{x}} \left[-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right] \varphi(t, \vec{x}). \end{aligned} \quad (7.22)$$

Annihilation operators at asymptotically large incoming and outgoing times differ by an integral over space-time of the Schrödinger operator acting on the field. In momentum space with $(px = -p^0x^0 + \vec{p}\vec{x} = -p^0t + \vec{p}\vec{x})$,

$$\varphi(t, \vec{x}) = \int \frac{dp^0}{2\omega} \frac{d^3\vec{p}}{(2\pi)^3} e^{ipx} \varphi(p), \quad (7.23)$$

this would read

$$a_{\vec{p}}(\infty) - a_{\vec{p}}(-\infty) = i \left[-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right] \varphi(p). \quad (7.24)$$

In a similar way one finds

$$\begin{aligned} a_{\vec{p}}^\dagger(\infty) - a_{\vec{p}}^\dagger(-\infty) &= -i \int_{-\infty}^{\infty} dt \int d^3x e^{-i\omega_{\vec{p}}t + i\vec{p}\vec{x}} \left[-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right] \varphi^*(t, \vec{x}) \\ &= -i \left[-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right] \varphi^*(p). \end{aligned} \quad (7.25)$$

Relation between S-matrix elements and correlation functions. For concreteness, let us consider $2 \rightarrow 2$ scattering with incoming state

$$|\vec{p}_1, \vec{p}_2; \text{in}\rangle = a_{\vec{p}_1}^\dagger(-\infty) a_{\vec{p}_2}^\dagger(-\infty) |0\rangle, \quad (7.26)$$

and out-going state

$$|\vec{q}_1, \vec{q}_2; \text{out}\rangle = a_{\vec{q}_1}^\dagger(\infty) a_{\vec{q}_2}^\dagger(\infty) |0\rangle. \quad (7.27)$$

The S-matrix element can be written as

$$\begin{aligned} S_{\vec{q}_1, \vec{q}_2, \vec{p}_1, \vec{p}_2} &= \langle \vec{q}_1, \vec{q}_2; \text{out} | \vec{p}_1, \vec{p}_2; \text{in} \rangle \\ &= \langle 0 | T \{ a_{\vec{q}_1}(\infty) a_{\vec{q}_2}(\infty) a_{\vec{p}_1}^\dagger(-\infty) a_{\vec{p}_2}^\dagger(-\infty) \} | 0 \rangle. \end{aligned} \quad (7.28)$$

We have inserted a time-ordering symbol but the operators are time-ordered already anyway. Now, one can use

$$a_{\vec{q}_1}(\infty) = a_{\vec{q}_1}(-\infty) + i \left[-q_1^0 + \frac{\vec{q}_1^2}{2m} + V_0 \right] \psi(q_1). \quad (7.29)$$

However, $a_{\vec{q}_1}(-\infty)$ is moved to the right by time ordering and leads to a vanishing contribution because of

$$a_{\vec{q}_1}(-\infty) |0\rangle = 0. \quad (7.30)$$

So, effectively under time ordering, one can replace

$$a_{\vec{q}_1}(\infty) \rightarrow i \left[-q_1^0 + \frac{\vec{q}_1^2}{2m} + V_0 \right] \varphi(q_1). \quad (7.31)$$

By a similar argument, one can replace creation operators for $t \rightarrow -\infty$ like

$$a_{\vec{p}_1}^\dagger(-\infty) \rightarrow i \left[-p_1^0 + \frac{\vec{p}_1^2}{2m} + V_0 \right] \varphi^*(p_1). \quad (7.32)$$

The above argument is not fully correct. There is one contribution from the operators $a_{\vec{q}}(-\infty)$ we have forgotten here. In fact, the replacements $a_{\vec{q}_1}(\infty) \rightarrow a_{\vec{q}_1}(-\infty)$ and $a_{\vec{q}_2}(\infty) \rightarrow a_{\vec{q}_2}(-\infty)$ give

$$\langle 0 | a_{\vec{q}_1}(-\infty) a_{\vec{q}_2}(-\infty) a_{\vec{p}_1}^\dagger(-\infty) a_{\vec{p}_2}(-\infty) | 0 \rangle. \quad (7.33)$$

We need to commute the annihilation operators to the right using the commutation relation

$$\left[a_{\vec{q}}(-\infty), a_{\vec{p}}^\dagger(-\infty) \right] = (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}). \quad (7.34)$$

This gives rise to a contribution to the S-matrix element

$$(2\pi)^6 \left[\delta^{(3)}(\vec{p}_1 - \vec{q}_1) \delta^{(3)}(\vec{p}_2 - \vec{q}_2) + \delta^{(3)}(\vec{p}_1 - \vec{q}_2) \delta^{(3)}(\vec{p}_2 - \vec{q}_1) \right]. \quad (7.35)$$

But this is just the “transition” amplitude for the case that no scattering has occurred! There is always this trivial part of the S-matrix and in fact one can write

$$S_{\alpha\beta} = \delta_{\alpha\beta} + \text{contributions from interactions}. \quad (7.36)$$

Let us keep this in mind and concentrate on the contribution from interactions in the following. We obtain thus for the S-matrix element

$$\begin{aligned} & \langle \vec{q}_1, \vec{q}_2; \text{out} | \vec{p}_1, \vec{p}_2; \text{in} \rangle \\ &= i^4 \left[-q_1^0 + \frac{\vec{q}_1^2}{2m} + V_0 \right] \left[-q_2^0 + \frac{\vec{q}_2^2}{2m} + V_0 \right] \left[-p_1^0 + \frac{\vec{p}_1^2}{2m} + V_0 \right] \left[-p_2^0 + \frac{\vec{p}_2^2}{2m} + V_0 \right] \\ & \times \langle 0 | T \{ \varphi(q_1) \varphi(q_2) \varphi^*(p_1) \varphi^*(p_2) \} | 0 \rangle. \end{aligned} \quad (7.37)$$

This shows how S-matrix elements are connected to time ordered correlation functions. This relation is known as the *Lehmann-Symanzik-Zimmermann (LSZ) reduction formula*, here applied to non-relativistic quantum field theory.

The time-ordered correlation functions can be written as functional integrals,

$$\langle 0 | T \{ \varphi(q_1) \varphi(q_2) \varphi^*(p_1) \varphi^*(p_2) \} | 0 \rangle = \frac{\int D\varphi \varphi(q_1) \varphi(q_2) \varphi^*(p_1) \varphi^*(p_2) e^{iS[\varphi]}}{\int D\varphi e^{iS[\varphi]}}. \quad (7.38)$$

We can now calculate S-matrix elements from functional integrals!

Relativistic scalar theories. Let us mention here that for a relativistic theory the LSZ formula is quite similar but one needs to replace

$$\left[-q^0 + \frac{\vec{q}^2}{2m} + V_0 \right] \rightarrow \left[-(q^0)^2 + \vec{q}^2 + m^2 \right], \quad (7.39)$$

and for particles $\varphi(q) \rightarrow \phi(q)$, $\varphi^*(q) \rightarrow \phi^*(q)$, while for anti-particles $\varphi(q) \rightarrow \phi^*(-q)$, $\varphi^*(q) \rightarrow \phi(-q)$.

7.3 Perturbation theory for interacting scalar fields

Let us now consider a non-relativistic theory with the action

$$S[\varphi] = \int dt d^3x \left\{ \varphi^* \left(i\partial_t + \frac{\nabla^2}{2m} - V_0 \right) \varphi - \frac{\lambda}{2} (\varphi^* \varphi)^2 \right\}. \quad (7.40)$$

Compared to equation (6.9) we have rescaled the interaction parameter, $\frac{\lambda}{4m^2} \rightarrow \lambda$. We introduce now the partition function in the presence of source terms J as

$$Z[J] = \int D\varphi \exp \left[iS[\varphi] + i \int_x \{ J^*(x) \varphi(x) + J(x) \varphi^*(x) \} \right], \quad (7.41)$$

with $x = (t, \vec{x})$ and $\int_x = \int dt \int d^3x$. The source term can also be written in momentum space,

$$\int_x \{J^*(x)\varphi(x) + J(x)\varphi^*(x)\} = \int_p \{J^*(p)\varphi(p) + J(p)\varphi^*(p)\}, \quad (7.42)$$

where

$$\varphi(x) = \int_p e^{ipx} \varphi(p), \quad \varphi^*(x) = \int_p e^{-ipx} \varphi^*(p), \quad (7.43)$$

with

$$\int_p = \int \frac{dp^0}{2\pi} \frac{d^3\vec{p}}{(2\pi)^3}, \quad (7.44)$$

and similar for J . Because the source term has the same form in position and momentum space, we will sometimes simply write it as

$$\int \{J^* \varphi + \varphi^* J\}. \quad (7.45)$$

One can generate correlation functions from functional derivatives of $Z[J]$, for example

$$\begin{aligned} \langle \varphi(x)\varphi^*(y) \rangle &= \langle 0|T \{\varphi(x)\varphi^*(y)\} |0\rangle \\ &= \frac{\int D\varphi \varphi(x) \varphi^*(y) e^{iS[\varphi]}}{\int D\varphi e^{iS[\varphi]}} \\ &= \left(\frac{(-i)^2}{Z[J]} \frac{\delta^2}{\delta J^*(x)\delta J(y)} Z[J] \right)_{J=0}. \end{aligned} \quad (7.46)$$

One can also take functional derivatives directly in momentum space, for example

$$\frac{\delta}{\delta J^*(P)} \exp \left[i \int \{J^* \varphi + \varphi^* J\} \right] = \frac{i}{(2\pi)^4} \varphi(p) \exp \left[i \int \{J^* \varphi + \varphi^* J\} \right]. \quad (7.47)$$

In this sense one can write

$$\langle \varphi(p) \varphi^*(q) \rangle = \left(\frac{(-i)^2}{Z[J]} (2\pi)^8 \frac{\delta^2}{\delta J^*(p)\delta J(q)} Z[J] \right)_{J=0}. \quad (7.48)$$

Perturbation theory for partition function. Let us write the partition function formally as

$$Z[J] = \int D\varphi \exp \left[-i\frac{\lambda}{2} \int_x \left(-i\frac{\delta}{\delta J(x)} \right)^2 \left(-i\frac{\delta}{\delta J^*(x)} \right)^2 \right] \exp \left[iS_2[\varphi] + i \int \{J^* \varphi + \varphi^* J\} \right], \quad (7.49)$$

where the quadratic action is

$$S_2[\varphi] = \int_x \varphi^* \left(i\partial_t + \frac{\vec{\nabla}^2}{2m} - V_0 \right) \varphi. \quad (7.50)$$

Note that when acting on the source term in the exponent, every functional derivative $-i\frac{\delta}{\delta J(x)}$ results in a field $\varphi^*(x)$ and so on. In this way, the quartic interaction term has been separated and written in terms of derivatives with respect to the source field. We can now pull it out of the functional integral and write

$$Z[J] = \exp \left[-i\frac{\lambda}{2} \int_x \left(-i\frac{\delta}{\delta J(x)} \right)^2 \left(-i\frac{\delta}{\delta J^*(x)} \right)^2 \right] Z_2[J], \quad (7.51)$$

with the partition function for the quadratic theory

$$Z_2[J] = \int D\varphi e^{iS_2[\varphi] + i \int \{J^* \varphi + \varphi^* J\}}. \quad (7.52)$$

The latter is rather easy to evaluate this in momentum space. One can write

$$\begin{aligned}
S_2 + \int \{J^* \varphi + \varphi^* J\} &= \int_p \left\{ -\varphi^* \left(-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right) \varphi + J^* \varphi + \varphi^* J \right\} \\
&= \int_p \left\{ - \left[\varphi^* - J^* \left(-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right)^{-1} \right] \left(-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right) \right. \\
&\quad \times \left. \left[\varphi - \left(-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right)^{-1} J \right] \right\} \\
&\quad + \int_p \left\{ J^*(p) \left(-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right)^{-1} J(p) \right\}.
\end{aligned} \tag{7.53}$$

Note that the last term is independent of the field φ and can be pulled out of the functional integral. The functional integral over φ is of Gaussian form. One can shift the integration variable

$$\left[\varphi - \left(-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right)^{-1} J \right] \rightarrow \varphi, \tag{7.54}$$

and perform the functional integration in $Z_2[\varphi]$. It yields then only an irrelevant constant and as a result one finds

$$Z_2[J] = \exp \left[i \int_p J^*(p) \left(-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right)^{-1} J(p) \right]. \tag{7.55}$$

In the following it will be useful to write also the interaction term in momentum space. One may use

$$\frac{\delta}{\delta J(x)} = \int d^4 p \frac{\delta J(p)}{\delta J(x)} \frac{\delta}{\delta J(p)} = \int \frac{d^4 p}{(2\pi)^4} e^{-ipx} (2\pi)^4 \frac{\delta}{\delta J(p)} = \int \frac{d^4 p}{(2\pi)^4} e^{-ipx} \delta_{J(p)} = \int_p e^{-ipx} \delta_{J(p)}. \tag{7.56}$$

Here we defined the abbreviation

$$\delta_{J(p)} = (2\pi)^4 \frac{\delta}{\delta J(p)}. \tag{7.57}$$

In a similar way

$$\frac{\delta}{\delta J^*(x)} = \int_p e^{ipx} \delta_{J^*(p)}. \tag{7.58}$$

Using also

$$\int_x e^{ipx} = (2\pi)^4 \delta^{(4)}(p), \tag{7.59}$$

one finds for the partition function

$$\begin{aligned}
Z[J] &= \exp \left[-i \frac{\lambda}{2} \int_x \left(\frac{\delta}{\delta J(x)} \right)^2 \left(\frac{\delta}{\delta J^*(x)} \right)^2 \right] Z_2[J] \\
&= \exp \left[-i \frac{\lambda}{2} \int_{k_1 \dots k_4} \left\{ (2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4) \delta_{J(k_1)} \delta_{J(k_2)} \delta_{J^*(k_3)} \delta_{J^*(k_4)} \right\} \right] \\
&\quad \times \exp \left[i \int_p J^*(p) \left(-p^0 + \frac{\vec{p}^2}{2m} + V_0 \right)^{-1} J(p) \right].
\end{aligned} \tag{7.60}$$

One can now expand the exponential to obtain a formal perturbation series in λ .

Let us now come back to the S-matrix element for $2 \rightarrow 2$ scattering

$$\begin{aligned}
&\langle \vec{q}_1, \vec{q}_2; \text{out} | \vec{p}_1, \vec{p}_2; \text{in} \rangle \\
&= i^4 \left[-q_1^0 + \frac{\vec{q}_1^2}{2m} + V_0 \right] \left[-q_2^0 + \frac{\vec{q}_2^2}{2m} + V_0 \right] \left[-p_1^0 + \frac{\vec{p}_1^2}{2m} + V_0 \right] \left[-p_2^0 + \frac{\vec{p}_2^2}{2m} + V_0 \right] \\
&\quad \times \langle \varphi(q_1) \varphi(q_2) \varphi^*(p_1) \varphi^*(p_2) \rangle \\
&= i^4 \left[-q_1^0 + \frac{\vec{q}_1^2}{2m} + V_0 \right] \left[-q_2^0 + \frac{\vec{q}_2^2}{2m} + V_0 \right] \left[-p_1^0 + \frac{\vec{p}_1^2}{2m} + V_0 \right] \left[-p_2^0 + \frac{\vec{p}_2^2}{2m} + V_0 \right] \\
&\quad \times \left(\frac{1}{Z[J]} \delta_{J^*(q_1)} \delta_{J^*(q_2)} \delta_{J(p_1)} \delta_{J(p_2)} Z[J] \right)_{J=0}.
\end{aligned} \tag{7.61}$$

If we now insert the perturbation expansion for $Z[J]$, we can concentrate on the contribution at order $\lambda^1 = \lambda$, because at order $\lambda^0 = 1$ we have only the trivial S-matrix element for no scattering that we already discussed. At order λ we have different derivatives acting on $Z_2[J]$,

- $\delta_{J(p_1)}$ for incoming particles with momentum \vec{p}_1
- $\delta_{J^*(q_1)}$ for outgoing particle with momentum \vec{q}_1
- $\delta_{J(k)}$ and $\delta_{J^*(k)}$ for the interaction term.

At the end, all these derivatives are evaluated at $J = J^* = 0$. Therefore, there must always be derivatives δ_J and δ_{J^*} acting together on one integral appearing in $Z_2[J]$. Note that

$$\delta_{J(p_1)}\delta_{J^*(q_1)}\left[i\int_p J^*(p)\left(-p^0 + \frac{\vec{p}^2}{2m} + V_0\right)^{-1}J(p)\right] = i\left(-p_1^0 + \frac{\vec{p}_1^2}{2m} + V_0\right)^{-1}(2\pi)^4\delta^{(4)}(p_1 - q_1). \quad (7.62)$$

This implies that if two derivatives representing external particles would hit the same integral in $Z_2[J]$, one would have no scattering because $\vec{p}_1 = \vec{q}_1$ and as a result of momentum conservation then also $\vec{p}_2 = \vec{q}_2$. Only if a derivative representing an incoming or outgoing particle is combined with a derivative from the interaction term, this is avoided. By doing the algebra one finds at order λ

$$\langle\vec{q}_1, \vec{q}_2; \text{out}|\vec{p}_1, \vec{p}_2; \text{in}\rangle = -i\frac{\lambda}{2}4(2\pi)^4\delta^{(4)}(q_1 + q_2 - p_1 - p_2). \quad (7.63)$$

The factor $4 = 2 \times 2$ comes from different ways to combine functional derivatives with sources. The overall Dirac function makes sure that the incoming four-momentum equals the out-going four-momentum,

$$p^{\text{in}} = p_1 + p_2 = q_1 + q_2 = p^{\text{out}}. \quad (7.64)$$

Quite generally, one can define for the non-trivial part of an S-matrix

$$\langle\beta; \text{out}|\alpha; \text{in}\rangle = (2\pi)^4\delta^{(4)}(p^{\text{out}} - p^{\text{in}})i\mathcal{T}_{\beta\alpha}. \quad (7.65)$$

Together with the trivial part from “no scattering”, one can write

$$S_{\beta\alpha} = \delta_{\beta\alpha} + (2\pi)^4\delta^{(4)}(p^{\text{out}} - p^{\text{in}})i\mathcal{T}_{\beta\alpha}. \quad (7.66)$$

By comparison of expressions we find for the $2 \rightarrow 2$ scattering of non-relativistic bosons at lowest order in λ simply

$$\mathcal{T} = -2\lambda, \quad (7.67)$$

independent of momenta. More generally, the transition amplitude \mathcal{T} is expected to depend on the momenta of incoming and outgoing particles.

7.4 From the S-matrix to a cross-section

Let us start from an S-matrix element in the form

$$\langle\beta; \text{out}|\alpha; \text{in}\rangle = (2\pi)^4\delta^{(4)}(p^{\text{out}} - p^{\text{in}})i\mathcal{T} \quad (7.68)$$

with transition amplitude \mathcal{T} which may depend on the momenta itself. (For $2 \rightarrow 2$ scattering of non-relativistic bosons, and at lowest order in λ , we found simply $\mathcal{T} = -2\lambda$.) Let us now discuss how one can relate S-matrix elements to actual scattering cross-sections that can be measured in an experiment. We start by writing the transition probability from a state α to a state β as

$$P = \frac{|\langle\beta; \text{out}|\alpha; \text{in}\rangle|^2}{\langle\beta; \text{out}|\beta; \text{out}\rangle\langle\alpha; \text{in}|\alpha; \text{in}\rangle} \quad (7.69)$$

The numerator contains a factor

$$\left[(2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}})\right]^2 = (2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}}) (2\pi)^4 \delta^{(4)}(0). \quad (7.70)$$

This looks ill defined but becomes meaningful in a finite volume V and for finite time interval ΔT . In fact

$$(2\pi)^4 \delta^4(0) = \int d^4x e^{i0x} = V\Delta T. \quad (7.71)$$

For the transition rate $\dot{P} = \frac{P}{\Delta T}$ we can therefore write

$$\dot{P} = \frac{V(2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}}) |\mathcal{T}|^2}{\langle \beta; \text{out} | \beta; \text{out} \rangle \langle \alpha; \text{in} | \alpha; \text{in} \rangle}. \quad (7.72)$$

Moreover, for incoming and outgoing two-particle states, their normalization is obtained from

$$\begin{aligned} \langle \vec{p}_1, \vec{p}_2; \text{in} | \vec{q}_1, \vec{q}_2; \text{in} \rangle &= \lim_{\vec{q}_j \rightarrow \vec{p}_j} \langle \vec{p}_1, \vec{p}_2; \text{in} | \vec{p}_1, \vec{p}_2; \text{in} \rangle \\ &= \lim_{\vec{q}_j \rightarrow \vec{p}_j} \left[(2\pi)^6 \left(\delta^{(3)}(\vec{p}_1 - \vec{q}_1) \delta^{(3)}(\vec{p}_2 - \vec{q}_2) + \delta^{(3)}(\vec{p}_1 - \vec{q}_2) \delta^{(3)}(\vec{p}_2 - \vec{q}_1) \right) \right] \\ &= \left[(2\pi)^3 \delta^{(3)}(0) \right]^2 \\ &= V^2. \end{aligned} \quad (7.73)$$

In a finite volume $V = L^3$, and with periodic boundary conditions, the final momenta are of the form

$$\vec{p} = \frac{2\pi}{L}(m, n, l), \quad (7.74)$$

with some integer numbers m, n, l . We can count final states according to

$$\sum_{m,n,l} = \sum_{m,n,l} \Delta m \Delta n \Delta l = L^3 \sum_{m,n,l} \frac{\Delta p_1 \Delta p_2 \Delta p_3}{(2\pi)^3}. \quad (7.75)$$

In the continuum limit this becomes

$$V \int \frac{d^3 p}{(2\pi)^3}. \quad (7.76)$$

The differential transition rate has one factor $V \frac{d^3 p}{(2\pi)^3}$ for each final state particle. For $2 \rightarrow 2$ scattering,

$$d\dot{P} = (2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}}) |\mathcal{T}|^2 \frac{1}{V} \frac{d^3 q_1}{(2\pi)^3} \frac{d^3 q_2}{(2\pi)^3}. \quad (7.77)$$

We can go from the transition probability to a cross-section by dividing through the flux of incoming particles

$$\mathcal{F} = \frac{1}{V} v = \frac{2|\vec{p}_1|}{mV}. \quad (7.78)$$

Here we have a density of one particle per volume V and the relative velocity of the two particles is $v = \frac{2|\vec{p}_1|}{m}$, in the center-of-mass frame where $|\vec{p}_1| = |\vec{p}_2|$, for identical particles with equal mass m . This cancels the last factor V and we find for the differential cross-section

$$d\sigma = \frac{|\mathcal{T}|^2 m}{2|\vec{p}_1|} (2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}}) \frac{d^3 q_1}{(2\pi)^3} \frac{d^3 q_2}{(2\pi)^3}. \quad (7.79)$$

In the center-of-mass frame one has also $\vec{p}^{\text{in}} = \vec{p}_1 + \vec{p}_2 = 0$ and accordingly

$$\delta^{(4)}(p^{\text{out}} - p^{\text{in}}) = \delta(E^{\text{out}} - E^{\text{in}}) \delta^{(3)}(\vec{q}_1 + \vec{q}_2). \quad (7.80)$$

The three-dimensional part can be used to perform the integral over \vec{q}_2 . In doing these integrals over final state momenta, a bit of care is needed because the two final state particles are indistinguishable. An outgoing state $|\vec{q}_1, \vec{q}_2; \text{out}\rangle$ equals the state $|\vec{q}_2, \vec{q}_1; \text{out}\rangle$. Therefore, in order to count only really different final states, one must divide by a factor 2 if one simply integrates d^3q_1 and d^3q_2 independently. Keeping this in mind, we find for the differential cross-section after doing the integral over \vec{q}_2 ,

$$d\sigma = \frac{|\mathcal{T}|^2 m}{2|\vec{p}_1|(2\pi)^2} \delta(E^{\text{out}} - E^{\text{in}}) d^3q_1. \quad (7.81)$$

We can now use

$$d^3\vec{q}_1 = |\vec{q}_1|^2 d|\vec{q}_1| d\Omega_{q_1} \quad (7.82)$$

where $d\Omega_{q_1}$ is the differential solid angle element. Moreover

$$E^{\text{out}} = \frac{\vec{q}_1^2}{2m} + \frac{\vec{q}_2^2}{2m} + 2V_0 = \frac{\vec{q}_1^2}{m} + 2V_0, \quad (7.83)$$

and

$$\frac{dE^{\text{out}}}{d|\vec{q}_1|} = 2\frac{|\vec{q}_1|}{m}. \quad (7.84)$$

With this, and using the familiar relation $\delta(f(x)) = \delta(x - x_0)/|f'(x_0)|$, one can perform the integral over the magnitude $|\vec{q}_1|$ using the Dirac function $\delta(E^{\text{out}} - E^{\text{in}})$. This yields $|\vec{q}_1| = |\vec{p}_1|$ and

$$d\sigma = \frac{|\mathcal{T}|^2 m^2}{16\pi^2} d\Omega_{q_1}. \quad (7.85)$$

For the simple case where \mathcal{T} is independent of the solid angle ω_{q_1} , we can calculate the total cross-section. Here we must now take into account that only half of the solid angle 4π corresponds to physically independent configurations. The total cross-section is therefore

$$\sigma = \frac{|\mathcal{T}|^2 m^2}{8\pi}. \quad (7.86)$$

In a final step we use $\mathcal{T} = -2\lambda$ to lowest order in λ (equivalent to the Born approximation in quantum mechanics) and find here the cross-section

$$\sigma = \frac{\lambda^2 m^2}{2\pi}. \quad (7.87)$$

Let us check the dimensions. The action

$$S = \int dt d^3x \left\{ \varphi^* \left(i\partial_t + \frac{\vec{\nabla}^2}{2m} - V_0 \right) \varphi - \frac{\lambda}{2} (\phi^* \phi)^2 \right\} \quad (7.88)$$

must be dimensionless. The field φ must have dimension

$$[\varphi] = \text{length}^{-\frac{3}{2}}. \quad (7.89)$$

The interaction strength λ must accordingly have dimension

$$[\lambda] = \frac{\text{length}^3}{\text{time}}. \quad (7.90)$$

Because

$$\left[\frac{\vec{\nabla}^2}{2m} \right] = \frac{1}{\text{time}}, \quad (7.91)$$

one has $[m] = \frac{\text{time}}{\text{length}^2}$ and therefore $[\lambda m] = \text{length}$. It follows that indeed

$$[\sigma] = \text{length}^2 \quad (7.92)$$

as appropriate for a cross-section.

8 Fermions

So far we have discussed bosonic fields and bosonic particles as their excitations. Let us now turn to fermions. Fermions as quantum particles differ in two central aspects from bosons. First, they satisfy fermionic statistics. Wave functions for several particles are anti-symmetric under the exchange of particles and occupation numbers of modes can only be 0 or 1. Second, fermionic particles have half integer spin, i. e. $1/2$, $3/2$, and so on, in contrast to bosonic particles which have integer spin 0, 1, 2 and so on. Both these aspects lead to interesting new developments. Half-integer spin in the context of relativistic theories leads to a new and deeper understanding of space-time symmetries and fermionic statistics leads to a new kind of functional integral based on anti-commuting numbers. The latter appears already for functional integral representations of non-relativistic quantum fields. We will start with this second-aspect and then turn to aspects of space-time symmetry for relativistic theories later on.

8.1 Non-relativistic fermions

Pauli spinor fields. In non-relativistic quantum mechanics, particles with spin $1/2$ are described by a variant of Schrödinger's equation with two-component fields. The fields are so-called Pauli spinors with components describing spin-up and spin-down parts with respect to some axis. One can write this as

$$\Psi(t, \vec{x}) = \begin{pmatrix} \psi_\uparrow(t, \vec{x}) \\ \psi_\downarrow(t, \vec{x}) \end{pmatrix} \quad (8.1)$$

We also use the notation $\psi_a(t, \vec{x})$ where $a = 1, 2$ and

$$\psi_1(t, \vec{x}) = \psi_\uparrow(t, \vec{x}), \quad \psi_2(t, \vec{x}) = \psi_\downarrow(t, \vec{x}). \quad (8.2)$$

The Pauli equation is a generalisation of Schrödinger's equation (neglecting spin-orbit coupling),

$$\left[\left(-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right) \mathbb{1} + \mu_B \vec{\sigma} \cdot \vec{B} \right] \Psi(t, \vec{x}) = 0, \quad (8.3)$$

or equivalently

$$\left[\left(-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right) \delta_{ab} + \mu_B \vec{\sigma}_{ab} \cdot \vec{B} \right] \psi_b(t, \vec{x}) = 0. \quad (8.4)$$

Here we use the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (8.5)$$

and $\vec{B} = (B_1, B_2, B_3)$ is the magnetic field, while μ_B is the magneton that quantifies the magnetic moment. Based on this, one would expect that the quadratic part of an action for a non-relativistic field describing spin- $1/2$ particles is of the form

$$S_2 = \int dt d^3x \left\{ -\Psi^\dagger \left[\left(-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right) \mathbb{1} + \mu_B \vec{\sigma} \cdot \vec{B} \right] \Psi \right\} \quad (8.6)$$

However, we also need to take care of fermionic (anti-symmetric) exchange symmetry, such that for fermionic states

$$|\vec{p}_1, \vec{p}_2; \text{in}\rangle = -|\vec{p}_2, \vec{p}_1; \text{in}\rangle. \quad (8.7)$$

To this aspect we turn next.

Grassmann variables. So-called *Grassmann variables* are generators θ_i of an algebra, and they are anti-commuting such that

$$\theta_i\theta_j + \theta_j\theta_i = 0. \quad (8.8)$$

An immediate consequence is that $\theta_j^2 = 0$. If there is a finite set of generators $\theta_1, \theta_2, \dots, \theta_n$, one can write general elements of the Grassmann algebra as a linear superposition (with coefficients that are ordinary complex (or real) numbers) of the following basis elements

$$\begin{aligned} &1, \\ &\theta_1, \theta_2, \dots, \theta_n, \\ &\theta_1\theta_2, \theta_1\theta_3, \dots, \theta_2\theta_3, \theta_2\theta_4, \dots, \theta_{n-1}\theta_n, \\ &\dots \\ &\theta_1\theta_2\theta_3 \dots \theta_n. \end{aligned}$$

There are 2^n such basis elements, because each Grassmann variable θ_j can be either present or absent.

Grade of monomial. To a monomial $\theta_{j_1} \dots \theta_{j_q}$ one can associate a *grade* q which counts the number of generators in the monomial. For A_p and A_q being two such monomials one has

$$A_p A_q = (-1)^{p \cdot q} A_q A_p. \quad (8.9)$$

In particular, the monomials of even grade

$$\begin{aligned} &1, \\ &\theta_1\theta_2, \theta_1\theta_3, \dots, \theta_2\theta_3, \dots, \theta_{n-1}\theta_n, \\ &\dots \end{aligned}$$

commute with other monomials, be the latter of even or odd grade.

Grassmann parity. One can define a Grassmann parity transformation P that acts on all generators according to

$$P(\theta_j) = -\theta_j, \quad P^2 = \mathbb{1}. \quad (8.10)$$

Even monomials are even, odd monomials are odd under this transformation. The parity even part of the algebra, spanned by the monomials of even grade, constitutes a sub-algebra. Because its elements commute with other elements of the algebra they behave “bosonic”, while elements of the Grassmann algebra that are odd with respect to P behave “fermionic”.

Functions of Grassmann variables. Because of $\theta^2 = 0$, functions of a Grassmann variable θ are always linear,

$$f(\theta) = f_0 + \theta f_1. \quad (8.11)$$

Note that f_0 and f_1 could depend on other Grassmann variables but not θ .

Differentiation for Grassmann variables. To define differentiation of $f(\theta)$ with respect to θ we first bring it to the form

$$f(\theta) = f_0 + \theta f_1 \quad (8.12)$$

and set then

$$\frac{\partial}{\partial \theta} f(\theta) = f_1. \quad (8.13)$$

Note that similar to $\theta^2 = 0$ one has also $(\frac{\partial}{\partial\theta})^2 = 0$. One may verify that the chain rule applies. Take $\sigma(\theta)$ to be an odd element and $x(\theta)$ an even element of the Grassmann algebra. One has then

$$\frac{\partial}{\partial\theta} f(\sigma(\theta), x(\theta)) = \frac{\partial\sigma}{\partial\theta} \frac{\partial f}{\partial\sigma} + \frac{\partial x}{\partial\theta} \frac{\partial f}{\partial x}. \quad (8.14)$$

The derivative we use here is a left derivative.

Consider for example

$$f = f_0 + \theta_1 \theta_2. \quad (8.15)$$

One has then

$$\frac{\partial}{\partial\theta_1} f = \theta_2, \quad \frac{\partial}{\partial\theta_2} f = -\theta_1, \quad \frac{\partial}{\partial\theta_2} \frac{\partial}{\partial\theta_1} f = 1, \quad \frac{\partial}{\partial\theta_1} \frac{\partial}{\partial\theta_2} f = -1. \quad (8.16)$$

One could also define a right derivative such that

$$f \overleftarrow{\frac{\partial}{\partial\theta_1}} = -\theta_2, \quad f \overleftarrow{\frac{\partial}{\partial\theta_2}} = \theta_1. \quad (8.17)$$

Integration for Grassmann variables. To define integration for Grassmann variables one takes orientation from two properties of integrals from $-\infty$ to ∞ for ordinary numbers. One such property is linearity,

$$\int_{-\infty}^{\infty} dx c f(x) = c \int_{-\infty}^{\infty} dx f(x). \quad (8.18)$$

The other is invariance under shifts of the integration variable,

$$\int_{-\infty}^{\infty} dx f(x+a) = \int_{-\infty}^{\infty} dx f(x). \quad (8.19)$$

For a function of a Grassmann variable

$$f(\theta) = f_0 + \theta f_1 \quad (8.20)$$

One sets therefore

$$\int d\theta f(\theta) = f_1. \quad (8.21)$$

Note that one has formally

$$\int d\theta f(\theta) = \frac{\partial}{\partial\theta} f(\theta). \quad (8.22)$$

In other words, we have defined

$$\int d\theta = 0, \quad \int d\theta \theta = 1. \quad (8.23)$$

This is indeed linear and makes sure that

$$\int d\theta f(\theta + \sigma) = \int d\theta \{(f_0 + \sigma f_1) + f_1 \theta\} = \int d\theta f(\theta) = f_1. \quad (8.24)$$

For functions of several variables one has

$$\int d\theta_1 \int d\theta_2 f(\theta_1, \theta_2) = \frac{\partial}{\partial\theta_1} \frac{\partial}{\partial\theta_2} f(\theta_1, \theta_2). \quad (8.25)$$

It is easy to see that derivatives with respect to Grassmann variables anti-commute

$$\frac{\partial}{\partial\theta_j} \frac{\partial}{\partial\theta_k} = -\frac{\partial}{\partial\theta_k} \frac{\partial}{\partial\theta_j}, \quad (8.26)$$

and accordingly also the differentials anti-commute

$$d\theta_j d\theta_k = -d\theta_k d\theta_j. \quad (8.27)$$

Functions of several Grassmann variables. A function that depends on a set of Grassmann variables $\theta_1, \dots, \theta_n$ can be written as

$$f(\theta) = f_0 + \theta_j f_1^j + \frac{1}{2} \theta_{j_1} \theta_{j_2} f_2^{j_1 j_2} + \dots + \frac{1}{n!} \theta_{j_1} \dots \theta_{j_n} f_n^{j_1 \dots j_n}. \quad (8.28)$$

We use here Einsteins summation convention with indices j_k being summed over. The coefficients $f_k^{j_1 \dots j_k}$ are completely anti-symmetric with respect to the interchange of any part of indices. In particular, the last coefficient can only be of the form

$$f_n^{j_1 \dots j_n} = \tilde{f}_n \epsilon_{j_1 \dots j_n}, \quad (8.29)$$

where $\epsilon_{j_1 \dots j_n}$ is the completely anti-symmetric Levi-Civita symbol in n dimensions with $\epsilon_{12 \dots n} = 1$. Let us now discuss what happens if we differentiate or integrate $f(\theta)$. One has

$$\frac{\partial}{\partial \theta_k} f(\theta) = f_1^k + \theta_{j_2} f_2^{kj_2} + \dots + \frac{1}{(n-1)!} \theta_{j_2} \dots \theta_{j_n} f_n^{kj_2 \dots j_n} \quad (8.30)$$

and similar for higher order derivatives. In particular

$$\frac{\partial}{\partial \theta_n} \dots \frac{\partial}{\partial \theta_1} f(\theta) = f_n^{12 \dots n} = \tilde{f}_n. \quad (8.31)$$

This defines also the integral with respect to all n variables,

$$\int d\theta_n \dots d\theta_1 f(\theta) = f_n^{12 \dots n} = \tilde{f}_n = \int d^n \theta f(\theta) \equiv \int D\theta f(\theta). \quad (8.32)$$

Linear change of Grassmann variables. Let us consider a linear change of the Grassmann variables in the form

$$\theta_j = J_{jk} \theta'_k \quad (8.33)$$

where J_{jk} is a matrix of commuting variables. We can write

$$f(\theta) = f_0 + \dots + \frac{1}{n!} (J_{i_1 j_1} \theta'_{j_1}) \dots (J_{i_n j_n} \theta'_{j_n}) \epsilon_{i_1 \dots i_n} \tilde{f}_n. \quad (8.34)$$

Now one can use the identity

$$\epsilon_{i_1 \dots i_n} J_{i_1 j_1} \dots J_{i_n j_n} = \det(J) \epsilon_{j_1 \dots j_n}. \quad (8.35)$$

This can actually be seen as the definition of the determinant. One can therefore write

$$f(\theta) = f_0 + \dots + \frac{1}{n!} \theta'_{j_1} \dots \theta'_{j_n} \epsilon_{j_1 \dots j_n} \det(J) \tilde{f}_n. \quad (8.36)$$

The integral with respect to θ' is

$$\int d^n \theta' f(\theta) = \det(J) \tilde{f}_n. \quad (8.37)$$

In summary, one has

$$\int d^n \theta f(\theta) = \frac{1}{\det(J)} \int d^n \theta' f(\theta). \quad (8.38)$$

One should compare this to the corresponding relation for conventional integrals with $x_j = J_{jk} x'_k$. In that case one has

$$\int d^n x f(x) = \det(J) \int d^n x' f(x'). \quad (8.39)$$

Note that the determinant appears in the denominator for Grassmann variables while it appears in the numerator for conventional integrals.

Gaussian integrals of Grassmann variables. Consider a Gaussian integral of two Grassmann variables

$$\int d\theta d\xi e^{-\theta\xi b} = \int d\theta d\xi (1 - \theta\xi b) = \int d\theta d\xi (1 + \xi\theta b) = b. \quad (8.40)$$

For a Gaussian integral over conventional complex variables one has instead

$$\int d(\operatorname{Re} x) d(\operatorname{Im} x) e^{-x^* x b} = \frac{\pi}{b}. \quad (8.41)$$

Again, integrals over Grassmann and ordinary variables behave in some sense “inverse”. For higher dimensional Gaussian integrals over Grassmann numbers we write

$$\int d^n \theta d^n \xi e^{-\theta_j a_{jk} \xi_k} = \int d\theta_n d\xi_n \cdots d\theta_1 d\xi_1 e^{-\theta_j a_{jk} \xi_k}. \quad (8.42)$$

One can now employ two unitary matrices to perform a change of variables

$$\theta_j = \theta'_i U_{lj}, \quad \xi_k = V_{km} \xi'_m \quad (8.43)$$

such that

$$U_{lj} a_{jk} V_{km} = \tilde{a}_l \delta_{lm} \quad (8.44)$$

is diagonal. This is always possible. The Gaussian integral becomes

$$d^n \theta d^n \xi e^{-\theta_j a_{jk} \xi_k} = \det(U)^{-1} \det(V)^{-1} \int d^n \theta' d^n \xi' e^{-\theta'_i \xi'_i \tilde{a}_i} = \prod_{l=1}^n \tilde{a}_l = \det(a_{jk}). \quad (8.45)$$

Again this is in contrast to a similar integral over commuting variables where the determinant would appear in the denominator.

Finally let us consider a Gaussian integral with source forms,

$$\int d^n \bar{\psi} d^n \psi \exp[-\bar{\psi} M \psi + \bar{\eta} \psi + \bar{\psi} \eta] = Z(\bar{\eta}, \eta). \quad (8.46)$$

We integrate here over independent Grassmann variables $\psi = (\psi_1, \dots, \psi_n)$ and $\bar{\psi} = (\bar{\psi}_1, \dots, \bar{\psi}_n)$ and we use the abbreviation

$$\bar{\psi} M \psi = \bar{\psi}_j M_{jk} \psi_k. \quad (8.47)$$

The source forms are also Grassmann variables $\eta = (\eta_1, \dots, \eta_n)$ and $\bar{\eta} = (\bar{\eta}_1, \dots, \bar{\eta}_n)$ with

$$\bar{\eta} \psi = \bar{\eta}_j \psi_j, \quad \bar{\psi} \eta = \bar{\psi}_j \eta_j. \quad (8.48)$$

As usual, we can write

$$Z(\bar{\eta}, \eta) = \int d^n \bar{\psi} d^n \psi \exp[-(\bar{\psi} - \eta M^{-1}) M (\psi - M^{-1} \eta) + \bar{\eta} M^{-1} \eta]. \quad (8.49)$$

A shift of integration variables does not change the result and thus we find

$$Z(\bar{\eta}, \eta) = \det(M) \exp[\bar{\eta} M^{-1} \eta]. \quad (8.50)$$

In this sense, Gaussian integrals over Grassmann variables can be manipulated similarly as Gaussian integrals over commuting variables. Note again that $\det(M)$ appears in the numerator while it would appear in the denominator of bosonic variables.

We can now take the limit $n \rightarrow \infty$ and write

$$\int d^n \bar{\psi} d^n \psi \rightarrow \int D\bar{\psi} D\psi, \quad Z(\bar{\eta}, \eta) \rightarrow Z[\bar{\eta}, \eta], \quad (8.51)$$

with

$$Z[\bar{\eta}, \eta] = \int D\bar{\psi} D\psi \exp[-\bar{\psi} M \psi + \bar{\eta} \psi + \bar{\psi} \eta] = \det(M) \exp[\bar{\eta} M^{-1} \eta]. \quad (8.52)$$

In this way we obtain a formalism that can be used for fermionic or Grassmann fields.

Action for free non-relativistic scalars. We can now write down an action for non-relativistic fermions with spin 1/2. It looks similar to what we have conjectured before,

$$S_2 = \int dt d^3x \left\{ -\bar{\psi} \left[\left(-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right) \mathbb{1} + \mu_B \vec{\sigma} \cdot \vec{B} \right] \psi \right\}, \quad (8.53)$$

but the two-component fields $\psi = (\psi_1, \psi_2)$ and $\bar{\psi} = (\bar{\psi}_1, \bar{\psi}_2)$ are in fact *Grassmann fields*. Such fields anti-commute, for example $\psi_1(x)\psi_2(y) = -\psi_2(y)\psi_1(x)$. One should see the field at different space-time positions x to be independent Grassmann numbers. Also, ψ_1 and $\bar{\psi}_1$ are independent as Grassmann fields. In particular $\psi_1(x)^2 = 0$ but $\bar{\psi}_1(x)\psi_1(x) \neq 0$. A partition function with sources for the above free theory could be written down as

$$Z_2[\bar{\eta}, \eta] = \int D\bar{\psi} D\psi \exp \left[iS_2[\bar{\psi}, \psi] + i \int_x \{ \bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x) \} \right] \quad (8.54)$$

Correlation functions can be obtained from functional derivatives of $Z[\bar{\eta}, \eta]$ with respect to the source field $\bar{\eta}(x)$ and $\eta(x)$. Some care is needed to take minus signs into account that may arise from possible commutation of Grassmann numbers. For the quadratic theory one can easily complete the square, perform the functional integral and write the partition function formally as

$$Z_2[\bar{\eta}, \eta] = \exp \left[i \int_x \bar{\eta}(x) \left[\left(-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right) \mathbb{1} + \mu_B \vec{\sigma} \cdot \vec{B} \right]^{-1} \eta(x) \right]. \quad (8.55)$$

The inverse of the operator

$$\left(-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right) \mathbb{1} + \mu_B \vec{\sigma} \cdot \vec{B} \quad (8.56)$$

is a matrix valued Greens function. For a magnetic field that is constant in space and time, for example pointing in z -direction, one can easily invert this operator in Fourier space,

$$\Upsilon(x-y) = \int \frac{d^4p}{(2\pi)^4} \left[\left(-p^0 + \frac{p^2}{2m} + V_0 \right) \mathbb{1} + \mu_B \vec{\sigma} \cdot \vec{B} \right]^{-1} e^{ip(x-y)}. \quad (8.57)$$

In the following we will set $\vec{B} = 0$ for simplicity such that

$$\Upsilon(x-y) = \mathbb{1} \int_p \frac{1}{-p^0 + \frac{p^2}{2m} + V_0 - i\epsilon} e^{ip(x-y)}. \quad (8.58)$$

The term $i\epsilon$ makes sure that we take the right Greens function with time ordering. For a non-relativistic theory at zero temperature and density, this equals the retarded Greens function.

Yukawa theory. Let us now investigate a theory for a non-relativistic fermion with spin 1/2 and a real, relativistic scalar boson

$$S = \int dt d^3x \left\{ -\bar{\psi} \left(-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 - i\epsilon \right) \psi - \frac{1}{2} \phi \left(\partial_t^2 - \vec{\nabla}^2 + M^2 - i\epsilon \right) \phi - g \phi \bar{\psi} \psi \right\}. \quad (8.59)$$

We will discuss this theory in terms of the partition function

$$Z[\bar{\eta}, \eta, J] = \int D\bar{\psi} D\psi D\phi e^{iS[\bar{\psi}, \psi, \phi] + i \int_x \{ \bar{\eta}\psi + \bar{\psi}\eta + J\phi \}}. \quad (8.60)$$

As usual, by taking functional derivatives with respect to the source fields, one can obtain various correlation functions. Our strategy will be to perform a perturbation expansion in the cubic term $\sim g$. Let us first concentrate on the quadratic theory and the corresponding partition function derived from the action

$$S_2 = \int dt d^3x \left\{ -\bar{\psi} \left(-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 - i\epsilon \right) \psi - \frac{1}{2} \phi \left(\partial_t^2 - \vec{\nabla}^2 + M^2 - i\epsilon \right) \phi \right\}. \quad (8.61)$$

By doing the Gaussian integration one finds

$$\begin{aligned} Z_2[\bar{\eta}, \eta, J] &= \int D\bar{\psi} D\psi D\phi e^{iS_2 + i \int_x \{\bar{\eta}\psi + \bar{\psi}\eta + J\phi\}} \\ &= \exp \left[i \int d^4x d^4y \left\{ \bar{\eta}(x) \Upsilon(x-y) \eta(y) + \frac{1}{2} J(x) \Delta(x-y) J(y) \right\} \right] \end{aligned} \quad (8.62)$$

where $\Upsilon(x-y)$ is the Greens function for fermions in eq. (8.58). For the scalar bosons, the Green function is

$$\Delta(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{1}{-(p^0)^2 + \vec{p}^2 + M^2 - i\epsilon} e^{ip(x-y)}. \quad (8.63)$$

Again, the $i\epsilon$ term makes sure that the Greens function corresponds to the time-ordered or Feynman boundary conditions. One can also obtain this from a careful consideration of analytic continuation from Euclidean space to real time /Minkowski space. Note that the $i\epsilon$ term has in the functional integral the form

$$e^{iS} = e^{[i\dots + i\epsilon \int_x \phi^2(x)]} = e^{-\epsilon \int_x \phi(x)^2 + i\dots}. \quad (8.64)$$

This is the same suppression term that also appears in the Euclidean functional integral. It makes sure that functional integrals are converging and that the theory approaches the ground state on long time scales. In the complex plane, positions of poles are shifted slightly away from the real axis. This is illustrated in the left panel of the following figure. In fact this is equivalent to keeping the poles at $p^0 = \pm\sqrt{\vec{p}^2 + M^2}$ but moving slightly in the integration contour. This is illustrated in the right panel of the following figure.

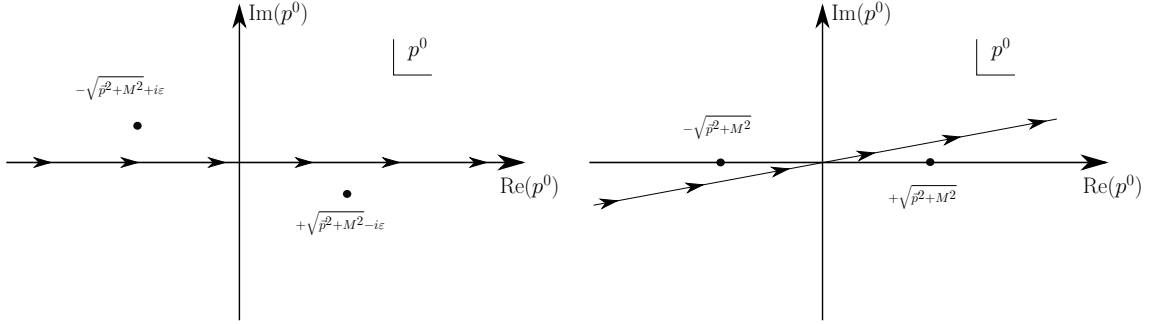


Figure 4. Illustration of the contour integral for the time-ordered Feynman propagator. In the left panel the poles are shifted slightly into the complex plane, in the right panel the integration contour is slightly shifted. Both prescriptions lead to equivalent results.

Let us use either of these prescriptions to calculate the scalar field propagator in position space

$$\Delta(x-y) = \int \frac{dp^0}{2\pi} \frac{d^3p}{(2\pi)^3} \frac{e^{-ip^0(x^0-y^0) + i\vec{p}(\vec{x}-\vec{y})}}{(-p^0 + \sqrt{\vec{p}^2 + M^2} - i\epsilon)(p^0 + \sqrt{\vec{p}^2 + M^2} - i\epsilon)} \quad (8.65)$$

The strategy will be to close the integration contour at $|p^0| \rightarrow \infty$ and to use the residue theorem. First, for $x^0 - y^0 > 0$, we can close the contour in the lower half of the complex p^0 -plane because $e^{-ip^0(x^0-y^0)} \rightarrow 0$ there. There is then only the residue at $p^0 = \sqrt{\vec{p}^2 + M^2}$ inside the integration contour (the $i\epsilon$ has already been dropped there). The residue theorem gives for the p^0 integral

$$\Delta(x-y) = \int \frac{d^3p}{(2\pi)^3} \frac{i}{2\sqrt{\vec{p}^2 + M^2}} e^{-i\sqrt{\vec{p}^2 + M^2}(x^0-y^0)} e^{i\vec{p}\vec{x}} \quad (\text{for } x^0 - y^0 > 0). \quad (8.66)$$

In contrast, for $x^0 - y^0 < 0$ we need to close the p^0 -integral in the upper half of the complex p^0 plane. The residue theorem given then

$$\Delta(x - y) = \int \frac{d^3 p}{(2\pi)^3} \frac{i}{2\sqrt{\vec{p}^2 + M^2}} e^{i\sqrt{\vec{p}^2 + M^2}(x^0 - y^0)} e^{i\vec{p}\vec{x}} \quad (\text{for } x^0 - y^0 < 0). \quad (8.67)$$

These results can be combined to

$$\begin{aligned} \Delta(x - y) &= \int \frac{d^3 p}{(2\pi)^3} \frac{i}{2\sqrt{\vec{p}^2 + M^2}} e^{-i\sqrt{\vec{p}^2 + M^2}|x^0 - y^0| + i\vec{p}\vec{x}} \\ &= i\theta(x^0 - y^0) \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2\sqrt{\vec{p}^2 + M^2}} e^{-i\sqrt{\vec{p}^2 + M^2}(x^0 - y^0) + i\vec{p}\vec{x}} \\ &\quad + i\theta(y^0 - x^0) \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2\sqrt{\vec{p}^2 + M^2}} e^{i\sqrt{\vec{p}^2 + M^2}(x^0 - y^0) + i\vec{p}\vec{x}} \end{aligned} \quad (8.68)$$

One can understand the first term as being due to particle-type excitations, while the second is due to anti-particle-type excitations. The above Greens function is known as time ordered or Feynmann propagator. For the non-relativistic fermion, the propagator integral over p^0 has just a single pole at $p^0 = \frac{\vec{p}^2}{2m} + V_0 - i\epsilon$,

$$\Upsilon(x - y) = \mathbb{1} \int \frac{dp^0}{2\pi} \frac{d^3 p}{(2\pi)^3} \frac{1}{-p^0 + \frac{\vec{p}^2}{2m} + V_0 - i\epsilon} e^{-ip^0(x^0 - y^0) + i\vec{p}\vec{x}}. \quad (8.69)$$

When $x^0 - y^0 > 0$ the contour can be closed below the real p^0 -axis, leading to

$$\Upsilon(x - y) = i \mathbb{1} \int \frac{d^3 p}{(2\pi)^3} e^{-i\left(\frac{\vec{p}^2}{2m} + V_0\right)(x^0 - y^0) + i\vec{p}\vec{x}} \quad (x^0 - y^0 > 0). \quad (8.70)$$

In contrast, for $x^0 - y^0 < 0$, the contour can be closed above and there is no contribution at all. In summary

$$\Upsilon(x - y) = i\theta(x^0 - y^0) \mathbb{1} \int \frac{d^3 p}{(2\pi)^3} e^{-i\left(\frac{\vec{p}^2}{2m} + V_0\right)(x^0 - y^0) + i\vec{p}\vec{x}}. \quad (8.71)$$

As a consequence of the absence of anti-particle-type excitations, the time-ordered and retarded propagators agree here.

Let us also note the relation between propagators and correlation functions. For the free (quadratic) theory one has

$$\begin{aligned} \langle \psi_a(x) \bar{\psi}_b(y) \rangle &= \left(\frac{1}{Z_2} \frac{\delta}{\delta \bar{\eta}_a(x)} \frac{\delta}{\delta \eta_b(y)} Z_2[\bar{\eta}, \eta, J] \right)_{\bar{\eta}=\eta=J=0} \\ &= -i\Upsilon_{ab}(x - y), \\ \langle \phi(x) \phi(y) \rangle &= \left(\frac{1}{Z_2} \frac{\delta}{\delta J(x)} \frac{\delta}{\delta J(y)} Z_2[\bar{\eta}, \eta, J] \right)_{\bar{\eta}=\eta=J=0} \\ &= -i\Delta(x - y). \end{aligned} \quad (8.72)$$

Note that some care is needed with interchanges of Grassmann variables to obtain the first expression. In a similar way one finds for the free theory

$$\begin{aligned} \langle \phi(x_1) \dots \phi(x_n) \rangle &= \left(\frac{1}{Z_2} \left(-i \frac{\delta}{\delta J(x_1)} \right) \dots \left(-i \frac{\delta}{\delta J(x_n)} \right) Z_2[\bar{\eta}, \eta, J] \right)_{\bar{\eta}=\eta=J=0} \\ &= \sum_{\text{pairings}} [-i\Delta(x_{j_1} - x_{j_2})] \dots [-i\Delta(x_{j_{n-1}} - x_{j_n})]. \end{aligned} \quad (8.73)$$

The sum in the last line goes over all possible ways to distribute x_1, \dots, x_n into pairs $(x_{j_1}, x_{j_2}), (x_{j_3}, x_{j_4}), \dots, (x_{j_{n-1}}, x_{j_n})$. This result is known as *Wick's theorem*. It follows directly from the

combinatorics of functional derivatives acting on Z_2 . For example,

$$\begin{aligned} \langle \phi(x_1) \phi(x_2) \phi(x_3) \phi(x_4) \rangle &= [-i\Delta(x_1 - x_2)][-i\Delta(x_3 - x_4)] \\ &\quad + [-i\Delta(x_1 - x_3)][-i\Delta(x_2 - x_4)] \\ &\quad + [-i\Delta(x_1 - x_4)][-i\Delta(x_2 - x_3)]. \end{aligned} \quad (8.74)$$

In a similar way correlation functions involving $\bar{\psi}$ and ψ can be written as sums over the possible ways to pair ψ and $\bar{\psi}$. For example

$$\begin{aligned} \langle \psi_{a_1}(x_1) \psi_{a_2} \bar{\psi}_{a_3}(x_3) \bar{\psi}_{a_4}(x_4) \rangle &= - \langle \psi_{a_1}(x_1) \bar{\psi}_{a_3}(x_3) \rangle \langle \psi_{a_2}(x_2) \bar{\psi}_{a_4}(x_4) \rangle \\ &\quad + \langle \psi_{a_1}(x_1) \bar{\psi}_{a_4}(x_4) \rangle \langle \psi_{a_2}(x_2) \bar{\psi}_{a_3}(x_3) \rangle \\ &= - [-i\Upsilon_{a_1 a_3}(x_1 - x_3)][-i\Upsilon_{a_2 a_4}(x_2 - x_4)] \\ &\quad + [-i\Upsilon_{a_1 a_4}(x_1 - x_4)][-i\Upsilon_{a_2 a_3}(x_2 - x_3)]. \end{aligned} \quad (8.75)$$

Note that correlation functions at quadratic level (for the free theory) need to involve as many fields ψ as $\bar{\psi}$, otherwise they vanish. Similarly, ϕ must appear an even number of times. For mixed correlation functions one can easily separate ϕ and $\psi, \bar{\psi}$ at quadratic level, because $Z_2[\bar{\eta}, \eta, J]$ factorizes. For example,

$$\langle \phi(x_1) \psi_a(x_2) \phi(x_3) \bar{\psi}_b(x_4) \rangle = [-i\Delta(x_1 - x_3)][-i\Upsilon_{ab}(x_2 - x_4)]. \quad (8.76)$$

It is useful to introduce also a graphical representation. We will represent the scalar propagator by a dashed line

$$-i\Delta(x - y) = x \text{-----} y \quad (8.77)$$

The Feynman propagator for the fermions will be represented by a solid line with arrow,

$$-i\Upsilon_{ab}(x - y) = (x, a) \text{-----} (y, b) \quad (8.78)$$

We can represent correlation functions graphically, for example, the mixed correlation function in eqn. (8.76) would be

$$\langle \phi(x_1) \psi_a(x_2) \phi(x_3) \bar{\psi}_b(x_4) \rangle = \left\{ \begin{array}{c} x_1 \text{-----} x_3 \\ (x_2, a) \text{-----} (x_4, b) \end{array} \right\} \quad (8.79)$$

Let us now also consider the interaction terms in the action. In the functional integral it contributes according to

$$e^{iS[\bar{\psi}, \psi, \phi]} = e^{iS_2[\bar{\psi}, \psi, \phi]} \exp \left[-ig \int d^4x \phi(x) \bar{\psi}_a(x) \psi_a(x) \right]. \quad (8.80)$$

Perturbation theory in g . We can assume that g is small and simply expand the exponential where it appears. This will add field factors $\sim \phi(x) \bar{\psi}_a(x) \psi_a(x)$ to correlation functions with an integral over x and an implicit sum over a . The resulting expression involving correlation functions can then be evaluated as in the free theory. For example,

$$\begin{aligned} \langle \phi(x_1) \psi_b(x_2) \bar{\psi}_c(x_3) \rangle &= \langle \phi(x_1) \psi_b(x_2) \bar{\psi}_c(x_3) \rangle_0 \\ &\quad + \left\langle \phi(x_1) \psi_b(x_2) \bar{\psi}_c(x_3) \left[-ig \int_y \phi(y) \bar{\psi}_a(y) \psi_a(y) \right] \right\rangle_0 + \dots \end{aligned} \quad (8.81)$$

The index 0 indicates that the correlation functions get evaluated in the free theory. Graphically, we can represent the interaction term as a vertex

$$-ig \int_y \sum_a = \begin{array}{c} \text{-----} \\ | \\ \text{-----} \\ (y, a) \end{array} \quad (8.82)$$

For each such vertex we need to include a factor $-ig$ as well as an integral over the space-time variable y and the spinor index a . To order g , we find for the example above

$$\begin{aligned}
\langle \phi(x_1)\psi_b(x_2)\bar{\psi}_c(x_3) \rangle = & \text{Diagram 1} + \text{Diagram 2} \\
& \text{Diagram 1: } (x_2, b) \xleftarrow{(y, a)} \overset{x_1}{\vdots} \xleftarrow{(y, a)} (x_3, c) \\
& \text{Diagram 2: } (x_2, b) \xleftarrow{(y, a)} \overset{x_1}{\vdots} \text{ (loop)} \xleftarrow{(y, a)} (x_3, c) \\
& \text{Diagram 3: } (x_2, b) \xleftarrow{(y, a)} \overset{x_1}{\vdots} \text{ (loop)} \xrightarrow{(y, a)} (x_3, c)
\end{aligned} \tag{8.83}$$

$$\begin{aligned}
&= -ig \int_y [-i\Delta(x_1 - y)][-i\Upsilon_{ba}(x_2 - y)][-i\Upsilon_{ac}(y - x_3)] \\
&+ ig \int_y [-i\Delta(x_1 - y)][-i\Upsilon_{bc}(x_2 - x_3)][-i\Upsilon_{aa}(y - y)].
\end{aligned}$$

The sign in the last line is due to an interchange of Grassmann fields. The last expression involves the fermion propagator for vanishing argument

$$\Upsilon_{ab}(0) = \delta_{ab} \int \frac{d^4p}{(2\pi)^4} \frac{1}{-p^0 + \frac{\vec{p}^2}{2m} + V_0 - i\epsilon} = i\theta(0)\delta_{ab}\delta^{(3)}(0). \tag{8.84}$$

We will set here $\theta(0) = 0$ so that the corresponding contribution vanishes. In other words, we will interpret

$$\Upsilon_{ab}(0) = \lim_{\Delta t \rightarrow 0} \Upsilon_{ab}(-\Delta t, \vec{0}) = 0. \tag{8.85}$$

Although this is a little ambiguous at this point, it turns out that this is the right way to proceed.

Feynmann rules in position space. To calculate a field correlation function in position space we need to

- have a scalar line ending on x for a factor $\phi(x)$, $x \text{-----}$
- have a fermion line ending on x for a factor $\psi_a(x)$, $(x, a) \xleftarrow{\quad}$
- have a fermion line starting on x for a factor $\bar{\psi}_a(x)$, $(x, a) \xrightarrow{\quad}$
- include a vertex $-ig \int_y$ for every order g , $\text{Diagram 1} + \text{Diagram 2}$ with integral over y .
- connect lines with propagators $[-i\Delta(x - y)]$ or $[-i\Upsilon_{ab}(x - y)]$
- determine the overall sign for interchanges of fermionic fields.

S-matrix elements. To calculate S-matrix elements from correlation functions, we need to use the LSZ formula. For an outgoing fermion, we need to apply the operator

$$i \left[-i\partial_t - \frac{\vec{\nabla}^2}{2m} + V_0 \right] \langle \cdots \psi_a(x) \cdots \rangle \tag{8.86}$$

and also go to momentum space by a Fourier transform

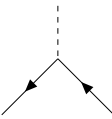
$$\int_x e^{+i\omega_p x^0 - i\vec{p}\vec{x}}. \tag{8.87}$$

The operator simply removes the propagator leading to x , because of

$$i \left[-i\partial_{x^0} - \frac{\vec{\nabla}_x^2}{2m} + V_0 \right] [-i\Upsilon_{ab}(x-y)] = \delta_{ab} \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \frac{-p^0 + \frac{\vec{p}^2}{2m} + V_0}{-p^0 + \frac{\vec{p}^2}{2m} + V_0} = \delta_{ab} \delta^{(4)}(x-y). \quad (8.88)$$

Moreover, all expressions are brought back to momentum space. One can formulate Feynmann rules directly for contributions to $i\mathcal{T}$ as follows.

- Incoming fermions are represented by an incoming line $\longleftarrow \bar{p}$ (to be read from right to left) associated with a momentum \vec{p} and energy $\omega_{\vec{p}} = \frac{\vec{p}^2}{2m} + V_0$.
- Outgoing fermions are represented by an outgoing line $\bar{p} \longleftarrow$
- Incoming or outgoing bosons are represented by $\cdots \leftarrow \bar{p}$ and $\bar{p} \leftarrow \cdots$ respectively.

- Vertices  contribute a factor $-ig$

- Internal lines that connect two vertices are represented by Feynmann propagators in momentum space, e.g.

$$\longleftarrow \xrightarrow{(p^0, \vec{p})} \longleftarrow = \frac{-i\delta_{ab}}{-p^0 + \frac{\vec{p}^2}{2m} + V_0}, \quad \cdots \xrightarrow{(p^0, \vec{p})} \cdots = \frac{-i}{-(p^0)^2 + \vec{p}^2 + M^2} \quad (8.89)$$

- Energy and momentum conservation are imposed on each vertex.
- For tree diagrams, all momenta are fixed by energy- and momenta conservation. For loop diagrams one must include an integral over the loop momentum l_j with measure $\frac{d^4l_j}{(2\pi)^4}$.
- Some care is needed to fix overall signs for fermions.
- Some care is needed to fix overall combinatoric factors from possible interchanges of lines / functional derivatives.

For the last two points it is often useful to go back to the algebraic expressions or to have some experience. We will later discuss very useful techniques based on generating functionals.

Fermion-fermion scattering We will now discuss an example, the scattering of (spin polarized) fermions of each other. The tree-level diagram is

$$\begin{array}{ccc} (\vec{q}_1, \uparrow) & & (\vec{q}_2, \downarrow) \\ & \searrow & \nearrow \\ & \text{---} & \\ & \nearrow & \searrow \\ (\vec{p}_1, \uparrow) & & (\vec{p}_2, \downarrow) \end{array} \quad (8.90)$$

Because the interaction with the scalar field does not change the spin, the outgoing fermion with momentum \vec{q}_1 will have spin \uparrow , the one with momentum \vec{q}_2 will have spin \downarrow . By momentum conservation the scalar line carries the four momentum

$$(\omega_{\vec{p}_1} - \omega_{\vec{q}_1}, \vec{p}_1 - \vec{q}_1) = \left(\frac{\vec{p}_1^2}{2m} - \frac{\vec{q}_1^2}{2m}, \vec{p}_1 - \vec{q}_1 \right) = (\omega_{\vec{q}_2} - \omega_{\vec{p}_2}, \vec{q}_2 - \vec{p}_2) \quad (8.91)$$

The last equality follows from overall momentum conservation, $p_1 + p_2 = q_1 + q_2$. The Feynmann rules give

$$i\mathcal{T} = (-ig)^2 \frac{-i}{-(\omega_{\vec{p}_1} - \omega_{\vec{q}_1})^2 + (\vec{p}_1 - \vec{q}_1)^2 + M^2}. \quad (8.92)$$

In the center-of-mass frame, one has $\omega_{\vec{p}_1} = \omega_{\vec{p}_2} = \omega_{\vec{q}_1} = \omega_{\vec{q}_2}$ and thus

$$\mathcal{T} = \frac{g^2}{(\vec{p}_1 - \vec{q}_1)^2 + M^2}. \quad (8.93)$$

Note that for $g^2 \rightarrow \infty$, $M^2 \rightarrow \infty$ with g^2/M^2 finite, \mathcal{T} becomes independent of momenta. This resembles closely the $\lambda(\phi^*\phi)^2$ interaction we discussed earlier for bosons.

More, generally, one can write

$$(\vec{p}_1 - \vec{q}_1)^2 = 2|\vec{p}_1|^2(1 - \cos(\vartheta)) = 4|\vec{p}_1|^2 \sin^2(\vartheta/2) \quad (8.94)$$

where we used $|\vec{p}_1| = |\vec{q}_1|$ in the center of mass frame and ϑ is the angle between \vec{p}_1 and \vec{q}_1 (incoming and outgoing momentum of the spin \uparrow particle). For the differential cross-section

$$\frac{d\sigma}{d\Omega_{q_1}} = \frac{|\mathcal{T}|^2 m^2}{16(\pi)^2}, \quad (8.95)$$

we find

$$\frac{d\sigma}{d\Omega_{q_1}} = \frac{g^4 m^2}{16\pi^2} \left[\frac{1}{4\vec{p}_1^2 \sin^2(\vartheta/2) + M^2} \right]^2. \quad (8.96)$$

Another interesting limit is $M^2 \rightarrow 0$. One has then

$$\frac{d\sigma}{d\Omega_{q_1}} = \frac{g^4 m^2}{64\pi^2 \vec{p}_1^2} \frac{1}{\sin^4(\vartheta/2)}. \quad (8.97)$$

This is the differential cross-section form found experimentally by Rutherford. It results from the exchange of a massless particle or force carrier which is here the scalar boson ϕ and in the case of Rutherford experiment (scattering of α -particles on Gold nuclei) it is the photon. This cross section has a strong peak at forward scattering $\vartheta \rightarrow 0$, and for $\vec{p}^2 \rightarrow 0$. These are known as colinear and soft singularities. Note that they are regulated by a small, nonvanishing mass $M > 0$.

9 Lorentz symmetry and the Dirac equation

Symmetries are basic concepts for the construction of a model. Particle physics in flat Minkowski space is invariant under Lorentz transformations. Even though the cosmological solutions are not Lorentz invariant, Lorentz invariance holds to a very good approximation on length and time scales that are small compared to the “size” (inverse Hubble parameter) of the universe. The functional integral formulation makes the implementation of symmetries easy. One imposes that the action S is invariant under the symmetry transformations. This is sufficient if the functional measure is also invariant. All symmetry properties follow the invariance of S and the functional measure.

9.1 Lorentz transformations and invariant tensors

Lorentz metric. The cartesian coordinates of space and time are t and \mathbf{x} . They are denoted as the contravariant vector

$$x^\mu = (t, \mathbf{x}), \quad t = x^0. \quad (9.1)$$

The corresponding covariant vector is

$$x_\mu = (-t, \mathbf{x}) = (-x^0, \mathbf{x}). \quad (9.2)$$

We can always lower and raise indices with the metric tensor $\eta_{\mu\nu}$ and its inverse $\eta^{\mu\nu}$, which are here actually the same,

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix}. \quad (9.3)$$

If we want to use a shorter notation, we can also say that our metric has the signature $(-, +, +, +)$. The explicit transformation equations are

$$x_\mu = \eta_{\mu\nu} x^\nu \quad \text{and} \quad x^\mu = \eta^{\mu\nu} x_\nu. \quad (9.4)$$

We want to know under which transformations $x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu$ the quantity $x^\mu x_\mu$ is invariant. So we calculate

$$x'^\mu x'_\mu = x'^\mu x'^\nu \eta_{\mu\nu} = \Lambda^\mu{}_\rho x^\rho \Lambda^\nu{}_\sigma x^\sigma \eta_{\mu\nu}. \quad (9.5)$$

This is equal to $x^\mu x_\mu$ if the condition

$$\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \eta_{\mu\nu} = \eta_{\rho\sigma} \quad (9.6)$$

is fulfilled. Equation (9.6) is the defining equation for Λ . All transformations that fulfill (9.6) are called *Lorentz transformations*. So-called *proper, orthochronous Lorentz transformations* that can be obtained as a sequence of infinitesimal transformations. Particle physics is invariant under the proper orthochronous Lorentz transformations for which we often use the shorthand “Lorentz transformations”. The general transformations (9.6), which we often call “extended Lorentz transformations”, comprise discrete transformations like parity and time reversal. Particle physics is not invariant under those discrete transformations.

Transformation of tensors. Let us consider the contravariant and covariant four-momenta

$$p^\mu = (E, \mathbf{p}) \quad (9.7)$$

$$p_\mu = (-E, \mathbf{p}) \quad (9.8)$$

As we already discussed, we can raise and lower indices of vectors with the metric tensor $\eta_{\mu\nu}$ and the inverse $\eta^{\mu\nu}$. We can interpret $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$ as matrices. As raising and lowering are inverse operations, the multiplication of both matrices is the identity,

$$\eta^{\mu\nu} \eta_{\nu\rho} = \delta^\mu{}_\rho. \quad (9.9)$$

If we perform a Lorentz transformation

$$p'^\mu = \Lambda^\mu{}_\nu p^\nu, \quad (9.10)$$

and lower indices on both sides, we get

$$\eta^{\mu\rho} p'_\rho = \Lambda^\mu{}_\nu \eta^{\nu\sigma} p_\sigma. \quad (9.11)$$

We multiply with an inverse metric

$$p'_\kappa = \eta_{\kappa\mu} \Lambda^\mu{}_\nu \eta^{\nu\sigma} p_\sigma \quad (9.12)$$

Obviously, the tensor product on the right hand side should be

$$\Lambda_\kappa{}^\nu = \eta_{\kappa\mu} \Lambda^\mu{}_\nu \eta^{\nu\sigma}. \quad (9.13)$$

The covariant vector (lower index) transforms as

$$p'_\mu = \Lambda_\mu^\nu p_\nu. \quad (9.14)$$

We use the metric tensor to raise and lower indices of tensors as well as for Λ_μ^ν . An example for the Lorentz transformation of a more complicated tensor is

$$A'^{\mu\nu\rho}_{\sigma\tau} = \Lambda^\mu_{\mu'} \Lambda^\nu_{\nu'} \Lambda^\rho_{\rho'} \Lambda_{\sigma'}^{\sigma} \Lambda_{\tau'}^{\tau} A^{\mu'\nu'\rho'}_{\sigma'\tau'} \quad (9.15)$$

The product of a covariant and a contravariant vector is a scalar: It is invariant under Lorentz transformations,

$$\begin{aligned} s &= a^\mu b_\mu, \\ \Rightarrow s' &= \Lambda^\mu_\rho a^\rho \Lambda_\mu^\sigma b_\sigma = a^\rho \underbrace{\Lambda^\mu_\rho \eta_{\mu\nu} \Lambda^\nu_\tau}_{\eta_{\rho\tau}} \eta^{\tau\sigma} b_\sigma = a^\rho b_\rho = s. \end{aligned} \quad (9.16)$$

Generalisation: Two contracted Lorentz indices do not contribute to transformations:

$$(A')_{\mu\rho} (B')^{\rho\nu} = \Lambda_\mu^\sigma \Lambda_\tau^\nu A_{\sigma\rho} B^{\rho\tau}, \quad (9.17)$$

e.g. $C_\mu^\nu = A_{\mu\rho} B^{\rho\nu}$ transforms as a 2-tensor.

Invariant tensors. We already mentioned that Lorentz transformations are defined in such a way that the metric tensor $\eta_{\mu\nu}$ is left invariant. Actually, there is only one more tensor that is invariant under Lorentz transformations, and this is the totally antisymmetric tensor $\epsilon_{\mu\nu\rho\sigma}$, the relativistic generalization of the Levi-Civita tensor ϵ_{ijkl} tensor. The Levi-Civita symbol with four indices $\epsilon_{\mu\nu\rho\sigma}$ is defined by total antisymmetry and

$$\epsilon_{0123} = 1. \quad (9.18)$$

It equals 1 for all cyclic permutations of (0, 1, 2, 3), and -1 for all anti-cyclic permutations. The ϵ -tensor with raised indices, $\epsilon^{\mu\nu\rho\sigma}$ has just the opposite signs, e. g. $\epsilon^{0123} = -1$.

Let us prove our statement that $\epsilon_{\mu\nu\rho\sigma}$ is invariant under Lorentz transformations. In the following lines we will use the short hand notation $\Lambda_\mu^\nu \rightarrow \Lambda$, $\Lambda^\nu_\sigma \rightarrow \Lambda^T$ and $\eta_{\mu\nu} \rightarrow \eta$. With this notation, the defining relation (9.13) reads

$$\Lambda \eta \Lambda^T = \eta, \quad (9.19)$$

If we compute the determinant on both sides, we find

$$\det(\Lambda) = \pm 1. \quad (9.20)$$

The determinant of Λ can also be calculated by

$$\det(\Lambda) = \frac{1}{4!} \Lambda_{\mu_1}^{\nu_1} \Lambda_{\mu_2}^{\nu_2} \Lambda_{\mu_3}^{\nu_3} \Lambda_{\mu_4}^{\nu_4} \epsilon_{\nu_1\nu_2\nu_3\nu_4} \epsilon^{\mu_1\mu_2\mu_3\mu_4} = \frac{1}{4!} \epsilon'_{\mu_1\mu_2\mu_3\mu_4} \epsilon^{\mu_1\mu_2\mu_3\mu_4} \quad (9.21)$$

Here ϵ' is the Lorentz transformed tensor. We can verify that $\epsilon'_{\mu\nu\rho\sigma}$ is totally antisymmetric, thus $\epsilon'_{\mu\nu\rho\sigma} = c \epsilon_{\mu\nu\rho\sigma}$ with constant c . Using $\epsilon_{\mu\nu\rho\sigma} \epsilon^{\mu\nu\rho\sigma} = 4!$ we obtain $\det(\Lambda) = c$ or

$$\epsilon'_{\mu_1\mu_2\mu_3\mu_4} = \det(\Lambda) \epsilon_{\mu\nu\rho\sigma} = \pm \epsilon_{\mu_1\mu_2\mu_3\mu_4} \quad (9.22)$$

Only Lorentz transformations with $\det(\Lambda) = +1$ will leave the ϵ -tensor invariant (the are called *proper*). The special Lorentz transformations obey $\det(\Lambda) = 1$ since they are continuously related to the unit transformation.

Analogy to Rotations. Equation (9.19) looks very similar to orthogonal transformations O_{mn} with

$$O\mathbb{1}O^T = OO^T = \mathbb{1}, \quad \mathbb{1}_{\mu\nu} = \delta_{\mu\nu}, \quad (9.23)$$

where $\mathbb{1}$ is the unit matrix. In (9.19) the ‘‘Euclidean metric’’ $\delta_{\mu\nu}$ is simply replaced by the metric tensor $\eta_{\mu\nu}$ for Minkowski space. In short,

- Orthogonal transformations : $\delta_{\mu\nu}$ invariant.
- Lorentz transformation: $\eta_{\mu\nu}$ invariant.
- Analytic continuation: $\delta_{\mu\nu} \rightarrow \eta_{\mu\nu}$.

The group of orthogonal transformations in three dimensions is denoted $O(3)$. The analogy that we just discussed motivates the name *Pseudo orthogonal transformations* $O(1,3)$ where the separated 1 indicates the special role of time in special relativity.

Derivatives. The derivative with respect to a contravariant vector is a covariant vector,

$$\partial_\mu = \frac{\partial}{\partial x^\mu}. \quad (9.24)$$

For example we have

$$\partial_\mu x^\mu = 4. \quad (9.25)$$

The momentum operator is

$$\hat{p}_\mu = -i\partial_\mu. \quad (9.26)$$

Four-dimensional Fourier transformation. The four-dimensional Fourier transformation of a function $\psi(x)$ is defined as

$$\psi(x) = \int_p e^{ip_\mu x^\mu} \psi(p). \quad (9.27)$$

With $p_\mu = (-\omega, \vec{p})$ and $p_\mu x^\mu = -\omega t + \vec{p}\vec{x}$ this reads

$$\psi(t, \vec{x}) = \int_\omega \int_{\vec{p}} e^{-i\omega t + i\vec{p}\vec{x}} \psi(\omega, \vec{p}). \quad (9.28)$$

Note that $p_\mu x^\mu$ is Lorentz invariant.

Covariant equations. For a covariant equation the left hand side and right hand side have the same transformation properties. An example is

$$\partial_\mu F^{\mu\nu} = J^\nu. \quad (9.29)$$

These are two of the four Maxwell equations.

9.2 Lorentz group

Group structure. If we have two elements g_1, g_2 that are elements of a group \mathcal{G} , the product of these two elements will still be an element of the group

$$g_3 = g_2 g_1 \in \mathcal{G}. \quad (9.30)$$

In particular, we can write for matrices

$$(\Lambda_3)^\mu{}_\nu = (\Lambda_2)^\mu{}_\rho (\Lambda_1)^\rho{}_\nu. \quad (9.31)$$

A group contains always a unit element e such that

$$ge = eg = g \quad (9.32)$$

for every group element g . For matrices, this unit element is $\delta^\mu{}_\nu$. Furthermore the inverse element g^{-1} exists. Every matrix $\Lambda^\mu{}_\nu$ has an inverse matrix because the determinant of Λ is ± 1 . Finally, for a group the multiplication law has to be associative, which is trivial for matrix multiplications.

Discrete symmetries. The Lorentz transformations contain some discrete symmetries that we discuss now.

Space reflection (parity). The space reflection transformation changes the sign of all space time coordinates, $x^j \rightarrow -x^j$ for $j \in \{1, 2, 3\}$ while time stays invariant $t \rightarrow t$. The corresponding matrix is

$$P = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (9.33)$$

The determinant is $\det(P) = -1$. The metric tensor $\eta_{\mu\nu}$ is kept invariant under a space reflection, $P\eta P^T = \eta$.

Time reflection. The time reflection transformation is $x^j \rightarrow x^j$ for $j \in \{1, 2, 3\}$ and $t \rightarrow -t$. The corresponding matrix is

$$T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix}. \quad (9.34)$$

The determinant of T is the same as for P , $\det(T) = \det(P) = -1$. Both transformations change the sign of the ϵ -tensor and are therefore improper Lorentz transformations. Again, the metric tensor is invariant under $T\eta T^T = \eta$.

Space-time reflection. The combination of both space and time reflection is

$$PT = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (9.35)$$

This time the determinant is $+1$.

Continuous Lorentz Transformations. A continuous Lorentz transformation can be obtained as a *product of infinitesimal* transformations. We use *Lorentz transformation* for the continuous Lorentz transformations. Since no jumps are possible, the continuous Lorentz transformations have a determinant $+1$, so we can immediately conclude that the discrete transformations P and T can't be described by continuous ones. As the product PT has a determinant $+1$, one could first think that this may be obtained by continuous transformations, but this is not the case. The reason is that infinitesimal transformations will never change the sign in front of time variable, but actually, PT does exactly this. However, a discrete transformation that can be obtained by infinitesimal ones is the reflection of x and y , so the product P_1P_2 with

$$P_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P_1P_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (9.36)$$

can be obtained as a continuous transformation, as familiar from rotations in two-dimensional space.

9.3 Generators and Lorentz Algebra

Infinitesimal Lorentz Transformations. Let us consider the difference δp^μ between a four-momentum and the transformed four-momentum,

$$\delta p^\mu = p'^\mu - p^\mu = (\Lambda^\mu{}_\nu - \delta^\mu{}_\nu)p^\nu = \delta\Lambda^\mu{}_\nu p^\nu, \quad (9.37)$$

with

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \delta\Lambda^\mu{}_\nu. \quad (9.38)$$

In a matrix representation, the infinitesimal Lorentz transformation is given by $\Lambda = \mathbb{1} + \delta\Lambda$. The defining relation of a Lorentz transformation ($\Lambda\eta\Lambda^T = \eta$) then leads to constraints for $\delta\Lambda$ as follows.

$$\begin{aligned} \Lambda\eta\Lambda^T &= \eta \\ \Leftrightarrow (1 + \delta\Lambda)\eta(1 + \delta\Lambda)^T &= \eta \\ \Leftrightarrow \delta\Lambda\eta + \eta\delta\Lambda^T &= 0. \end{aligned} \quad (9.39)$$

In this last line we neglected the 2nd order term in $\delta\Lambda$. If we write down this equation in the index notation of eq. (9.6), we have

$$\begin{aligned} \delta\Lambda^\mu{}_\rho\eta_{\mu\sigma} + \delta\Lambda^\nu{}_\sigma\eta_{\rho\nu} &= 0, \\ \text{or } \delta\Lambda_{\mu\nu} + \delta\Lambda_{\nu\mu} &= 0. \end{aligned} \quad (9.40)$$

This equation tells us that $\delta\Lambda_{\mu\nu}$ is antisymmetric, but note that $\delta\Lambda^\mu{}_\nu$ is not antisymmetric. The matrices have six independent elements, what is obvious for $\delta\Lambda_{\mu\nu} = -\delta\Lambda_{\nu\mu}$. The number of independent elements in a (antisymmetric) matrix is of course equal to the number of linear independent (antisymmetric) matrices we can build. The physical meaning of these six matrices is that they represent the possible *three infinitesimal rotations* and *three infinitesimal boosts*.

Generators. Let us write the infinitesimal transformation of the momentum vector in the following way,

$$\delta p^\mu = i\epsilon_z(T_z)^\mu{}_\nu p^\nu, \quad z = 1 \dots 6, \quad (9.41)$$

where a sum over z is implied. Any infinitesimal Lorentz transformation can be represented as a linear combination in this form

$$\delta\Lambda^\mu{}_\nu = i\epsilon_z(T_z)^\mu{}_\nu. \quad (9.42)$$

For the six independent generators we choose

$$\text{rotations : } (T_1)_{\mu\nu} = (T_1)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \quad (9.43)$$

$$(T_2)_{\mu\nu} = (T_2)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, \quad (T_3)_{\mu\nu} = (T_3)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (9.44)$$

$$\text{boosts : } (T_4)_{\mu\nu} = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (T_4)^\mu{}_\nu = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (9.45)$$

$$(T_5)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (T_6)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}. \quad (9.46)$$

Some remarks

- T_1 is a rotation around the x -axis (only y and z components change). Similarly T_2 is a rotation around the y -axis and T_3 a rotation around the z -axis.
- For the rotation matrices, raising and lowering of indices doesn't change anything. The reason is that the metric tensor has a -1 only in the zero component and the rotation matrices are zero in the first row.
- For the boost matrices, raising of the first index changes the sign of the first row of the matrix (see T_4). After raising the index, the boost matrices are not any longer antisymmetric. Explicitly,

$$(T_4)^\mu{}_\nu = \eta^{\mu\rho}(T_4)_{\rho\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (9.47)$$

- To see that T_1 , T_2 and T_3 are really rotations, compare them to the well-known rotation matrix in two dimensions,

$$R = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}. \quad (9.48)$$

If $\phi = \epsilon$ is infinitesimal, it becomes

$$R = \begin{pmatrix} 1 & -\epsilon \\ \epsilon & 1 \end{pmatrix}. \quad (9.49)$$

The difference to the identity is

$$\delta R = \begin{pmatrix} 0 & -\epsilon \\ \epsilon & 0 \end{pmatrix}, \quad (9.50)$$

But this is now equivalent to what we have in (9.6) when we write $i\epsilon$ in front of the matrix. The i in the definition of the generators is chosen such that T_1, T_2, T_3 are hermitian matrices.

- Similarly, you can convince yourself that T_4, T_5 and T_6 are boosts in x , y and z direction.

Lorentz algebra. The product of two group elements is again a group element. From this we can conclude that the commutator of two generators must again be a generator. In general we can therefore write

$$[T_x, T_y] = i f_{xyz} T_z, \quad (9.51)$$

where the sum over z is implied. The f_{xyz} are called the *structure constants* of a group. Whenever one has to deal with groups, the structure constants are very important, because once we know them, we know all algebraic relations for this group.

The central relation (9.7) can be shown as follows. Consider transformations

$$e^{iA}, e^{iB}, \quad A = \epsilon_z^{(A)} T_z, \quad B = \epsilon_y^{(B)} T_y \quad (9.52)$$

The combined transformation

$$e^{-iA} e^{-iB} e^{iA} e^{iB} = e^{iC} \quad (9.53)$$

is again an element of the group and therefore $C = \epsilon_w^{(C)} T_w$. Use

$$e^{iA} e^{iB} = e^{iB} e^{iA} + [B, A] + \dots \quad (9.54)$$

for showing in the combined transformation

$$\begin{aligned} \mathbb{1} + [B, A] &= \mathbb{1} + iC \\ [B, A] &= iC \\ -\epsilon_y^{(B)} \epsilon_z^{(A)} [T_y, T_z] &= i\epsilon_w^{(C)} T_w. \end{aligned} \quad (9.55)$$

It follows that the commutator $-i[T_z, T_y]$ is a linear combination of generators,

$$-i[T_z, T_y] = c_w^{(zy)} T_w. \quad (9.56)$$

The coefficients $c_w^{(zy)} = f_{zyw}$ can be identified with the structure constants.

Example Let us consider a rotation in three dimensional space. We want to rotate a system

- by an angle α around the y -axis,
- by an angle β around the x -axis,
- by an angle $-\alpha$ around the y -axis,
- and finally by an angle $-\beta$ around the x -axis.

The result of a product of infinitesimal rotations is again an infinitesimal rotation,

$$\begin{aligned} (1 - i\beta T_x - \frac{1}{2}\beta^2 T_x^2) (1 - i\alpha T_y - \frac{1}{2}\alpha^2 T_y^2) (1 + i\beta T_x - \frac{1}{2}\beta^2 T_x^2) (1 + i\alpha T_y - \frac{1}{2}\alpha^2 T_y^2) \\ = 1 - \alpha\beta(T_x T_y - T_y T_x) = 1 - i\alpha\beta T_z \end{aligned} \quad (9.57)$$

The first order is zero, and the terms $\propto T_x^2$ and $\propto T_y^2$ cancel, too. The product $\alpha\beta$ is the parameter of the resulting infinitesimal transformation.

For the special case of a rotation in three dimensional space, one can show the commutation relation

$$[T_1, T_2] = iT_3 \quad (9.58)$$

by multiplication of the matrices specified before. More generally, the generators of rotations obey

$$[T_k, T_l] = i\epsilon_{klm} T_m \quad \text{for } k, l, m \in \{1, 2, 3\}. \quad (9.59)$$

The calculation of this example gives us already some commutation relations of the generators of the Lorentz group, if we consider the T_i as 4 x 4 matrices with zeroes in all elements of the first column and row. This is of course not surprising, as the three dimensional rotations are a subgroup of the Lorentz group. The other structure constants f_{xyz} where one element of x, y, z is 0 can also be found from the specified matrices. We will give them later.

9.4 Representations of the Lorentz group

Remembering the spin matrices and their commutation relations, we discover that they are exactly the same as for the generators of the rotation group $SO(3)$:

$$\begin{aligned} s_k &= \frac{1}{2}\tau_k, \quad \text{where } \tau_k \text{ are Pauli matrices and,} \\ [T_k, T_l] &= 2i\epsilon_{klm} \tau_m, \quad [s_k, s_l] = i\epsilon_{klm} s_m, \end{aligned} \quad (9.60)$$

Note the difference between spin matrices s_i and Pauli matrices τ_i

The important thing we learn here is that the spin matrices $\tau_i/2$ and the generators of rotations in 3D space have the same algebraic relations. They correspond to *different representations of the rotation group*. The T_m are a three-dimensional and the $\tau_m/2$ are a two-dimensional representation of the group $SO(3)$.

Representations and Matrices Let us summarize what we know about the Lorentz group: It is $SO(1,3)$ and is generated by a set of 6 independent matrices T_z , which obey the commutation relations

$$[T_x, T_y] = i f_{xyz} T_z. \quad (9.61)$$

For $x, y, z \in 1, 2, 3$ we know already that $f_{xyz} = \epsilon_{xyz}$. The dimension of the matrices T_z depends on the representation of the group: If we have a d -dimensional representation, the matrices will be $d \times d$.

d-dimensional representation: The set of $d \times d$ -matrices T_z obey the commutation relations of a given group. In the physics literature, one often uses the representation to design (somewhat improperly) also a d -component object on which the matrices T_z act.

For a vector, the dimension is $d = 4$ because we have three space and one time coordinate. What happens if we want to transform a tensor? Consider the symmetric energy-momentum-tensor $T^{\mu\nu} = T^{\nu\mu}$. We know that it has 10 independent elements: 4 diagonal and 6 off-diagonal ones. Let us write all independent elements into a 10 dimensional vector ψ^α . The generator T_z that transforms this vector into a new vector $\psi^\alpha + \delta\psi^\alpha$ (in complete analogy to the momentum p^μ that transformed into $p^\mu + \delta p^\mu$) must now be a 10×10 matrix:

$$\delta\psi^\alpha = i \epsilon_z (T_z)^\alpha_\beta \psi^\beta \quad (9.62)$$

The elements of ψ are the elements of the energy-momentum tensor $T^{\mu\nu}$ and we therefore know the Lorentz transformations.

$$\delta T^{\mu\nu} = i \epsilon_z (T_z)^{\mu\nu}_{\mu'\nu'} T^{\mu'\nu'}. \quad (9.63)$$

Here $(\mu\nu) = (\nu\mu)$ is considered as a double index, $\alpha = (\mu\nu)$. In this equation, don't mix up the energy-momentum tensor and the generator! The elements of $(T_z)^{\mu\nu}_{\mu'\nu'}$ can easily be computed from the known Lorentz transformation of a tensor.

Irreducible Representations We can decompose T into the trace and the remaining traceless part \tilde{T} :

$$\tilde{T}^{\mu\nu} = T^{\mu\nu} - \frac{1}{4} \theta \eta^{\mu\nu}. \quad (9.64)$$

Here θ is the trace of the energy-momentum tensor and $\tilde{T}^{\mu\nu}$ is the traceless part. For the trace we can also write

$$\theta = T^\mu_\mu = \eta_{\nu\mu} T^{\mu\nu}. \quad (9.65)$$

The trace is a scalar and thus doesn't change under Lorentz transformations:

$$T'^\rho_\rho = T^\rho_\rho. \quad (9.66)$$

Furthermore, the traceless tensor \tilde{T} remains traceless when it is transformed. It has nine independent components. In this way, we have reduced the 10 representation to $9 + 1$. The transformation of traceless, symmetric tensors is represented by 9×9 matrices as generators.

As an intermediate result we can now summarize:

	Representation	Dimension
	scalar	1
	vector	4
	symmetric and traceless tensors	9
	antisymmetric tensors	6
	spinor	?

9.5 Transformation of Fields

Scalar Field $\varphi(x)$ How do scalar fields $\varphi(x)$ transform? To find an answer to this question, we recall the transformation of the space-time vector x^μ :

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu, \quad x'^\mu = x^\mu + \delta x^\mu \quad (9.67)$$

The value of the transformed field φ' at the transformed coordinate x' is the same as the field value before the transformation.

$$\varphi'(x') = \varphi(x). \quad (9.68)$$

From this we find the transformed field value φ' at the assigned coordinate

$$\varphi'(x) = \varphi(x - \delta x), \quad (9.69)$$

since $x - \delta x$ is transformed to x . We can visualise this by the following picture.

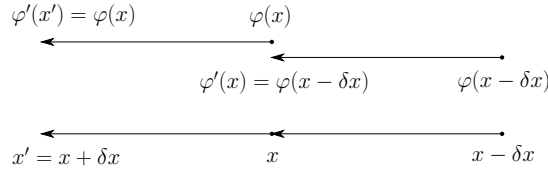


Figure 5. Transformation of a scalar field

We want to consider field transformations of fixed coordinates and therefore employ

$$\varphi'(x) = \varphi(x) + \delta\varphi(x) = \varphi(x - \delta x). \quad (9.70)$$

The transformation of φ at fixed x is called an **active transformation**. In contrast, leaving φ fixed and changing coordinates would be a **passive transformation**. (The combination of both does not change the field, $\varphi'(x') = \varphi(x)$.)

The difference of the field $\delta\varphi$ can be expressed as follows.

$$\begin{aligned} \delta\varphi &= \varphi(x - \delta x) - \varphi(x) \\ &= -\partial_\mu \varphi(x) \delta x^\mu. \end{aligned} \quad (9.71)$$

The second line comes from the definition of the derivative.

If we insert $\delta x^\mu = \delta\Lambda^\mu{}_\nu x^\nu$ we get

$$\begin{aligned} \delta\varphi &= x^\nu \delta\Lambda^\mu{}_\nu \partial_\mu \varphi(x) \\ &= -\delta\Lambda^\mu{}_\nu x^\nu \partial_\mu \varphi(x) \\ &= -i\epsilon_z (T_z)^\mu{}_\nu x^\nu \partial_\mu \varphi(x) \\ &= i\epsilon_z L_z \varphi(x). \end{aligned} \quad (9.72)$$

In the second to the last line we use (9.42) and in the last line we introduce the definition

$$L_z = -(T_z)^\mu{}_\nu x^\nu \partial_\mu. \quad (9.73)$$

For the second equation one lowers one index of Λ and uses the antisymmetry.

For fields the L_z are the generators and not T_z because (9.72) is of the form (9.41). The generators L_z contain a differential operator in this case! Fields are infinite dimensional representations in this sense.

The letter L was not chosen arbitrary, as L_1 , L_2 and L_3 are the angular momenta. For instance L_1 can be written as

$$L_1 = -(T_1)^\mu{}_\nu x^\nu \partial_\mu. \quad (9.74)$$

T_1 has only two non-zero elements: $(T_1)^\mu{}_\nu = -i$ and $(T_1)^\nu{}_\mu = i$, so

$$L_1 = -ix^2 \frac{\partial}{\partial x^3} + ix^3 \frac{\partial}{\partial x^2} = -i(y\partial_z - z\partial_y). \quad (9.75)$$

This is obviously the angular momentum as we know it from classical mechanics: $L = r \times p$. The transformation of fields with Lorentz indices has two ingredients. The first arises from the transformation of coordinates, the second is related to the Lorentz indices. For scalars one has only the coordinate part.

Vector Field $A^\mu(x)$ Contravariant vectors transform as :

$$A^\mu(x) \rightarrow A'^\mu(x) = A^\mu(x) + \delta A^\mu(x), \quad (9.76)$$

$$\delta A^\mu(x) = \delta \Lambda^\mu{}_\nu A^\nu(x) + x^\rho \delta \Lambda_\rho{}^\sigma \partial_\sigma A^\mu(x). \quad (9.77)$$

Here, $\delta \Lambda^\mu{}_\nu A^\nu$ is the usual transformation law for covariant vectors and $x^\rho \delta \Lambda_\rho{}^\sigma \partial_\sigma A^\mu$ reflects the change of the coordinates. This second term is always there, no matter what kind of field we are transforming.

Covariant vectors transform as:

$$\delta A_\mu(x) = \delta \Lambda_\mu{}^\nu A_\nu(x) + x^\rho \delta \Lambda_\rho{}^\sigma \partial_\sigma A_\mu(x). \quad (9.78)$$

The covariant derivative transforms as

$$\begin{aligned} \partial_\mu \varphi(x) &\rightarrow (\partial_\mu \varphi)'(x) = \partial_\mu(\varphi(x) + \delta \varphi(x)) = \partial_\mu \varphi(x) + \delta \partial_\mu \varphi(x), \\ \delta \partial_\mu \varphi(x) &= \partial_\mu \delta \varphi(x) = \partial_\mu(x^\rho \delta \Lambda_\rho{}^\sigma \partial_\sigma \varphi(x)) = \delta \Lambda_\mu{}^\sigma \partial_\sigma \varphi(x) + (x^\rho \delta \Lambda_\rho{}^\sigma \partial_\sigma)(\partial_\mu \varphi(x)), \end{aligned} \quad (9.79)$$

So, $\partial_\mu \varphi$ transforms as a covariant vector. With a similar argument one finds that the contravariant derivative transforms as a contravariant vector. This implies

$$\delta(\partial^\mu \varphi(x) \partial_\mu \varphi(x)) = (x^\rho \delta \Lambda_\rho{}^\sigma \partial_\sigma)(\partial^\mu \varphi(x) \partial_\mu \varphi(x)), \quad (9.80)$$

i.e. $\partial^\mu \varphi(x) \partial_\mu \varphi(x)$ transforms as a scalar.

Invariant Action This is a central piece, but with all the machinery we have developed it is almost trivial. Now our works pays off.

Let $f(x)$ be some (composite) scalar function

$$\delta f = x^\rho \delta \Lambda_\rho{}^\sigma \partial_\sigma f, \quad (9.81)$$

examples are $f = \varphi^2$ or $f = V(\phi)$ or $f = \partial^\mu \varphi \partial_\mu \varphi$. It follows that

$$S = \int d^4x f(x) \text{ is invariant, i.e. } \delta S = 0. \quad (9.82)$$

Proof:

$$\begin{aligned} \delta S &= \int d^4x \delta f(x) \\ &= \int d^4x x^\rho \delta \Lambda_\rho{}^\sigma \partial_\sigma f \\ &= \int d^4x \partial_\sigma (x^\rho \delta \Lambda_\rho{}^\sigma f) - \int d^4x \delta \Lambda_\rho{}^\sigma \partial_\sigma f \\ &= 0 \end{aligned} \quad (9.83)$$

The first integral is zero because there are no boundary contributions. Total derivatives in \mathcal{L} will always be neglected, i.e. always $\int d^4x \partial_\mu A = 0$. The second integral is zero because of the antisymmetry of $\Lambda_{\rho\sigma}$:

$$\delta_\sigma^\rho \delta \Lambda_\rho^\sigma = \eta^{\rho\sigma} \delta \Lambda_{\rho\sigma} = 0. \quad (9.84)$$

Examples It is now very easy to construct quantum field theories! Simply write down actions S that are Lorentz invariant.

We will consider actions of the form

$$S = \int d^4x \sum_k \mathcal{L}_k(x), \quad (9.85)$$

where \mathcal{L}_k are (composite) scalar quantities.

Here are some examples:

$$\bullet \mathcal{L} = \partial^\mu \varphi^* \partial_\mu \varphi + m^2 \varphi^* \varphi \quad (9.86)$$

This is a **free charged scalar field**, it describes particles with mass m like e.g. pions π^\pm with interactions neglected.

$$\bullet \mathcal{L} = \frac{1}{4} F^{\mu\nu} F_{\mu\nu}, \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (9.87)$$

F is the **electromagnetic field**. This describes free photons.

$$\bullet \mathcal{L} = (\partial^\mu + ieA^\mu) \varphi^* (\partial_\mu - ieA_\mu) \phi + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \quad (9.88)$$

This describes a charged scalar field interacting with photons and is called **scalar QED**. (We need one more concept to do **QED**, we have to account for the spin of the electrons.)

9.6 Functional Integral, Correlation Functions

Measure

$$\int D\varphi(x) \quad \text{is invariant.} \quad (9.89)$$

To prove this, we use the equivalence of active and passive transformations,

$$\varphi'(x) = \varphi(\Lambda^{-1}x). \quad (9.90)$$

For vectors we have

$$\int DA^\mu = \int DA'^\mu \times \text{Jacobian} \quad (9.91)$$

But the Jacobian is $\det \Lambda = 1$.

Comment : Is it always possible to find an invariant measure? There is a possible conflict with regularization, i.e. with taking the continuum limit. E.g. lattice regularization is *not* Lorentz invariant.

The answer to that question is that in all experience physicists have so far, lattice QFTs do work. We assume in this lecture that $\int D\varphi$ is invariant under Lorentz transformations.

Partition Function $Z = \int D\phi e^{-S}$ is invariant if $\int D\varphi$ and S are invariant.

Correlation Function

$$\langle \varphi(x) | \varphi(x') \rangle = Z^{-1} \int D\phi \phi(x) \phi(x') e^{-S} \quad \text{transforms as } \varphi(x) \varphi(x'). \quad (9.92)$$

This is a *covariant construction*. This makes it easy to construct an invariant S-matrix. Thus e.g. scattering cross sections are Lorentz invariant.

Summary Explicit Lorentz covariance is an important advantage of the functional formulation! This is not so early implemented in the operator formalism! Recall that H is not invariant, it is a three-dimensional object. S is a four-dimensional object.

Spinor representations of the Lorentz group Electrons have half-integer spin. We first look at the rotation group $SO(3)$, which is a subgroup of the Lorentz group. For *nonrelativistic* electrons this subgroup is all that matters.

We look at a two-dimensional representation of the rotation group :

$$\chi = \begin{pmatrix} \chi_1(x) \\ \chi_2(x) \end{pmatrix} = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}. \quad (9.93)$$

The rotation subgroup $SO(3)$ is given by

$$\delta\chi = i\epsilon_z T_z \chi + \delta'\chi, \quad z = 1, 2, 3, \quad (9.94)$$

$$\delta'\chi(x) = x^\rho \delta\Lambda_\rho^\sigma \partial_\sigma \chi(x). \quad (9.95)$$

We will omit δ' in the notation from now on. This universal contribution is the same for all fields. It is implicitly added if we transform fields. The spinor representation of $SO(3)$ is two-dimensional. The three 2×2 matrices T_z are given by the Pauli matrices:

$$T_z = \frac{1}{2}\tau_z, \quad z = 1, 2, 3. \quad (9.96)$$

Comment The $\chi(x)$ are Grassmann variables. This is not relevant for symmetry transformations. Now we ask the question for relativistic electrons or neutrons:

- What are the spinor representations of the full Lorentz group, i.e. what are T_z for $z = 1, \dots, 6$?
- Are there two-dimensional representations, i.e. are there six 2×2 matrices that obey

$$[T_x, T_y] = i f_{xyz} T_z? \quad (9.97)$$

These questions belong to the mathematical field of **representation** theory. We do not attempt to find the representation ourselves. Dirac, Pauli and Weyl did that for us. We only give the representations and verify that they really are representations of the Lorentz group.

Dirac Spinors By Dirac spinors we mean the four-dimensional representation

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}, \quad (9.98)$$

with generators

$$i\epsilon_z T_z = \frac{i}{2}\epsilon_{\hat{\mu}\hat{\nu}} T^{\hat{\mu}\hat{\nu}} \quad (9.99)$$

The six generators $T^{\hat{\mu}\hat{\nu}}$ are now labeled by $\hat{\mu}\hat{\nu}$ instead of z . The factor $\frac{1}{2}$ accounts for $\frac{1}{2}(\epsilon_{12}T^{12} + \epsilon_{21}T^{21}) = \epsilon_{12}T^{12}$ etc. with

$$T^{\hat{\mu}\hat{\nu}} = -T^{\hat{\nu}\hat{\mu}}, \quad \epsilon_{\hat{\mu}\hat{\nu}} = -\epsilon_{\hat{\nu}\hat{\mu}}. \quad (9.100)$$

We put the hats on $\hat{\mu}$ and $\hat{\nu}$ to avoid confusion: the matrices $T^{\hat{\mu}\hat{\nu}}$ are fixed 4×4 matrices and Lorentz transformations do not act on them as they do on fields. Once again: e.g. T^{12} is itself a 4×4 matrix, $\hat{\mu}\hat{\nu} = 12$ is just a convenient label for this matrix, we could also have labelled it by

$z = 3$.

The matrices $T^{\hat{\mu}\hat{\nu}}$ are obtained as the commutators of the Dirac matrices γ^μ

$$T^{\hat{\mu}\hat{\nu}} = -\frac{i}{4} [\gamma^{\hat{\mu}}, \gamma^{\hat{\nu}}]. \quad (9.101)$$

The **Dirac matrices** $\gamma^{\hat{\mu}}$ are complex 4×4 matrices. There are four of them:

$$\gamma^k = \begin{pmatrix} 0 & -i\tau_k \\ i\tau_k & 0 \end{pmatrix}, \quad k = 1, 2, 3 \quad \text{and} \quad \gamma^0 = \begin{pmatrix} 0 & -i\mathbf{1} \\ -i\mathbf{1} & 0 \end{pmatrix}, \quad (9.102)$$

where τ_k , $k = 1, 2, 3$ are the Pauli matrices.

In the following, we often omit the hat for γ^μ , but remember that Lorentz transformations act only on fields, whereas the matrices γ^μ are fixed.

If you compute the T-matrices(exercise!), you will find that they are of the form

$$T^{\mu\nu} = \begin{pmatrix} T_+^{\mu\nu} & 0 \\ 0 & T_-^{\mu\nu} \end{pmatrix} \quad (9.103)$$

where the $T_\pm^{\mu\nu}$ are 2×2 matrices. The ij-components are rotations,

$$T_+^{ij} = T_-^{ij} = \frac{1}{2} \varepsilon^{ijk} \tau_k, \quad i, j, k \in \{1, 2, 3\}. \quad (9.104)$$

E.g. for a rotation around the z-axis ($\varepsilon_{12} = -\varepsilon_{21} \equiv \varepsilon_3$), we have

$$\begin{aligned} \varepsilon_3 T_3 &\equiv \frac{1}{2} (\varepsilon_{12} T^{12} + \varepsilon_{21} T^{21}), \\ &= \varepsilon_3 T^{12} = \varepsilon_3 \left(\frac{1}{2} \varepsilon^{123} \tau_3 \right) = \varepsilon_3 \frac{\tau_3}{2}, \end{aligned} \quad (9.105)$$

confirming $T_3 = \frac{\tau_3}{2}$. If we denote

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \psi_L, \quad \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix} = \psi_R, \quad \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \psi, \quad (9.106)$$

then ψ_L and ψ_R transform as 2-component spinors with respect to rotations.

The T^{0k} generators are boosts,

$$T_+^{0k} = -T_-^{0k} = -\frac{i}{2} \tau_k. \quad (9.107)$$

The boost generators are not hermitian.

The commutation relations can be computed as

$$[T^{\mu\nu}, T^{\rho\sigma}] = i (\eta^{\mu\rho} T^{\nu\sigma} - \eta^{\mu\sigma} T^{\nu\rho} + \eta^{\nu\sigma} T^{\mu\rho} - \eta^{\nu\rho} T^{\mu\sigma}) \quad (9.108)$$

These are indeed the commutation relations of the Lorentz group.

We can compare with the defining vector representation by the identification

$$\begin{aligned} T_1 &= T^{23}, & T_2 &= T^{31}, & T_3 &= T^{12}, \\ T_4 &= T^{01}, & T_5 &= T^{02}, & T_6 &= T^{03}. \end{aligned} \quad (9.109)$$

In the vector representation one has

$$(T^{\hat{\mu}\hat{\nu}})^\mu{}_\nu = -i (\eta^{\hat{\mu}\mu} \delta_\nu^{\hat{\nu}} - \eta^{\hat{\nu}\mu} \delta_\nu^{\hat{\mu}}), \quad (9.110)$$

e.g.

$$\begin{aligned} (T_1)^\mu{}_\nu &= (T^{23})^\mu{}_\nu = -i (\delta^{2\mu} \delta_\nu^3 - \delta^{3\mu} \delta_\nu^2) = \begin{pmatrix} 0, 0, 0, 0 \\ 0, 0, 0, 0 \\ 0, 0, 0, -i \\ 0, 0, i, 0 \end{pmatrix}, \\ (T_4)^\mu{}_\nu &= (T^{01})^\mu{}_\nu = -i (-\delta^{0\mu} \delta_\nu^1 - \delta^{1\mu} \delta_\nu^0). \end{aligned} \quad (9.111)$$

In this representation the commutation relation is easily established.

Weyl Spinors As we have seen before, the matrices $T^{\mu\nu}$ are block-diagonal, which means that they do not mix all components of a 4-spinor Ψ into each other, but only the first two and the last two. Mathematically speaking, there are two invariant subspaces, so the Dirac representation is called reducible. This is why we introduce now the Weyl representation, which will be a two-dimensional irreducible representation (irrep). We define

$$\Psi_L = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ 0 \\ 0 \end{pmatrix}, \quad \Psi_R = \begin{pmatrix} 0 \\ 0 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}. \quad (9.112)$$

From now on, we will suppress the 0's in the Weyl spinors and just write

$$\Psi_L = \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}, \quad \Psi_R = \begin{pmatrix} \Psi_3 \\ \Psi_4 \end{pmatrix}. \quad (9.113)$$

We will later use Weyl spinors to describe neutrinos. For electrons we will need Dirac Spinors. This is related to the fact that the parity transformation maps between Ψ_L and Ψ_R .

Parity Transformation The parity transformation is defined by

$$\Psi(x) \rightarrow \gamma^0 \Psi(Px), \quad Px = (x^0, -\vec{x}). \quad (9.114)$$

How do the individual Weyl spinors transform? We observe that

$$\gamma^0 \begin{pmatrix} \Psi_L \\ \Psi_R \end{pmatrix} = -i \begin{pmatrix} \Psi_R \\ \Psi_L \end{pmatrix}. \quad (9.115)$$

and therefore

$$(\Psi')_L = -i\Psi_R, \quad (\Psi')_R = -i\Psi_L, \quad (9.116)$$

Parity exchanges left and right components. This is indeed one of the reasons why we will need a left-handed and a right-handed Weyl spinor to describe electrons. Neutrinos are described only by a left-handed Weyl spinor, so obviously they violate parity!

Projection Matrix Now we introduce a matrix γ^5 , such that we can make a projection from the Dirac to Weyl representation by

$$\Psi_L = \frac{1}{2}(1 + \gamma^5)\Psi, \quad (9.117)$$

$$\Psi_R = \frac{1}{2}(1 - \gamma^5)\Psi. \quad (9.118)$$

This is obviously fulfilled by

$$\gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (9.119)$$

where the 1 represents a 2×2 -unit-matrix. One can check that

$$[\gamma^5, T^{\mu\nu}] = 0, \quad (\gamma^5)^2 = 1. \quad (9.120)$$

However, we want to treat the matrix γ^5 in a more general way and express it in terms of the other γ -matrices, so that we know it independently of the particular representation of the Dirac matrices. First we show that for the relations (9.120) to hold, it is sufficient that

$$\{\gamma^\mu, \gamma^5\} = 0. \quad (9.121)$$

Proof:

$$\gamma^5 T^{\mu\nu} = -\frac{i}{4}\gamma^5(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) = \frac{i}{4}(\gamma^\mu\gamma^5\gamma^\nu - \gamma^\nu\gamma^5\gamma^\mu) = -\frac{i}{4}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\gamma^5 = T^{\mu\nu}\gamma^5. \quad (9.122)$$

One can check as an exercise that this relation is indeed fulfilled when we define

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (9.123)$$

In our particular representation one has the properties

$$\frac{1 + \gamma^5}{2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \frac{1 - \gamma^5}{2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (9.124)$$

or

$$\gamma^5\Psi_L = \Psi_L, \quad \gamma^5\Psi_R = -\Psi_R. \quad (9.125)$$

Dirac Matrices Let's look in some more detail at the Dirac matrices we have used so far. Their defining property is given by

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}. \quad (9.126)$$

This is known as the Clifford algebra. From this relation one can derive all the commutator relations for the $T^{\mu\nu}$ and γ^5 ! For instance one can obviously see that $(\gamma^i)^2 = 1$, $i = 1, 2, 3$ and $(\gamma^0)^2 = -1$. This is quite useful, since different books will use different representations of the Clifford algebra (however, also take care for the signature of the metric in different books!). We can go from one representation to another using a similarity transformation

$$\gamma^\mu \rightarrow \gamma'^\mu = A\gamma^\mu A^{-1}. \quad (9.127)$$

We can easily check that such a transformation does not change the anticommutator relations:

$$\{\gamma'^\mu, \gamma'^\nu\} = A\{\gamma^\mu, \gamma^\nu\}A^{-1} = 2A\eta^{\mu\nu}A^{-1} = \eta^{\mu\nu}. \quad (9.128)$$

10 Quantum electrodynamics

10.1 Action and propagators

We are now ready to construct the action for quantum electrodynamics (QED). We have Grassmann variables for fermions and the spinor representation of the Lorentz group. We start with free electrons, and add the interactions with photons subsequently.

a) Invariant action for free electrons.

Kinetic term: We want to use the spinor representation discussed in the previous section to write down Lorentz invariant actions for fermions. In fact, it is possible to write down an action consisting only of a kinetic term with only one derivative:

$$S = \int d^4x \mathcal{L}, \quad \mathcal{L} = i\bar{\Psi}\gamma^\mu\partial_\mu\Psi = i\bar{\Psi}_\alpha(\gamma^\mu)_{\alpha\beta}\partial_\mu\Psi_\beta. \quad (10.1)$$

As usual, Ψ denotes a column vector and $\bar{\Psi}$ is a line vector,

$$\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}, \quad \bar{\Psi} = (\bar{\Psi}_1, \bar{\Psi}_2, \bar{\Psi}_3, \bar{\Psi}_4). \quad (10.2)$$

Here Ψ_α and $\bar{\Psi}_\alpha$ are independent Grassmann variables. The kinetic term for fermions involves only one derivative. This is simpler than the kinetic term for scalars, where we must use two derivatives. Under a Lorentz transformation, $\bar{\Psi}$ and $\bar{\Psi}$ transform as

$$\delta\Psi = \frac{i}{2}\epsilon_{\mu\nu}T^{\mu\nu}\Psi, \quad (10.3)$$

$$\delta\bar{\Psi} = -\frac{i}{2}\epsilon_{\mu\nu}\bar{\Psi}T^{\mu\nu}. \quad (10.4)$$

One can introduce a complex structure in the Grassmann algebra by defining Ψ^* through

$$\bar{\Psi} = \Psi^\dagger\gamma^0 = (\Psi^*)^T\gamma^0. \quad (10.5)$$

This is the defining relation for Ψ^* in terms of $\bar{\Psi}$. One can check the consistency of complex conjugation with Lorentz transformations,

$$\delta\Psi^* = -\frac{i}{2}\epsilon_{\mu\nu}T^{\mu\nu}\Psi^*, \quad \delta\bar{\Psi} = (\delta\Psi)^\dagger\gamma^0. \quad (10.6)$$

Having defined Ψ^* , one could define real and imaginary parts $\Psi_{\text{Re}} = \frac{1}{2}(\Psi + \Psi^*)$ and $\Psi_{\text{Im}} = -\frac{i}{2}(\Psi - \Psi^*)$ and use those as independent Grassmann variables.

Transformation of Spinor Bilinears In order to verify the invariance of S we consider general bilinear forms of spinors and check their properties under Lorentz transformations. We will only consider infinitesimal Lorentz transformations here. The first relation we proof is

$$\delta(\bar{\Psi}\Psi) = 0. \quad (10.7)$$

Indeed,

$$\delta(\bar{\Psi}\Psi) = \delta\bar{\Psi}\Psi + \bar{\Psi}\delta\Psi = -\frac{i}{2}(\bar{\Psi}T^{\mu\nu}\Psi - \bar{\Psi}T^{\mu\nu}\Psi) = 0. \quad (10.8)$$

This means that $\bar{\Psi}\Psi$ transforms as a scalar under Lorentz transformations. Next we will show that

$$\delta(\bar{\Psi}\gamma^\mu\Psi) = \delta\Gamma^\mu_\nu(\bar{\Psi}\gamma^\nu\Psi) = \epsilon^\mu_\nu\bar{\Psi}\gamma^\nu\Psi, \quad (10.9)$$

i.e. it transforms as a contravariant vector under Lorentz transformations. This can be seen in three steps. First we note that

$$\delta(\bar{\Psi}\gamma^\rho\Psi) = \delta\bar{\Psi}\gamma^\rho\Psi + \bar{\Psi}\gamma^\rho\delta\Psi = -\frac{i}{2}\epsilon_{\mu\nu}(\bar{\Psi}T^{\mu\nu}\gamma^\rho\Psi - \bar{\Psi}\gamma^\rho T^{\mu\nu}\Psi) = -\frac{i}{2}\epsilon_{\mu\nu}\bar{\Psi}[T^{\mu\nu}, \gamma^\rho]\Psi. \quad (10.10)$$

Second, we employ

$$\gamma^\mu\gamma^\nu\gamma^\rho = \gamma^\mu\{\gamma^\nu, \gamma^\rho\} - \gamma^\mu\gamma^\rho\gamma^\nu = 2\eta^{\nu\rho}\gamma^\mu - \gamma^\mu\gamma^\rho\gamma^\nu. \quad (10.11)$$

Using this, we find

$$\begin{aligned} [T^{\mu\nu}, \gamma^\rho] &= -\frac{i}{4}(\gamma^\mu\gamma^\nu\gamma^\rho - \gamma^\nu\gamma^\mu\gamma^\rho - \gamma^\rho\gamma^\mu\gamma^\nu + \gamma^\rho\gamma^\nu\gamma^\mu) \\ &= -\frac{i}{4}(2\eta^{\nu\rho}\gamma^\mu - \gamma^\mu\gamma^\rho\gamma^\nu - 2\eta^{\mu\rho}\gamma^\nu + \gamma^\nu\gamma^\rho\gamma^\mu - 2\eta^{\mu\rho}\gamma^\nu \\ &\quad + \gamma^\mu\gamma^\rho\gamma^\nu + 2\eta^{\nu\rho}\gamma^\mu - \gamma^\nu\gamma^\rho\gamma^\mu) \\ &= -i(\eta^{\nu\rho}\gamma^\mu - \eta^{\mu\rho}\gamma^\nu) \end{aligned} \quad (10.12)$$

Insertion of this commutation relation yields

$$\delta(\bar{\Psi}\gamma^\rho\Psi) = -\frac{i}{2}\bar{\Psi}\epsilon_{\mu\nu}(-i)(\eta^{\nu\rho}\gamma^\mu - \eta^{\mu\rho}\gamma^\nu)\Psi = -\frac{1}{2}\bar{\Psi}(\epsilon_\mu^\rho\gamma^\mu - \epsilon_\nu^\rho\gamma^\nu)\Psi = \epsilon_\nu^\rho\bar{\Psi}\gamma^\nu\Psi \quad (10.13)$$

Since we also know the transformation properties of ∂_ρ , we can easily check that $\bar{\Psi}\gamma^\rho\partial_\rho\Psi$ transforms as a scalar:

$$\delta(\bar{\Psi}\gamma^\rho\partial_\rho\Psi) = \epsilon_\nu^\rho\bar{\Psi} + \epsilon_\rho^\nu\bar{\Psi}\gamma^\rho\partial_\nu\Psi = \epsilon_{\rho\nu}\bar{\Psi}\gamma^\nu\partial^\rho\Psi + \epsilon_{\nu\rho}\bar{\Psi}\gamma^\nu\partial^\rho\Psi = 0. \quad (10.14)$$

Electrons with mass m We would now like to look at a system of free electrons. Such a system is described by

$$\mathcal{L}_H = i\bar{\Psi}\gamma^\mu\partial_\mu\Psi + \text{Im}(\bar{\Psi}\Psi) \quad (10.15)$$

b) Dirac Equation The functional variation of the associated action S with regard to $\bar{\Psi}$ leads to the famous Dirac equation

$$\frac{\delta S}{\delta \bar{\Psi}} = 0 \Rightarrow (\gamma^\mu\partial_\mu + m)\Psi = 0. \quad (10.16)$$

The equation is relativistic covariant, because \mathcal{L} is invariant. For a single particle state, this is also the Schrödinger equation, with Ψ interpreted as a wave function. Then Ψ is a complex function (not a Grassmann variable). This does not hold with interactions. But an external electromagnetic field can be added.

Energy-Momentum Relation To get to the energy momentum relation for a relativistic particle, we square the Dirac equation

$$\gamma^\nu\partial_\nu\gamma^\mu\partial_\mu\Psi = m^2\Psi. \quad (10.17)$$

To evaluate this ansatz, we make use of the anticommutator relation for the γ matrices

$$\frac{1}{2}\{\gamma^\nu, \gamma^\mu\}\partial_\nu\partial_\mu\Psi = \eta^{\nu\mu}\partial_\nu\partial_\mu\Psi = \partial^\mu\partial_\mu\Psi = m^2\Psi \quad (10.18)$$

Now the last equation is exactly the Klein-Gordon equation $(\partial^\mu\partial_\mu - m^2)\Psi = 0$, so we see that solutions for (10.16) solve this equation.

Plane waves of the form $\Psi = \Psi_0 e^{ip^\mu x_\mu} = \Psi_0 e^{-i(Et - \mathbf{p}\mathbf{x})}$ are the easiest solutions, they lead to

$$(E^2 - \mathbf{p}^2 - m^2)\Psi = 0 \Rightarrow E^2 = \mathbf{p}^2 + m^2. \quad (10.19)$$

So, $E = \pm\sqrt{p^2 + m^2}$ are both solutions. What does the solution with the negative energy describe?

Hamiltonian Formulation We multiply (10.16) with $-i\gamma^0$

$$-i\gamma^0\gamma^\mu\partial_\mu\Psi = -i(\gamma^0)^2\partial_0\Psi - i\gamma^0\gamma^k\partial_k\Psi = i\gamma^0 m\Psi \quad (10.20)$$

and introduce

$$\alpha^k = -\gamma^0\gamma^k = \gamma^k\gamma^0 = \begin{pmatrix} -\tau_k & 0 \\ 0 & \tau_k \end{pmatrix} \quad \beta = i\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (10.21)$$

which lead to

$$i\dot{\Psi} = -i\alpha^k\partial_k\Psi + m\beta\Psi \quad (10.22)$$

In the momentum basis, we get the “relativistic Schrödinger equation”

$$i\dot{\Psi} = H\Psi \quad \text{with} \quad H = \alpha^k p_k + m\beta. \quad (10.23)$$

Let’s switch to the rest frame of the particle ($\mathbf{p} = 0$). For the Hamiltonian we get

$$H = m \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (10.24)$$

This matrix mixes the Weyl spinors Ψ_L and Ψ_R

$$i\partial_t \begin{pmatrix} \Psi_L \\ \Psi_R \end{pmatrix} = m\beta \begin{pmatrix} \Psi_L \\ \Psi_R \end{pmatrix} = m \begin{pmatrix} \Psi_R \\ \Psi_L \end{pmatrix} \quad (10.25)$$

We can verify that H has 2 eigenvectors with positive energy ($E = +m$), and 2 with negative energy ($E = -m$). One sees again the negative energy states!

Interpretation of Dirac Equation, Positrons We construct linear combinations of Ψ_L and Ψ_R , which are mass eigenstates

$$\Psi_{\pm} = \frac{1}{\sqrt{2}}(\Psi_L \pm \Psi_R) \quad \text{and} \quad i\dot{\Psi}_{\pm} = \pm m\Psi_{\pm} \quad (10.26)$$

By conjugating the equation for Ψ_-

$$-i\dot{\Psi}_-^* = -m\Psi_-^* \Rightarrow i\dot{\Psi}_-^* = m\Psi_-^* \quad (10.27)$$

we see that Ψ_-^* is a mass eigenstate with positive eigenvalue $E = +m$. This field can be interpreted as a new particle field, called the positron field. The positron is the antiparticle to the electron. An essential consequence of Lorentz symmetry is the existence of antiparticles! We will see that Ψ_-^* has electric charge -e, while Ψ_+ has charge e. We use Ψ_+ for electrons and therefore $e < 0$.

c) Electrons and Positrons in the Electromagnetic Field We want to see, how electrons and positrons act in the electromagnetic field, that means why they have opposite charges. The electromagnetic field is given by $A_{\mu} = (-\phi, \mathbf{A})$, and the covariant Lagrangian by

$$\mathcal{L} = i\bar{\Psi}\gamma^{\mu}(\partial_{\mu} - ieA_{\mu})\Psi + im\bar{\Psi}i\Psi. \quad (10.28)$$

This leads via least action principle to the following modifications of the Dirac equation

$$\begin{aligned} \partial_t\Psi &\rightarrow (\partial_0 + ie\phi)\Psi, \\ \partial_k\Psi &\rightarrow (\partial_k - ieA_k)\Psi, \end{aligned} \quad (10.29)$$

or

$$i\dot{\Psi} = \left(\alpha^k(p_k - eA_k) + e\phi + \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} \right) \Psi \quad (10.30)$$

with

$$\alpha^k \begin{pmatrix} \Psi_L \\ \Psi_R \end{pmatrix} = \begin{pmatrix} -\tau_k & 0 \\ 0 & \tau_k \end{pmatrix} \begin{pmatrix} \Psi_L \\ \Psi_R \end{pmatrix} = \begin{pmatrix} -\tau_k\Psi_L \\ \tau_k\Psi_R \end{pmatrix}. \quad (10.31)$$

The action of α^k on the linear combinations (10.26) is as follows

$$\begin{aligned} \alpha^k\Psi_+ &= -\tau_k\Psi_-, \\ \alpha^k\Psi_- &= -\tau_k\Psi_+. \end{aligned} \quad (10.32)$$

Now we can insert Ψ_+, Ψ_- into (10.30), and we get

$$\begin{aligned} i\dot{\Psi}_+ &= (m + e\phi)\Psi_+ + i(\partial_k - ieA_k)\tau_k\Psi_-, \\ i\dot{\Psi}_- &= (-m + e\phi)\Psi_- + i(\partial_k - ieA_k)\tau_k\Psi_+. \end{aligned} \quad (10.33)$$

For the complex conjugate of Ψ_- one finds

$$i\dot{\Psi}_-^* = (m - e\phi)\Psi_-^* + i(\partial_k + ieA_k)\tau_k^*\Psi_+^* \quad (10.34)$$

Thus Ψ_-^* (positrons) has indeed the opposite charge as Ψ_+ (electrons).

d) Quantum electrodynamics We finally add a kinetic term for the photons

$$\mathcal{L}_F = \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \quad F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}. \quad (10.35)$$

Taking things together, we arrive at the functional integral for QED

$$Z = \int D\varphi \exp(iS_M), \quad (10.36a)$$

$$S_M = - \int_x \mathcal{L}_M, \quad (10.36b)$$

$$\mathcal{L}_M = i\bar{\Psi}\gamma^{\mu}(\partial_{\mu} - ieA_{\mu})\Psi + \text{Im}(\bar{\Psi}\Psi) + \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \quad (10.36c)$$

From there, all correlation functions can be computed! Precise computations with many decimal places agree perfectly with observation.

e) **Gauge symmetry** The action of QED is invariant under local gauge transformations.

$$\Psi'(x) = e^{i\alpha(x)\Psi(x)}, \quad (10.37)$$

$$A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu\alpha(x) \quad (10.38)$$

$\alpha(x)$: depends on x , leads to change of ψ at every x independently \rightarrow “local symmetry”.

$$\begin{aligned} \bar{\Psi}\gamma^\mu\partial_\mu\Psi &\rightarrow \bar{\Psi}\gamma^\mu\partial_\mu\Psi + i\partial_\mu\alpha\bar{\Psi}\gamma^\mu\Psi \\ -ie\bar{\Psi}\gamma^\mu A_\mu\Psi &\rightarrow -ie\bar{\Psi}\gamma^\mu A_\mu\Psi - i\partial_\mu\alpha\bar{\Psi}\gamma^\mu\Psi \Rightarrow i\bar{\Psi}\gamma^\mu(\partial_\mu - ieA_\mu)\Psi \text{ is invariant} \\ F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu &\Rightarrow F_{\mu\nu} + F_{\mu\nu} + \frac{1}{e}\partial_\mu\partial_\nu\alpha - \frac{1}{\alpha}\partial_\nu\partial_\mu\alpha = 0 \end{aligned} \quad (10.39)$$

$F_{\mu\nu}$ is invariant. It follows that \mathcal{L} is invariant.

Local gauge invariance is an important principle for finding the action of a quantum field theory. It is also related to renormalizability.

Renormalizability Gauge symmetry is a powerful restriction for the choice of the action. Is it sufficient? Consider a possible term

$$\Delta\mathcal{L} = \frac{b}{m}\bar{\Psi}[\gamma^\mu, \gamma^\nu]\Psi F_{\mu\nu} \quad (10.40)$$

It is Lorentz invariant and gauge invariant. If we add it with an unknown coefficient b , predictions will depend on this coefficient. Predictivity of QED, which only involves m and $\alpha = e^2/4\pi$, is lost. Why is such a term not allowed? This is again related to renormalizability.

g) Non-relativistic limit of Dirac equation For the two-component spinor χ for the electron one finds

$$i\partial_t\chi = \frac{1}{2m}(\vec{p} - e\vec{A})^2 + e\varphi - \frac{e}{m}\vec{S}\vec{B}, \quad \vec{S} = \frac{1}{2}\vec{\tau} \quad (10.41)$$

For “weak \vec{A} ” one linearizes in \vec{A} . For a constant \vec{B} one takes $\vec{A} = -\frac{1}{2}\vec{r} \times \vec{B}$ and obtains

$$\frac{1}{2m}(\vec{p} - e\vec{A})^2 = \frac{\vec{p}^2}{2m} - \frac{e}{2m}\vec{L}\vec{B}, \quad \vec{L} : \text{angular momentum} \quad (10.42)$$

This is the Schrödinger equation for atomic physics!

The magnetic field couples to

$$\vec{L} + g\vec{S}, \quad g = 2. \quad (10.43)$$

QED corrections from fluctuations yield small correction to $g - 2$, which is computed to many decimal places and well tested.

The derivation of the non-relativistic limit is done in several steps.

Step 1: Square the Dirac equation,

$$\gamma^\nu(\partial_\nu - ieA_\nu)\gamma^\mu(\partial_\mu - ieA_\mu)\Psi = m^2\Psi. \quad (10.44)$$

Step 2: From the Dirac algebra we use $[\gamma^\mu, \gamma^\nu] = 4iT^{\mu\nu}$ and obtain

$$((\partial^\mu - ieA^\mu)(\partial_\mu - ieA_\mu) + eT^{\mu\nu}F_{\mu\nu} - m^2)\Psi = 0 \quad (10.45)$$

Step 3: We use $T^{\mu\nu}F_{\mu\nu} = \frac{1}{2}B_k\tau_k + \frac{i}{2}E_k\tau_k\gamma^5$, with $\tau_k = \begin{pmatrix} \tau_k & 0 \\ 0 & \tau_k \end{pmatrix}$. Also using Ψ_\pm , we get

$$\{(\partial^\mu - ieA^\mu)(\partial_\mu - ieA_\mu) - m^2 + eB_k\tau_k\}\Psi_+ = -ieE_k\tau_k\Psi_- \quad (10.46)$$

Step 4: We forget about positrons: For $\Psi_- = 0$ one obtains an equation for a two component spinor Ψ_+ . Introduce the non-relativistic wave function χ by

$$\Psi_+ = e^{-imt} \chi. \quad (10.47)$$

This yields

$$i\partial_t \Psi = H\chi = (E - m)\chi \quad (10.48)$$

with non-relativistic Hamiltonian H . The non-relativistic limit is given by $H \ll m$. In this limit one can neglect

$$\frac{\partial_t^2}{m}, \quad \frac{A_0 \partial_t}{m}, \quad \frac{(\partial_t A_0)}{m}, \quad \frac{A_0^2}{m}. \quad (10.49)$$

This yields the above non-relativistic result.

Functional integral for photons. For photons, the field one integrates over in the functional integral is the gauge field $A_\mu(x)$. The field theory is described by the partition function

$$\begin{aligned} Z_2[J] &= \int DA \exp \left[iS_2[A] + i \int J^\mu A_\mu \right] \\ &= \int DA \exp \left[i \int d^4x \left\{ -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + J^\mu A_\mu \right\} \right] \end{aligned} \quad (10.50)$$

One can go to momentum space as usual

$$A_\mu(x) = \int \frac{d^4p}{(2\pi)^4} e^{ipx} A_\mu(p), \quad (10.51)$$

and finds for the term in the exponential

$$\begin{aligned} &\int_x \left\{ -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + J^\mu A_\mu \right\} \\ &= \frac{1}{2} \int \frac{d^4p}{(2\pi)^4} \left\{ -A_\mu(-p) (p^2 \eta^{\mu\nu} - p^\mu p^\nu) A_\nu(p) + J^\mu(-p) A_\mu(p) + A_\mu(-p) J^\mu(p) \right\}. \end{aligned} \quad (10.52)$$

The next step would now be to perform the Gaussian integral over A_μ by completing the square. However, a problem arises here: The “inverse propagator” for the gauge field

$$p^2 \eta^{\mu\nu} - p^\mu p^\nu = p^2 \mathcal{P}^{\mu\nu}(p) \quad (10.53)$$

is not invertible. We wrote it here in terms of

$$\mathcal{P}_\mu{}^\nu(p) = \delta_\mu{}^\nu - \frac{p_\mu p^\nu}{p^2}, \quad (10.54)$$

which is in fact a projector to the space orthogonal to p_ν

$$\mathcal{P}_\mu{}^\nu(p) \mathcal{P}_\nu{}^\rho(p) = \mathcal{P}_\mu{}^\rho(p). \quad (10.55)$$

As a projector matrix it has eigenvalues 0 and 1, only. However,

$$\mathcal{P}_\mu{}^\nu(p) p_\nu = 0. \quad (10.56)$$

The field $A_\nu(p)$ can be decomposed into two parts,

$$A_\nu(p) = \frac{i}{e} p_\nu \beta(p) + \hat{A}_\nu(p) \quad (10.57)$$

with

$$\hat{A}_\nu(p) = \mathcal{P}_\nu{}^\rho(p) A_\rho(p) \quad (10.58)$$

such that $p^\nu \hat{A}_\nu(p) = 0$. Moreover

$$\beta(p) = \frac{e}{ip^2} p^\nu A_\nu(p). \quad (10.59)$$

When acting on $\hat{A}_\nu(p)$, the projector $\mathcal{P}_\mu{}^\nu(p)$ is simply the unit matrix.

Recall that gauge transformations shift the field according to

$$A_\mu(x) \rightarrow \frac{1}{e} \partial_\mu \alpha + A_\mu(x) \quad (10.60)$$

or in momentum space

$$A_\mu(p) \rightarrow \frac{i}{e} p_\mu \alpha(p) + A_\mu(p). \quad (10.61)$$

One can therefore always perform a gauge transformation such that $\beta(p) = 0$ or

$$\partial^\mu A_\mu(x) = 0. \quad (10.62)$$

This is known as Lorenz gauge or Landau gauge. We will use this gauge in the following and restrict the functional integral to field configurations that fulfil the gauge condition.

Now we can easily perform the Gaussian integral,

$$\begin{aligned} Z_2[J] &= \int DA \exp \left[\frac{i}{2} \int_p \left\{ - \left(A_\mu(-p) - J_\rho(-p) \frac{\mathcal{P}^\rho{}_\mu}{p^2} \right) p^2 \mathcal{P}^{\mu\nu} \left(A_\nu(p) - \frac{\mathcal{P}_\nu{}^\sigma}{p^2} J_\sigma(p) \right) \right\} \right] \\ &\quad \times \exp \left[\frac{i}{2} \int_p J_\mu(-p) \frac{\mathcal{P}^{\mu\nu}(p)}{p^2} J_\nu(p) \right] \\ &= \text{const} \times \exp \left[\frac{i}{2} \int_{x,y} J_\mu(x) \Delta^{\mu\nu}(x-y) J_\nu(y) \right]. \end{aligned} \quad (10.63)$$

In the last line we used the photon propagator in position space (in Landau gauge)

$$\Delta^{\mu\nu}(x-y) = \int \frac{d^4 p}{(2\pi)^4} e^{ip(x-y)} \frac{\mathcal{P}^{\mu\nu}(p)}{p^2 - i\epsilon}. \quad (10.64)$$

In the last step we have inserted the $i\epsilon$ term as usual.

In the free theory one has

$$\langle A_\mu(x) A_\nu(y) \rangle = \frac{1}{i^2} \left(\frac{1}{Z[J]} - \frac{\delta^2}{\delta J^\mu(x) \delta J^\nu(y)} Z[J] \right)_{J=0} = \frac{1}{i} \Delta_{\mu\nu}(x-y). \quad (10.65)$$

We use the following graphical notation

$$(x, \mu) \text{---} \text{wavy line} \text{---} (y, \nu) = \frac{1}{i} \Delta_{\mu\nu}(x-y) \quad (10.66)$$

or with sources $iJ^\mu(x)$ at the end points

$$\bullet \text{---} \text{wavy line} \text{---} \bullet = \frac{1}{2} \int_{x,y} iJ^\mu(x) \frac{1}{i} \Delta_{\mu\nu}(x-y) iJ^\nu(y). \quad (10.67)$$

Mode decomposition for free photons. To describe incoming and outgoing photons we need to discuss free solutions for the gauge field. In momentum space, and for the gauge-fixed field (Landau gauge), the linear equation of motion (Maxwell's equation) is simply

$$p^2 \mathcal{P}_\mu{}^\nu(p) \hat{A}_\nu(p) = p^2 \hat{A}_\mu(p) = 0. \quad (10.68)$$

Non-trivial solutions satisfy $p^2 = 0$. Without loss of generality we assume now $p^\mu = (E, 0, 0, E)$; all other light like momenta can be obtained from this via Lorentz-transformations. Quite generally, a four-vector can be written as

$$\hat{A}_\nu(p) = \left(b, \frac{a_1 + a_2}{\sqrt{2}}, \frac{-ia_1 + ia_2}{\sqrt{2}}, c \right). \quad (10.69)$$

From the Landau gauge condition $p^\nu \hat{A}_\nu = 0$ it follows that $b = -c$, so that one can write

$$\hat{A}_\nu(p) = c \times (-1, 0, 0, 1) + a_1 \epsilon_\nu^{(1)} + a_2 \epsilon_\nu^{(2)} \quad (10.70)$$

with

$$\epsilon_\nu^{(1)} = \left(0, \frac{1}{\sqrt{2}}, \frac{-i}{\sqrt{2}}, 0 \right), \quad \epsilon_\nu^{(2)} = \left(0, \frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0 \right). \quad (10.71)$$

However, the term $\sim c$ is in fact proportional to $p_\nu = (-E, 0, 0, E)$. We can do another gauge transformation such that $c = 0$. This does not violate the Landau gauge condition because of $p^\nu p_\nu = 0$. In other words, the photon field has only two independent polarization states, chosen here as positive and negative circular polarizations, or helicities.

In summary, we can expand free solutions of the photon field like

$$A_\mu(x) = \sum_{\lambda=1}^2 \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left\{ a_{\vec{p},\lambda} \epsilon_\mu^{(\lambda)}(p) e^{ipx} + a_{\vec{p},\lambda}^\dagger \epsilon_\mu^{(\lambda)*}(p) e^{-ipx} \right\} \quad (10.72)$$

where $E_p = |\vec{p}|$ is the energy of a photon. The index λ labels the two polarization states.

In the current setup, $a_{\vec{p},\lambda}$ and $a_{\vec{p},\lambda}^\dagger$ are simply expansion coefficients, while they become annihilation and creation operators in the operator picture. The non-trivial commutation relation becomes then

$$\left[a_{\vec{p},\lambda}, a_{\vec{p}',\lambda'}^\dagger \right] = (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{p}') \delta_{\lambda\lambda'}. \quad (10.73)$$

LSZ reduction formula for photons. We also need a version of the Lehmann-Symanzik-Zimmermann reduction formula for photons. Recall that for non-relativistic bosons we could replace for the calculation of the interacting part of the S-matrix

$$a_{\vec{q}}(\infty) \rightarrow i \left[-q^0 + \frac{q^2}{2m} + V_0 \right] \varphi(q), \quad (10.74)$$

$$a_{\vec{q}}^\dagger(-\infty) \rightarrow i \left[-q^0 + \frac{q^2}{2m} + V_0 \right] \varphi^*(q). \quad (10.75)$$

For relativistic fields this is in general somewhat more complicated because of renormalization. This will be discussed in more detail in the second part of the course. In the following we will discuss only tree level diagrams where this plays no role. For photons one can replace for outgoing states

$$\begin{aligned} \sqrt{2E_p} a_{\vec{p},\lambda}(\infty) &\rightarrow i \epsilon_{(\lambda)}^{\nu*}(p) \int d^4x e^{-ipx} [-\partial_\mu \partial^\mu] A_\nu(x) \\ \sqrt{2E_p} a_{\vec{p},\lambda}^\dagger(-\infty) &\rightarrow i \epsilon'_{(\lambda)}(p) \int d^4x e^{ipx} [-\partial_\mu \partial^\mu] A_\nu(x). \end{aligned} \quad (10.76)$$

These formulas can be used to write S-matrix elements as correlation functions of fields. Note that $[-\partial_\mu \partial^\nu]$ is essentially the inverse propagator in Landau gauge.

Mode expansion for Dirac fields. We also need a mode expansion for free Dirac fields in order to describe asymptotic (incoming and outgoing) fermion states. We write the fields as

$$\begin{aligned}\psi(x) &= \sum_{s=1}^2 \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left\{ b_{\vec{p},s} u_s(p) e^{ipx} + d_{\vec{p},s}^\dagger v_s(p) e^{-ipx} \right\}, \\ \psi(\bar{x}) &= \sum_{s=1}^2 \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left\{ -i b_{\vec{p},s}^\dagger \bar{u}_s(p) e^{-ipx} - i d_{\vec{p},s}^\dagger \bar{v}_s(p) e^{ipx} \right\}.\end{aligned}\tag{10.77}$$

Again, $b_{\vec{p},s}$, $d_{\vec{p},s}$ etc. can be seen as expansion coefficients and become operators in the operator picture. The Dirac equation

$$(\gamma^\mu \partial_\mu + m)\psi(x) = 0,\tag{10.78}$$

becomes for the plane waves

$$(i\gamma^\mu p_\mu + m) u_s(\vec{p}) e^{ipx} + (-i\gamma^\mu p_\mu + m) v_s(\vec{p}) e^{-ipx} = 0.\tag{10.79}$$

To solve this one needs

$$\begin{aligned}(i\not{p} + m) u_s(\vec{p}) &= 0, \\ (-i\not{p} + m) v_s(\vec{p}) &= 0,\end{aligned}\tag{10.80}$$

with $\not{p} = \gamma^\mu p_\mu$. We consider this first in the frame where the spatial momentum vanishes, $\vec{p} = 0$, such that $p_\mu = (-m, 0, 0, 0)$,

$$\not{p} = -\gamma^0 m = im \begin{pmatrix} \mathbb{1} & \\ & \mathbb{1} \end{pmatrix}.\tag{10.81}$$

The last equation holds in the chiral basis where

$$\gamma^\mu = -i \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}.\tag{10.82}$$

with $\sigma^\mu = (\mathbb{1}, \vec{\sigma})$ and $\bar{\sigma}^\mu = (\mathbb{1}, -\vec{\sigma})$. For the spinor u_s one has the equation

$$(i\not{p} + m)u_s = m \begin{pmatrix} +\mathbb{1} & -\mathbb{1} \\ -\mathbb{1} & +\mathbb{1} \end{pmatrix} u_s = 0.\tag{10.83}$$

The two independent solutions are

$$u_1^{(0)} = \sqrt{m} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad u_2^{(0)} = \sqrt{m} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.\tag{10.84}$$

The normalization has been chosen for later convenience. Similarly

$$(-i\not{p} + m)v_s(0) = m \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix} v_s(0) = 0\tag{10.85}$$

has the two independent solutions

$$v_1^{(0)} = \sqrt{m} \begin{pmatrix} 0 \\ +1 \\ 0 \\ -1 \end{pmatrix}, \quad v_2^{(0)} = \sqrt{m} \begin{pmatrix} -1 \\ 0 \\ +1 \\ 0 \end{pmatrix}.\tag{10.86}$$

We see here that the Dirac equation has two independent solutions (for spin up and down with respect to some basis) for particles and two more for anti-particles. One can now go to an arbitrary reference frame by performing a Lorentz transformation. That gives

$$u_s(\vec{p}) = \begin{pmatrix} \sqrt{-p_\mu \sigma^\mu} \xi_s \\ \sqrt{-p_\mu \bar{\sigma}^\mu} \xi_s \end{pmatrix}, \quad v_s(\vec{p}) = \begin{pmatrix} \sqrt{-p_\mu \sigma^\mu} \xi_s \\ -\sqrt{-p_\mu \bar{\sigma}^\mu} \xi_s \end{pmatrix}, \quad (10.87)$$

with a two-dimensional orthonormal basis ξ_s such that

$$\xi_s^\dagger \xi_r = \delta_{rs}, \quad \sum_{s=1}^2 \xi_s \xi_s^\dagger = \mathbb{1}_2. \quad (10.88)$$

Other identities involving $u_s(\vec{p})$, $v_s(\vec{p})$ as well as

$$\begin{aligned} \bar{u}_s(\vec{p}) &= u_s^\dagger(\vec{p}) i \gamma^0 = u_s^\dagger(p) \begin{pmatrix} \mathbb{1} \\ \mathbb{1} \end{pmatrix}, \\ \bar{v}_s(\vec{p}) &= v_s^\dagger(\vec{p}) i \gamma^0 = v_s^\dagger(p) \begin{pmatrix} \mathbb{1} \\ \mathbb{1} \end{pmatrix}, \end{aligned} \quad (10.89)$$

have been discussed in exercises. They will be mentioned here once they are needed.

LSZ reduction for Dirac fermions. Finally, let us give the LSZ reduction formulas for Dirac fermions (again neglecting renormalization effects)

$$\begin{aligned} \sqrt{2E_p} b_{\vec{p},s}(\infty) &\rightarrow i \int d^4x e^{-ipx} \bar{u}_s(\vec{p}) (\gamma^\mu \partial_\mu + m) \psi(x), \\ \sqrt{2E_p} d_{\vec{p},s}^\dagger(-\infty) &\rightarrow -i \int d^4x e^{-ipx} \bar{v}_s(\vec{p}) (\gamma^\mu \partial_\mu + m) \psi(x), \\ \sqrt{2E_p} d_{\vec{p},s}(\infty) &\rightarrow -i \int d^4x i \bar{\psi}_s(x) (-\gamma^\mu \overleftarrow{\partial}_\mu + m) v_s(x) e^{-ipx}, \\ \sqrt{2E_p} b_{\vec{p},s}(-\infty) &\rightarrow -i \int d^4x i \bar{\psi}_s(x) (-\gamma^\mu \overleftarrow{\partial}_\mu + m) u_s(x) e^{ipx}. \end{aligned} \quad (10.90)$$

The left-pointing arrows indicate here that these derivatives act to the left (on the field $\bar{\psi}_s(x)$). These relations have been obtained as part of the exercises.

10.2 Feynman rules and Feynman diagrams

Feynman rules of QED. We are now ready to formulate the Feynman rules for a perturbative treatment of quantum electrodynamics. The microscopic action is

$$\begin{aligned} S &= \int d^4x \left\{ -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - i \bar{\psi} \gamma^\mu (\partial_\mu - ie A_\mu) \psi - im \bar{\psi} \psi \right\} \\ &= S_2[\bar{\psi}, \psi, A] - \int d^4x e \bar{\psi} \gamma^\mu A_\mu \psi. \end{aligned} \quad (10.91)$$

The last term is cubic in the fields $\bar{\psi}, \psi$ and A_μ , while all others terms are quadratic. We will perform a perturbative expansion in the electric charge e .

Let us write the partition function as

$$Z[\bar{\eta}, \eta, J] = \int D\bar{\psi} D\psi DA \exp \left[i S[\bar{\psi}, \psi, A] + i \int \{ \bar{\eta} \psi + \bar{\psi} \eta + J^\mu A_\mu \} \right] \quad (10.92)$$

with $\bar{\eta} \psi = \bar{\eta}_\alpha \psi_\alpha$ where $\alpha = 1, \dots, 4$ sums over spinor components. Formally, one can write

$$Z[\bar{\eta}, \eta, J] = \exp \left[-e \int d^4x \left(\frac{1}{i} \frac{\delta}{\delta J^\mu(x)} \right) \left(i \frac{\delta}{\delta \eta_\alpha(x)} \right) (\gamma^\mu)_{\alpha\beta} \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}_\beta(x)} \right) \right] Z_2[\bar{\eta}, \eta, J], \quad (10.93)$$

with quadratic partition function

$$\begin{aligned} Z_2 &= \int D\bar{\psi}D\psi DA \exp \left[iS_2[\bar{\psi}, \psi, A] + i \int \{ \bar{\eta}\psi + \bar{\psi}\eta + J^\mu A_\mu \} \right] \\ &= \exp \left[i \int d^4x d^4y \bar{\eta}(x)S(x-y)\eta(y) \right] \times \exp \left[\frac{i}{2} \int d^4x d^4y J^\mu(x)\Delta_{\mu\nu}(x-y)J^\nu(y) \right]. \end{aligned} \quad (10.94)$$

We have introduced here also the propagator for Dirac fermions, which is in fact a matrix in spinor space,

$$\begin{aligned} S_{\alpha\beta}(x-y) &= \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} (ip_\mu\gamma^\mu + m)_{\alpha\beta}^{-1} \\ &= \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \frac{(-i\not{p} + m\mathbb{1})_{\alpha\beta}}{p^2 + m^2 - i\epsilon}. \end{aligned} \quad (10.95)$$

We can now calculate S-matrix elements by first expressing them as correlation functions which get then evaluated in a perturbative expansion of the functional integral. These perturbative expressions have an intuitive graphical representation as we have briefly discussed before. We concentrate here on tree diagrams for which renormalization is not needed yet.

From the quadratic function one can obtain also Dirac field propagator

$$\langle \psi_\alpha(x)\bar{\psi}_\beta(y) \rangle = \frac{1}{Z_2} \left(\frac{1}{i} \frac{\delta}{\delta\bar{\eta}_\alpha(x)} \right) \left(i \frac{\delta}{\delta\eta_\beta(y)} \right) Z_2 \Big|_{\bar{\eta}=\eta=J=0} = \frac{1}{i} S_{\alpha\beta}(x-y). \quad (10.96)$$

We introduce a graphical representation for thus, as well,

$$(x, \alpha) \longleftarrow (y, \beta) = \frac{1}{i} S_{\alpha\beta}(x-y). \quad (10.97)$$

With sources $i\bar{\eta}_\alpha(x)$ and $i\eta_\beta(y)$ at the end this would be

$$\bullet \longleftarrow \bullet = \int_{x,y} i\bar{\eta}_\alpha(x) \frac{1}{i} S_{\alpha\beta}(x-y) i\eta_\beta(y) = i \int_{x,y} \bar{\eta}(x)S(x-y)\eta(y). \quad (10.98)$$

The conventions are such that the arrow points away from the source η and to the source $\bar{\eta}$. It can also be seen as denoting the direction of fermions while anti-fermions move against the arrow direction. The Dirac indices α, β are sometimes left implicit when there is no doubt about them.

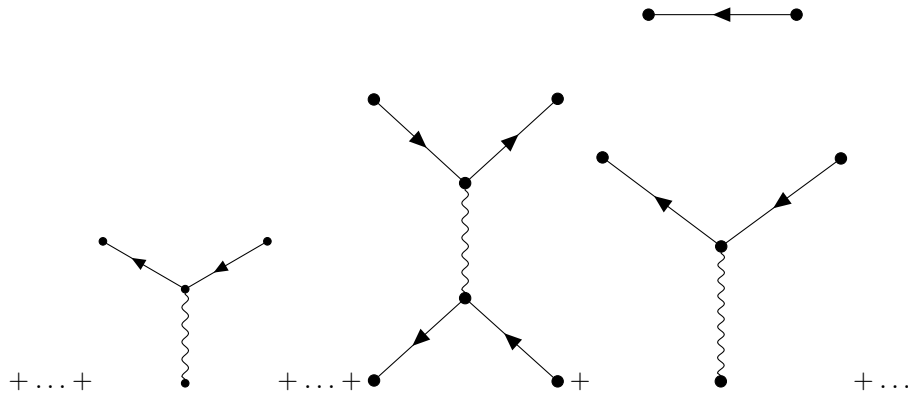
We now consider the full partition function and expand out the exponentials,

$$\begin{aligned} Z[\bar{\eta}, \eta, J] &= \sum_{V=0}^{\infty} \frac{1}{V!} \left[\int_x \left(\frac{1}{i} \frac{\delta}{\delta J^\mu(x)} \right) \left(i \frac{\delta}{\delta\eta_\alpha(x)} \right) (-e\gamma_{\alpha\beta}^\mu) \left(\frac{1}{i} \frac{\delta}{\delta\bar{\eta}_\beta(x)} \right) \right]^V \\ &\quad \times \sum_{F=0}^{\infty} \frac{1}{F!} \left[\int_{x',y'} i\bar{\eta}_\alpha(x') \left(\frac{1}{i} S_{\alpha\beta}(x'-y') \right) i\eta_\beta(y') \right]^F \\ &\quad \times \sum_{P=0}^{\infty} \frac{1}{P!} \left[\frac{1}{2} \int_{x'',y''} iJ^\mu(x'') \left(\frac{1}{i} \Delta_{\mu\nu}(x''-y'') \right) iJ^\nu(y'') \right]^P. \end{aligned} \quad (10.99)$$

The index F counts the number of fermion propagators (corresponding to fermion lines in a graphical representation), the index P counts the number of photon propagators (photon lines). The index V counts vertices that connect fermion and photon in a specific way. More specifically, each power of this term removes one of each kind of sources and introduces $-e\gamma_{\alpha\beta}^\mu$ to connect the lines in the graphical representation. In the full expression for $Z[\eta, \bar{\eta}, J]$ many terms are present, in fact all graphs one can construct with fermion lines, photon lines and the vertex.

For example

(10.100)



One distinguishes connected diagrams where all endpoints are connected with lines to each other, for example

(10.101)

Disconnected diagrams can be decomposed into several connected diagrams.

One also distinguishes tree diagrams and loop diagrams. Loop diagrams have closed loops of

particle flow, for example

$$(10.102)$$

Tree diagrams have no closed loop, for example

$$(10.103)$$

To each of these diagrams with sources one can associate an expression, for example

$$(10.104)$$

$$= \int_{x,y,z,w} i\bar{\eta}(x) \left[\frac{1}{i} S(x-z) \right] (-e\gamma^\mu) \left[\frac{1}{i} S(z-y) \right] i\eta(y) \left[\frac{1}{i} \Delta_{\mu\nu}(z-w) \right] iJ^\nu(w).$$

To calculate S-matrix elements we are mainly interested in the connected diagrams because disconnected diagrams describe events where not all particles scatter. Also, we concentrate here on tree diagrams. Loop diagrams will be discussed somewhat later.

Now that we have seen how to represent $Z[\bar{\eta}, \eta, J]$, let us discuss how to obtain S-matrix elements. For example, for an outgoing photon we had the LSZ rule

$$\sqrt{2E_p} a_{\bar{p},\lambda}(\infty) \rightarrow i\epsilon_{(\lambda)}^{\nu*}(p) \int d^4x e^{-ipx} [-\partial_\mu \partial^\mu] A_\nu(x) \quad (10.105)$$

To obtain the field A_ν under the functional integral we can use

$$A_\nu(x) \rightarrow \frac{1}{i} \frac{\delta}{\delta J^\nu(x)}, \quad (10.106)$$

acting on $Z[\bar{\eta}, \eta, J]$. Moreover, $i[-\partial_\mu \partial^\mu]$ will remove one propagator line for the outgoing photon,

$$\begin{aligned} i[-\partial_\mu \partial^\mu] \frac{1}{i} \Delta_{\rho\sigma}(x-y) &= [-\partial_\mu \partial^\mu] \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \frac{\mathcal{P}_{\rho\sigma}(p)}{p^2 - i\epsilon} \\ &= \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \mathcal{P}_{\rho\sigma}(p) \rightarrow \eta_{\rho\sigma} \delta^{(4)}(x-y). \end{aligned} \quad (10.107)$$

The projector has no effect if the photon couples to conserved currents and the result is simply $\eta_{\rho\sigma}\delta^{(4)}(x-y)$. What remains is to multiply with the polarization vector

$$\epsilon_{(\lambda)\mu}^*(p) \quad (10.108)$$

for the out-going photon with momentum p . Also, the Fourier transform brings the expression to momentum space. The out-going momentum is on-shell, i. e. it satisfies $p_\mu p^\mu = 0$ for photons. Similarly, for incoming photons we need to remove the external propagator line and contract with

$$\epsilon_{(\lambda)\mu}(p) \quad (10.109)$$

instead.

For out-going electrons we need to remove the external fermion propagator and multiply with $\bar{u}_s(\vec{p})$ where p is the momentum of the out-going electron satisfying $p^2 + m^2 = 0$ and s labels its spin state. Similarly, for an incoming electron we need to contract with $iu_s(p)$.

For out-going positrons we need to contract with $i\bar{v}_s(p)$ (and include here one factor i because $i\bar{\psi}$ appears in the LSZ rule in our conventions). For an incoming positron the corresponding external spinor is $\bar{v}_s(p)$.

Working now directly in momentum space, the photon propagator is represented by

$$-i \frac{\mathcal{P}_{\mu\nu}(p)}{p^2 - i\epsilon} = -i \frac{\eta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}}{p^2 - i\epsilon}. \quad (10.110)$$

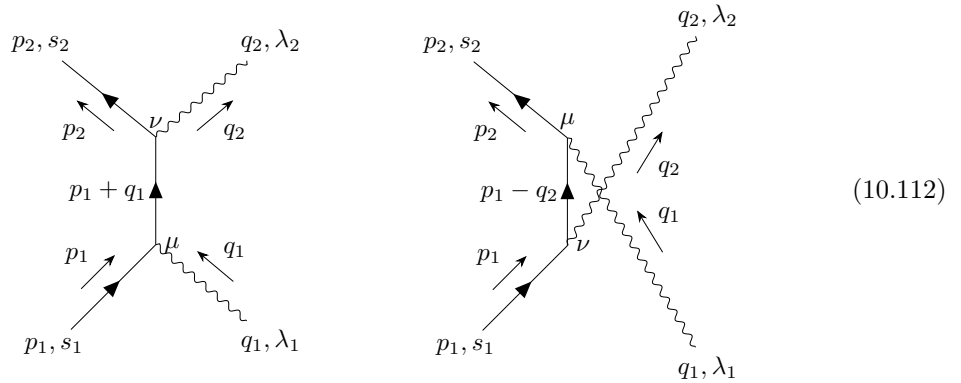
The fermion propagator is

$$-i \frac{-i\not{p} + m}{p^2 + m^2 - i\epsilon}. \quad (10.111)$$

The vertex is as before $-e\gamma^\mu$. Momentum conservation must be imposed at each vertex. Together these rules constitute the Feynman rules of QED. One can work with the graphical representation and then translate to formula at a convenient point. However, when in doubt, one can always go back to the functional representation.

10.3 Elementary scattering processes

Compton Scattering. As a first example let us consider Compton scattering $e^-\gamma \rightarrow e^-\gamma$



These are two diagrams at order e^2 , as shown above. The first diagram corresponds to the expression

$$\bar{u}_{s_2}(p_2)(-e\gamma^\nu) \left(-i \frac{-i(\not{p}_1 + \not{q}_1) + m}{(p_1 + q_1)^2 + m^2} \right) (-e\gamma^\mu) iu_s(p_1) \epsilon_{(\lambda_1)\mu}(q_1) \epsilon_{(\lambda_2)\nu}^*(q_2). \quad (10.113)$$

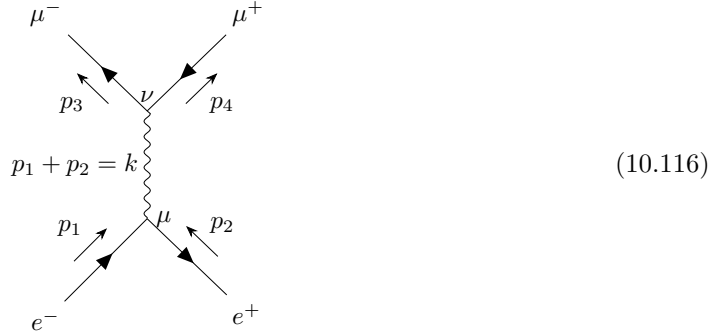
Similarly, the second diagram gives

$$\bar{u}_{s_2}(p_2)(-e\gamma^\mu) \left(-i \frac{-i(\not{p}_1 - \not{q}_1) + m}{(p_1 - q_1)^2 + m^2} \right) (-e\gamma^\nu) iu_s(p_1) \epsilon_{(\lambda_1)\mu}(q_1) \epsilon_{(\lambda_2)\nu}^*(q_2). \quad (10.114)$$

Combining terms and simplifying a bit leads to

$$i\mathcal{T} = e^2 \epsilon_{(\lambda_1)\mu}(q_1) \epsilon_{(\lambda_2)\nu}^*(q_2) \bar{u}_{s_2}(p_2) \left[\gamma^\nu \frac{-i(\not{p}_1 + \not{q}_1) + m}{(p_1 + q_1)^2 + m^2} \gamma^\mu + \gamma^\mu \frac{-i(\not{p}_1 - \not{q}_2) + m}{(p_1 - q_2)^2 + m^2} \gamma^\nu \right] u_{s_1}(p_1). \quad (10.115)$$

Electron-positron to muon-anti-muon scattering. As another example for an interesting process in QED we consider $e^-e^+ \rightarrow \mu^- \mu^+$. From the point of view of QED, the muon behaves like the electron but has a somewhat larger mass. Diagrams contributing to this process are (we keep the polarizations implicit)



The corresponding expression is

$$i\mathcal{T} = \bar{v}(p_2)(-e\gamma^\mu) iu(p_1) \left(-i \frac{\eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}}{(k^2)} \right) \bar{u}(p_2) (-e\gamma^\nu) iv(p_4), \quad (10.117)$$

with $k = p_1 + p_2 = p_3 + p_4$. The external momenta are on-shell and the spinors $u(p_1)$ etc. satisfy the Dirac equation,

$$(i\not{p}_1 + m_e)u(p_1) = 0, \quad (-i\not{p}_4 + m_\mu)v(p_4) = 0, \quad (10.118)$$

$$\bar{u}(p_3)(i\not{p}_3 + m_\mu) = 0, \quad \bar{v}(p_2)(-i\not{p}_2 + m_e) = 0. \quad (10.119)$$

This allows to write

$$\begin{aligned} i\bar{v}(p_2) \gamma^\mu k_\mu u(p_1) &= i\bar{v}(p_2) (\not{p}_1 + \not{p}_2) u(p_1) = \bar{v}(p_2) (-m_e + m_e) u(p_1) = 0, \\ i\bar{u}(p_3) \gamma^\nu k_\nu v(p_4) &= i\bar{u}(p_3) (\not{p}_3 + \not{p}_4) v(p_4) = \bar{u}(p_3) (-m_\mu + m_\mu) v(p_4) = 0. \end{aligned} \quad (10.120)$$

These arguments show that the term $\sim k_\mu k_\nu$ can be dropped. This is essentially a result of gauge invariance. One is left with

$$\mathcal{T} = \frac{e^2}{k^2} \bar{v}(p_2) \gamma^\mu u(p_1) \bar{u}(p_3) \gamma_\mu v(p_4). \quad (10.121)$$

To calculate $|\mathcal{T}|^2$ we also need \mathcal{T}^* which follows from hermitian conjugation

$$\mathcal{T}^* = \frac{e^2}{k^2} v^\dagger(p_4) \gamma_\mu^\dagger \bar{u}^\dagger(p_3) u^\dagger(p_1) \gamma^{\mu\dagger} \bar{v}^\dagger(p_2). \quad (10.122)$$

Recall that $\bar{u}(p) = u(p)^\dagger \beta$ with $\beta = i\gamma^0$. With the explicit representation

$$\gamma^\mu = \begin{pmatrix} & -i\bar{\sigma}^\mu \\ -i\sigma^\mu & \end{pmatrix}, \quad (10.123)$$

it is also easy to prove $\beta\gamma^{\mu\dagger}\beta = -\gamma^\mu$. By inserting $\beta^2 = \mathbb{1}$ at various places we find thus

$$\mathcal{T}^* = \frac{e^2}{k^2} \bar{v}(p_4)\gamma_\mu u(p_3) \bar{u}(p_1)\gamma^\mu v(p_2) \quad (10.124)$$

Putting together and using $s = -k^2 = -(p_1 + p_2)^2$ we obtain

$$|\mathcal{T}|^2 = \frac{e^4}{s^2} \bar{u}(p_1)\gamma^\mu v(p_2) \bar{v}(p_2)\gamma^\nu u(p_1) \bar{u}(p_3)\gamma_\nu v(p_1) \bar{v}(p_4)\gamma_\mu u(p_3). \quad (10.125)$$

To proceed further, we need to specify also the spins of the incoming and outgoing particles. The simplest case is the one of unpolarized particles so that we need to average the spins of the incoming electrons, and to sum over possible spins in the final state. Summing over the spins of the μ^+ can be done as follows (exercise)

$$\sum_{s=1}^2 v_s(p_4)\bar{v}_s(p_4) = -i\not{p}_4 - m_\mu, \quad (10.126)$$

and similarly for μ^-

$$\sum_{s=1}^2 u_s(p_3)\bar{u}_s(p_3) = -i\not{p}_3 + m_\mu. \quad (10.127)$$

We can therefore write

$$\bar{u}(p_3)\gamma_\nu v(p_4) \bar{v}(p_4)\gamma_\mu u(p_3) = \text{tr} \left\{ (-i\not{p}_3 + m_\mu)\gamma_\nu (-i\not{p}_4 - m_\mu)\gamma_\mu \right\}. \quad (10.128)$$

Spins of the electron and positron must be averaged instead,

$$\begin{aligned} \frac{1}{2} \sum_{s=1}^2 u(p_1)\bar{u}(p_1) &= \frac{1}{2}(-i\not{p}_1 + m_e), \\ \frac{1}{2} \sum_{s=1}^2 v(p_2)\bar{v}(p_2) &= \frac{1}{2}(-i\not{p}_2 - m_e). \end{aligned} \quad (10.129)$$

This leads to

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{T}|^2 = \frac{e^4}{4s^2} \text{tr} \left\{ (-i\not{p}_1 + m_e)\gamma^\mu (-i\not{p}_2 - m_e)\gamma^\nu \right\} \times \text{tr} \left\{ (-i\not{p}_3 + m_\mu)\gamma_\nu (-i\not{p}_4 - m_\mu)\gamma_\mu \right\}. \quad (10.130)$$

In order to proceed further, we need to know how to evaluate traces of up to four gamma matrices.

Traces of gamma matrices. We need to understand how to evaluate traces of the form $\text{tr}\{\gamma^{\mu_1} \dots \gamma^{\mu_n}\}$. To work them out we can use $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$, $\gamma_5^2 = \mathbb{1}$ and $\{\gamma^\mu, \gamma_5\} = 0$. Also, $\text{tr}\{\mathbb{1}\} = 4$. First we prove that traces of an odd number of gamma matrices must vanish,

$$\begin{aligned} \text{tr}\{\gamma^{\mu_1} \dots \gamma^{\mu_n}\} &= \text{tr}\{\gamma_5^2 \gamma^{\mu_1} \gamma_5^2 \dots \gamma_5^2 \gamma^{\mu_n}\} \\ &= \text{tr}\{(\gamma_5 \gamma^{\mu_1} \gamma_5) \dots (\gamma_5 \gamma^{\mu_n} \gamma_5)\} \\ &= \text{tr}\{(-\gamma_5^2 \gamma_1^{\mu_1}) \dots (-\gamma_5^2 \gamma_n^{\mu_n})\} \\ &= (-1)^n \text{tr}\{\gamma^{\mu_1} \dots \gamma^{\mu_n}\}. \end{aligned} \quad (10.131)$$

This implies what we claimed. Now for even numbers

$$\text{tr}\{\gamma^\mu \gamma^\nu\} = \text{tr}\{\gamma^\nu \gamma^\mu\} = \frac{1}{2} \text{tr}\{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu\} = \eta^{\mu\nu} \text{tr}\{\mathbb{1}\} = 4\eta^{\mu\nu}. \quad (10.132)$$

From this it also follows that

$$\text{tr}\{\not{p}\not{q}\} = 4p \cdot q. \quad (10.133)$$

Now consider $\text{tr}\{\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\}$. This idea is to commute γ^μ to the right using $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$. Thus

$$\begin{aligned}\text{tr}\{\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\} &= -\text{tr}\{\gamma^\nu\gamma^\mu\gamma^\rho\gamma^\sigma\} + 2\eta^{\mu\nu} \text{tr}\{\gamma^\rho\gamma^\sigma\} \\ &= \text{tr}\{\gamma^\nu\gamma^\rho\gamma^\mu\gamma^\sigma\} - 2\eta^{\rho\mu}\text{tr}\{\gamma^\nu\gamma^\sigma\} + 2\eta^{\mu\nu} \text{tr}\{\gamma^\rho\gamma^\sigma\} \\ &= -\text{tr}\{\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^\mu\} + 2\eta^{\sigma\mu} \text{tr}\{\gamma^\nu\gamma^\rho\} - 2\eta^{\rho\mu} \text{tr}\{\gamma^\nu\gamma^\sigma\} + 2\eta^{\mu\nu} \text{tr}\{\gamma^\rho\gamma^\sigma\}.\end{aligned}\tag{10.134}$$

But by the cyclic property of the trace

$$\text{tr}\{\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^\mu\} = \text{tr}\{\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\}\tag{10.135}$$

which is also on the left hand side. Bringing it to the left and dividing by 2 gives

$$\begin{aligned}\text{tr}\{\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\} &= \eta^{\sigma\mu} \text{tr}\{\gamma^\nu\gamma^\rho\} - \eta^{\rho\mu} \text{tr}\{\gamma^\nu\gamma^\sigma\} + \eta^{\mu\nu}\text{tr}\{\gamma^\rho\gamma^\sigma\} \\ &= 4(\eta^{\sigma\mu}\eta^{\nu\rho} - \eta^{\rho\mu}\eta^{\nu\sigma} + \eta^{\mu\nu}\eta^{\rho\sigma}).\end{aligned}\tag{10.136}$$

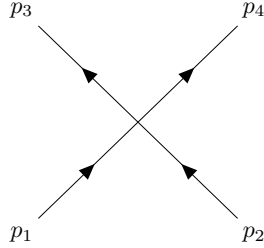
This is the result we were looking for. Clearly by using this trick we can in principle evaluate traces of an arbitrary number of gamma matrices.

Coming back to $e^-e^+ \rightarrow \mu^-\mu^+$ we find

$$\begin{aligned}\frac{1}{4} \sum_{\text{spins}} |\mathcal{T}|^2 &= \frac{4e^4}{s^2} [-p_1^\mu p_2^\nu - p_1^\nu p_2^\mu + (p_1 \cdot p_2 - m_e^2)\eta^{\mu\nu}] \\ &\quad \times [- (p_3)_\nu (p_4)_\mu - (p_3)_\mu (p_4)_\nu + (p_3 \cdot p_4 - m_\mu^2)\eta_{\mu\nu}] \\ &= \frac{8e^4}{s^2} [(p_1 \cdot p_4)(p_2 \cdot p_3) + (p_1 \cdot p_3)(p_2 \cdot p_4) - m_\mu^2(p_1 \cdot p_2) - m_e^2(p_3 \cdot p_4) + 2m_e^2 m_\mu^2]\end{aligned}\tag{10.137}$$

This looks already quite decent but it can be simplified even further in terms of Mandelstam variables.

Mandelstam Variables.



$$\tag{10.138}$$

$$\begin{aligned}s &= -(p_1 + p_2)^2 = -(p_3 + p_4)^2 \\ t &= -(p_1 - p_3)^2 = -(p_2 - p_4)^2 \\ u &= -(p_1 - p_4)^2 = -(p_2 - p_3)^2\end{aligned}\tag{10.139}$$

Together with the squares $p_1^2, p_2^2, p_3^2, p_4^2$, the Mandelstam variables can be used to express all Lorentz invariant bilinears in the momenta. Incoming and outgoing momenta are on-shell such that $p_1^2 + m_1^2 = 0$ etc. The sum of Mandelstam variables is

$$s + t + u = -(p_1^2 + p_2^2 + p_3^2 + p_4^2) = m_1^2 + m_2^2 + m_3^2 + m_4^2.\tag{10.140}$$

Using these variables for example through

$$p_1 \cdot p_4 = -\frac{1}{2} [(p_1 - p_4)^2 - p_1^2 - p_4^2] = \frac{1}{2} [u - m_e^2 + m_\mu^2],\tag{10.141}$$

one finds for $e^-e^+ \rightarrow \mu^-\mu^+$

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{T}|^2 = \frac{2e^4}{s^2} [t^2 + u^2 + 4s(m_e^2 + m_\mu^2) - 2(m_e^2 + m_\mu^2)^2].\tag{10.142}$$

From the squared matrix element we can calculate the differential cross section in the center of mass frame. For relativistic kinematics of $2 \rightarrow 2$ scattering and the normalization conventions we employ here one has in the center of mass frame

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s} \frac{|\vec{p}_3|}{|\vec{p}_1|} |\mathcal{T}|^2 = \frac{1}{64\pi^2 s} \frac{\vec{p}_3}{\vec{p}_1} \frac{1}{4} \sum_{\text{spins}} |\mathcal{T}|^2. \quad (10.143)$$

Let us express everything in terms of the energy E of the incoming particles and the angle θ between the incoming e^- electron momenta and outgoing μ^- muon.

$$\begin{aligned} |\vec{p}_1| &= \sqrt{E^2 - m_e^2} & s &= 4E^2, \\ |\vec{p}_3| &= \sqrt{E^2 - m_\mu^2} & t &= m_e^2 + m_\mu^2 - 2E^2 + 2\vec{p}_1 \cdot \vec{p}_3, \\ \vec{p}_1 \cdot \vec{p}_3 &= |\vec{p}_1| |\vec{p}_3| \cos \theta & u &= m_e^2 + m_\mu^2 - 2E^2 - 2\vec{p}_1 \cdot \vec{p}_3. \end{aligned} \quad (10.144)$$

With these relations we can express $\frac{d\sigma}{d\Omega}$ in terms of E and θ only. Let us concentrate on the ultrarelativistic limit $E \gg m_e, m_\mu$ so that we can set $m_e = m_\mu = 0$. One has then $|\vec{p}_1| = |\vec{p}_3|$ and

$$t^2 + u^2 = 8E^4(1 + \cos^2 \theta), \quad \frac{2(t^2 + u^2)}{s^2} = 1 + \cos^2 \theta, \quad (10.145)$$

which leads to

$$\frac{d\sigma}{d\Omega} = \frac{e^4}{64\pi^2 s} (1 + \cos^2 \theta) = \frac{\alpha^2}{4s} (1 + \cos^2 \theta). \quad (10.146)$$

In the last equation we used $\alpha = e^2/(4\pi)$.

Electron-Muon Scattering. We can also consider the scattering process $e^- \mu^- \rightarrow e^- \mu^-$,

$$k = q_1 - q_3 \quad (10.147)$$

$$i\mathcal{T} = \bar{u}(q_3)(-e\gamma^\mu)iu(q_1) \left(-i \frac{\eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}}{k^2} \right) \bar{u}(q_4)(-e\gamma^\nu)iu(q_2). \quad (10.148)$$

By a similar argument as before the term $\sim k_\mu k_\nu$ drops out,

$$\mathcal{T} = \frac{e^2}{(q_1 - q_3)^2} \bar{u}(q_3)\gamma^\mu u(q_1)\bar{u}(q_4)\gamma_\mu u(q_2) \quad (e^- \mu^- \rightarrow e^- \mu^-). \quad (10.149)$$

Compare this to what we have found for $e^- e^+ \rightarrow \mu^- \mu^+$

$$\mathcal{T} = \frac{e^2}{(p_1 + p_2)^2} \bar{v}(p_2)\gamma^\mu u(p_1)\bar{u}(p_3)\gamma_\mu v(p_4) \quad (10.150)$$

where the conventions were according to

$$(10.151)$$

There is a close relation and the expressions agree if we put

$$\begin{aligned}
q_1 &= +p_1, & u(q_1) &= u(p_1), \\
q_2 &= -p_4, & u(q_2) &= u(-p_4) \rightarrow v(p_4), \\
q_3 &= -p_2, & \bar{u}(q_3) &= \bar{u}(-p_2) \rightarrow \bar{v}(p_2), \\
q_4 &= +p_3, & \bar{u}(q_4) &= \bar{u}(p_3).
\end{aligned}
\tag{10.152}$$

This identification makes sense, recall that

$$(i\not{p} + m) u(p) = 0 \quad \text{but} \quad (-i\not{p} + m) v(p) = 0. \tag{10.153}$$

However one sign arises from the spin sums

$$\begin{aligned}
\sum_{s=1}^2 u_s(p) \bar{u}_s(p) &= -i\not{p} + m, \\
\sum_{s=1}^2 v_s(p) \bar{v}_s(p) &= -i\not{p} - m = -\sum_s u_s(-p) \bar{u}_s(-p).
\end{aligned}
\tag{10.154}$$

Because it appears twice, the additional sign cancels for $|\mathcal{T}|^2$ after spin averaging and one finds indeed the same result as for $e^-e^+ \rightarrow \mu^-\mu^+$ but with

$$\begin{aligned}
s_q &= -(q_1 + q_2)^2 = -(p_1 - p_4)^2 = u_p, \\
t_q &= -(q_1 - q_3)^2 = -(p_1 + p_2)^2 = s_p, \\
u_q &= -(q_1 - q_4)^2 = -(p_1 - p_3)^2 = t_p.
\end{aligned}
\tag{10.155}$$

We can take what we had calculated but must change the role of s , t and u ! This is an example of *crossing symmetries*.

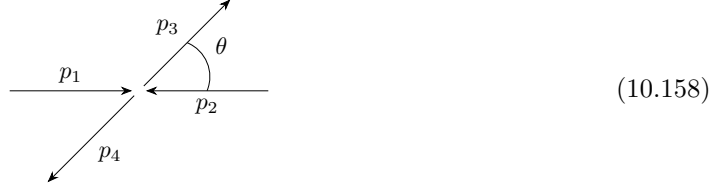
Recall that we found for $e^-e^+ \rightarrow \mu^-\mu^+$ in the massless limit $m_e = m_\mu = 0$ simply

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{T}|^2 = \frac{2e^4}{s^2} [t^2 + u^2]. \tag{10.156}$$

For $e^-\mu^- \rightarrow e^-\mu^-$ we find after the replacements $u \rightarrow s$, $s \rightarrow t$, $t \rightarrow u$,

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{T}|^2 = \frac{2e^4}{t^2} [u^2 + s^2]. \tag{10.157}$$

To get a better feeling for s , t and u , let us evaluate them in the center of mass frame for a situation where all particles have mass m .



$$(10.158)$$

$$\begin{aligned} p_1^\mu &= (E, \vec{p}), & p_2^\mu &= (E, -\vec{p}), \\ p_3^\mu &= (E, \vec{p}'), & p_4^\mu &= (E, -\vec{p}'). \end{aligned} \quad (10.159)$$

While s measures the center of mass energy, t is a momentum transfer that vanishes in the soft limit $\vec{p}^2 \rightarrow 0$ and in the colinear limit $\theta \rightarrow 0$. Similarly, u vanishes for $\vec{p}^2 \rightarrow 0$ and for backward scattering $\theta \rightarrow \pi$.

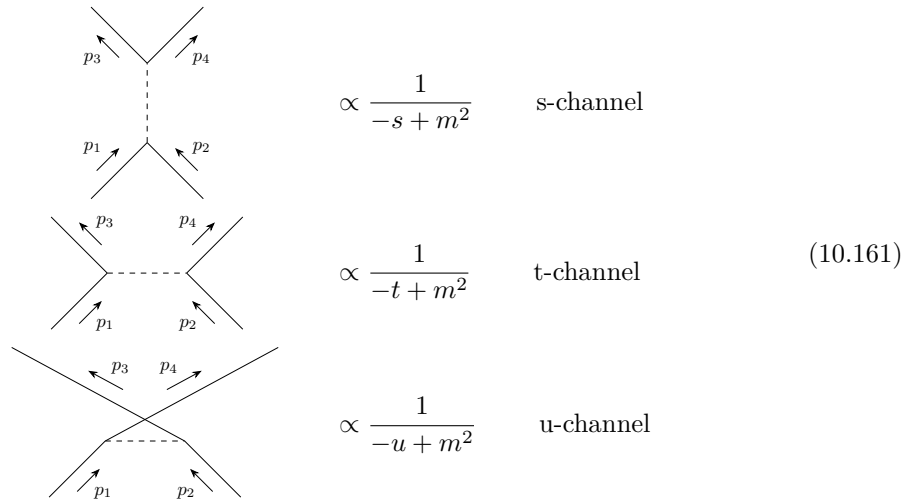
For the cross section we find for $e^- \mu^- \rightarrow e^- \mu^-$ in the massless limit

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s} \frac{1}{4} \sum_{spins} |\mathcal{T}|^2 = \frac{\alpha^2 [4 + (1 + \cos\theta)^2]}{2s(1 - \cos\theta)^2} \quad (10.160)$$

This diverges in the colinear limit $\theta \rightarrow 0$ as we had already seen for Yukawa theory in the limit where the exchange particle becomes massless.

Note that by the definition $s \geq 0$ while u and t can have either sign. Replacements of the type used for crossing symmetry are in this sense always to be understood as analytic continuation.

s -, t - and u -channels. One speaks of interactions in different channels for tree diagrams of the following generic types,



10.4 Relativistic scattering and decay kinematics

Covariant normalization of asymptotic states. For non-relativistic physics this we have used a normalization of single particle states in the asymptotic incoming and out-going regimes such that

$$\langle \vec{p} | \vec{q} \rangle = (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}). \quad (10.162)$$

For relativistic physics this has the drawback that it is not Lorentz invariant. To see this let us consider a boost in z -direction

$$\begin{aligned} E' &= \gamma(E + \beta p^3), \\ p^{1'} &= p^1, \\ p^{2'} &= p^2, \\ p^{3'} &= \gamma(p^3 + \beta E). \end{aligned} \tag{10.163}$$

Using the identity

$$\delta(f(x) - f(x_0)) = \frac{1}{|f'(x_0)|} \delta(x - x_0), \tag{10.164}$$

one finds

$$\begin{aligned} \delta^{(3)}(\vec{p} - \vec{q}) &= \delta^{(3)}(\vec{p}' - \vec{q}') \frac{dp^{3'}}{dp^3} = \delta^{(3)}(\vec{p}' - \vec{q}') \gamma \left(1 + \beta \frac{dE}{dp^3} \right) \\ &= \delta^{(3)}(\vec{p}' - \vec{q}') \frac{1}{E} \gamma (E + \beta p^3) \\ &= \frac{E'}{E} \delta^{(3)}(\vec{p}' - \vec{q}'). \end{aligned} \tag{10.165}$$

This shows, however, that $E \delta^{(3)}(\vec{p} - \vec{q})$ is in fact Lorentz invariant. This motivates to change the normalization such that

$$|p; \text{in}\rangle = \sqrt{2E_p} a_p^\dagger(-\infty) |0\rangle = \sqrt{2E_{\vec{p}}} |\vec{p}; \text{in}\rangle. \tag{10.166}$$

Note the subtle difference in notation between $|p; \text{in}\rangle$ (relativistic normalization) and $|\vec{p}; \text{in}\rangle$ (non-relativistic normalization). This implies for example

$$\langle p; \text{in} | q; \text{in} \rangle = 2E_p (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}). \tag{10.167}$$

With this normalization we must divide by $2E_p$ at the same places. In particular the completeness relation for single particle incoming states is

$$\mathbb{1}_{1\text{-particle}} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} |p; \text{in}\rangle \langle p; \text{in}|. \tag{10.168}$$

In fact, what appears here is a Lorentz invariant momentum measure. To see this consider

$$\int \frac{d^4p}{(2\pi)^4} (2\pi) \delta(p^2 + m^2) \theta(p^0) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}}. \tag{10.169}$$

The left hand side is explicitly Lorentz invariant and so is the right hand side.

Covariantly normalized S-matrix. We can use the covariant normalization of states also in the definition of S-matrix elements. The general definition is as before

$$S_{\beta\alpha} = \langle \beta; \text{out} | \alpha; \text{in} \rangle = \delta_{\beta\alpha} + i \mathcal{T}_{\beta\alpha} (2\pi)^4 \delta^{(4)}(p^{\text{in}} - p^{\text{out}}). \tag{10.170}$$

But now we take elements with relativistic normalization, e.g. for $2 \rightarrow 2$ scattering

$$S_{q_1 q_2, p_1 p_2} = \langle q_1, q_2; \text{out} | p_1, p_2; \text{in} \rangle. \tag{10.171}$$

We can calculate these matrix elements as before using the LSZ reduction formula to replace $\sqrt{2E_p} a_p^\dagger(-\infty)$ by fields. For example, for relativistic scalar fields

$$\sqrt{2E_{\vec{p}}} a_{\vec{p}}^\dagger(-\infty) = \sqrt{2E_{\vec{p}}} a_{\vec{p}}^\dagger(\infty) + i [-(p^0)^2 + \vec{p}^2 + m^2] \phi^*(p). \tag{10.172}$$

This allows to calculate S-matrix elements through correlation functions.

Cross sections for $2 \rightarrow n$ scattering. Let us now generalize our discussion of $2 \rightarrow 2$ scattering of non-relativistic particles to a scattering $2 \rightarrow n$ of relativistic particles. The transition probability is as before

$$P = \frac{|\langle \beta; \text{out} | \alpha; \text{in} \rangle|^2}{\langle \beta; \text{out} | \beta; \text{out} \rangle \langle \alpha; \text{in} | \alpha; \text{in} \rangle}. \quad (10.173)$$

Rewriting the numerator in terms of $\mathcal{T}_{\beta\alpha}$ and going over to the transition rate we obtain as before

$$\dot{P} = \frac{V(2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}}) |\mathcal{T}|^2}{\langle \beta; \text{out} | \beta; \text{out} \rangle \langle \alpha; \text{in} | \alpha; \text{in} \rangle}. \quad (10.174)$$

But now states are normalized in a covariant way

$$\begin{aligned} \langle p | q \rangle &= \lim_{q \rightarrow p} \langle p | q \rangle \\ &= \lim_{q \rightarrow p} 2E_p (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) \\ &= 2E_p (2\pi)^3 \delta^{(3)}(0) \\ &= 2E_p V \end{aligned} \quad (10.175)$$

One has thus for the incoming state of two particles

$$\langle \alpha; \text{in} | \alpha; \text{in} \rangle = 4E_1 E_2 V^2. \quad (10.176)$$

For the outgoing state of n particles one has instead

$$\langle \beta; \text{out} | \beta; \text{out} \rangle = \prod_{j=1}^n \{2q_j^0 V\}. \quad (10.177)$$

The product goes over final state particles which have the four-momentum q_j^n . So, far we have thus

$$\dot{P} = \frac{V(2\pi)^4 \delta^{(4)}(p^{\text{in}} - p^{\text{out}}) |\mathcal{T}|^2}{4E_1 E_2 V^2 \prod_{j=1}^n \{2q_j^0 V\}}. \quad (10.178)$$

To count final state momenta appropriately we could go back to finite volume and then take the continuum limit. This leads to an additional factor

$$\sum_{\vec{n}_j} \rightarrow V \int \frac{d^3 q}{(2\pi)^3} \quad (10.179)$$

for each final state particle. The transition rate becomes

$$\dot{P} = \frac{|\mathcal{T}|^2}{4E_1 E_2 V} \left[(2\pi)^4 \delta^{(4)}\left(p^{\text{in}} - \sum_j q_j\right) \prod_{j=1}^n \left\{ \frac{d^3 q_j}{(2\pi)^3 2q_j^0} \right\} \right] \quad (10.180)$$

The expression in square brackets is known as the *Lorentz-invariant phase space* measure (sometimes "LIPS"). To go from there to a differential cross section we need to divide by a flux of particles. There is one particle per volume V with velocity $v = v_1 - v_2$, so the flux is

$$\mathcal{F} = \frac{|v|}{V} = \frac{|v_1 - v_2|}{V} = \frac{\left| \frac{p_1^3}{p_1^0} - \frac{p_2^3}{p_2^0} \right|}{V}. \quad (10.181)$$

In the last equality we chose the beam axis to coincide with the z -axis. For the differential cross section we obtain

$$d\sigma = \frac{|\mathcal{T}|^2}{4E_1 E_2 |v_1 - v_2|} [\text{LIPS}]. \quad (10.182)$$

The expression in the prefactor can be rewritten like

$$\frac{1}{E_1 E_2 |v_1 - v_2|} = \frac{1}{p_1^0 p_2^0 \left| \frac{p_1^3}{p_1^0} - \frac{p_2^3}{p_2^0} \right|} = \frac{1}{|p_2^0 p_1^3 - p_1^0 p_2^3|} = \frac{1}{|\epsilon_{\mu xy\nu} p_2^\mu p_1^\nu|}. \quad (10.183)$$

This is not Lorentz invariant in general but invariant under boosts in the z -direction. In fact it transforms as a two-dimensional area element as it should. In the center of mass frame one has $p_2^3 = -p_1^3 = \pm |\vec{p}_1|$ and

$$\frac{1}{|p_2^0 p_1^3 - p_1^0 p_2^3|} = \frac{1}{|\vec{p}_1| (p_1^0 + p_2^0)} = \frac{1}{|\vec{p}_1|_{\text{COM}} \sqrt{s}} \quad (10.184)$$

This leads finally to the result for the differential cross section

$$d\sigma = \frac{|\mathcal{T}|^2}{4|\vec{p}_1|_{\text{COM}} \sqrt{s}} \left[(2\pi)^4 \delta^{(4)} \left(p^{\text{in}} - \sum_j q_j \right) \prod_{j=1}^n \left\{ \frac{d^3 q_j}{(2\pi)^3 2q_j^0} \right\} \right]. \quad (10.185)$$

2 → 2 scattering. For the case of $n = 2$ one can write the Lorentz invariant differential phase space element in the center of mass frame (exercise)

$$\left[(2\pi)^4 \delta^{(4)} (p^{\text{in}} - q_1 - q_2) \frac{d^3 q_1}{(2\pi)^3 2q_1^0} \frac{d^3 q_2}{(2\pi)^3 2q_2^0} \right] = \frac{|\vec{q}_1|}{16\pi^2 \sqrt{s}} d\Omega \quad (10.186)$$

such that

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s} \frac{|\vec{q}_1|}{|\vec{p}_1|} |\mathcal{T}|^2. \quad (10.187)$$

Decay rate. Let us now consider the decay rate of a single particle, i. e. a process $1 \rightarrow n$. We can still use equation (10.174), but now the initial state is normalized like

$$\langle \alpha; \text{in} | \alpha; \text{in} \rangle = 2E_1 V. \quad (10.188)$$

We find then for the differential transition or decay rate $d\Gamma = \dot{P}$

$$d\Gamma = \frac{|\mathcal{T}|^2}{2E_1} \left[(2\pi)^4 \delta^{(4)} \left(p^{\text{in}} - \sum_j q_j \right) \prod_{j=1}^n \left\{ \frac{d^3 q_j}{(2\pi)^3 2q_j^0} \right\} \right] \quad (10.189)$$

In the center of mass frame one has $E_1 = m_1$. For the special case of $1 \rightarrow 2$ decay one finds in the center of mass frame

$$d\Gamma = \frac{|\mathcal{T}|^2 |\vec{q}_1|}{32\pi^2 m_1^2} d\Omega. \quad (10.190)$$

10.5 Higgs/Yukawa theory

Let us consider the following extension of QED by a neutral scalar field (with $m = gv$)

$$S[\bar{\psi}, \psi, A, \phi] = \int_x \left\{ -\bar{\psi} \gamma^\mu (\partial_\mu - ieA_\mu) \psi - im\bar{\psi}\psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{1}{2} \phi (-\partial_\mu \partial^\mu + M^2) \phi - ig\phi \bar{\psi}\psi \right\}. \quad (10.191)$$

Note that a constant (homogeneous) scalar field ϕ modifies the fermion mass according to

$$m_{\text{eff}} = m + g\phi = g(v + \phi) \quad (10.192)$$

In fact, one can understand the masses of elementary fermions (leptons and quarks) in the standard model of elementary particle physics as being due to such a scalar field expectation value for the

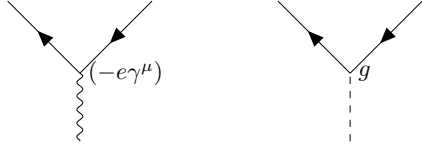
Higgs field. In the theory above we have now different propagators

$$\begin{array}{ll}
\text{~~~~~} & \frac{1}{i}\Delta_{\mu\nu}(x-y) \\
\text{-----} & \frac{1}{i}S_{\alpha\beta}(x-y) \\
\text{-----} & \frac{1}{i}\Delta(x-y)
\end{array} \tag{10.193}$$

with scalar propagator

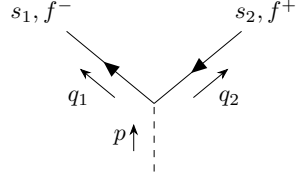
$$\Delta(x-y) = \int_p e^{ip(x-y)} \frac{1}{p^2 + M^2}. \tag{10.194}$$

The vertices are



$$\tag{10.195}$$

Higgs decay to fermions. Let us discuss first the process $\phi \rightarrow f^- f^+$. The fermions could be leptons (e, μ, τ) or quarks (u, d, s, c, b, t). The Feynman diagram for the decay is simply



$$\tag{10.196}$$

According to the Feynman rules we obtain

$$i\mathcal{T} = g \bar{u}_{s_1}(q_1) i v_{s_2}(q_2), \quad \mathcal{T}^* = g \bar{v}_{s_2}(q_2) u_{s_1}(q_1). \tag{10.197}$$

For the absolute square one finds

$$|\mathcal{T}|^2 = g^2 \bar{u}_{s_1}(q_1) v_{s_2}(q_2) \bar{v}_{s_2}(q_2) u_{s_1}(q_1). \tag{10.198}$$

We will assume that the final spins are not observed and sum them

$$\sum_{\text{spins}} |\mathcal{T}|^2 = g^2 \text{tr} \left\{ (-i\not{q}_2 - m)(-i\not{q}_1 + m) \right\} \tag{10.199}$$

We used here again the spin sum formula

$$\sum_s v_s(p) \bar{v}_s(p) = -i\not{p} - m, \quad \sum_s u_s(p) \bar{u}_s(p) = -i\not{p} + m. \tag{10.200}$$

Performing also the Dirac traces gives

$$\sum_{\text{spins}} |\mathcal{T}|^2 = g^2 (-4q_1 \cdot q_2 - 4m^2). \tag{10.201}$$

Let us now go into the rest frame of the decaying particle where

$$p = (M, 0, 0, 0), \quad q_1 = \left(\frac{M}{2}, \vec{q}\right), \quad q_2 = \left(\frac{M}{2}, -\vec{q}\right), \tag{10.202}$$

with

$$\vec{q}^2 = -m^2 + \frac{M^2}{4}, \quad q_1 \cdot q_2 = -\frac{M^2}{4} - \vec{q}^2 = -\frac{M^2}{m^2}, \tag{10.203}$$

and

$$\sum_{\text{spins}} |\mathcal{T}|^2 = 2g^2 M^2 \left(1 - 4\frac{m^2}{M^2}\right). \quad (10.204)$$

Note that the decay is kinematically possible only for $M > 2m$ so that the bracket is always positive. For the particle decay rate we get

$$\frac{d\Gamma}{d\Omega} = \frac{|\vec{q}_1|}{32\pi^2 M^2} \sum_{\text{spins}} |\mathcal{T}|^2 = \frac{g^2 M}{32\pi^2} \left(1 - 4\frac{m^2}{M^2}\right)^{3/2}. \quad (10.205)$$

Because this is independent of the solid angle Ω one can easily integrate to obtain the decay rate

$$\Gamma = \frac{g^2 M}{8\pi} \left(1 - 4\frac{m^2}{M^2}\right)^{3/2}. \quad (10.206)$$

If the scalar boson ϕ is the Higgs boson, the Yukawa coupling is in fact proportional to the fermion mass m ,

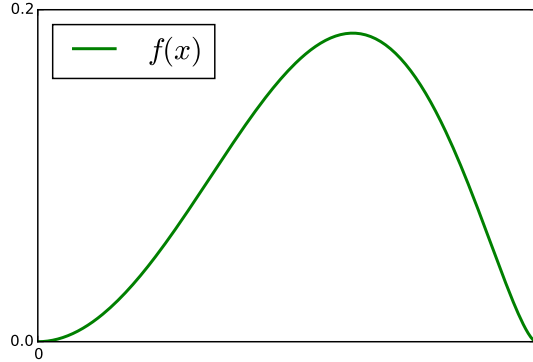
$$g = \frac{m}{V}. \quad (10.207)$$

One has then

$$\Gamma = \frac{M^3}{32\pi v^2} f\left(\frac{2m}{M}\right) \quad (10.208)$$

where

$$f(x) = x^2(1 - x^2)^{3/2} \quad (10.209)$$



Decay into light fermions is suppressed because of small coupling while decay into very heavy fermions is suppressed by small phase space or even kinematically excluded for $2m > M$.

For Higgs boson mass of $M = 125$ GeV the largest decay rate to fermions is to $b\bar{b}$ (bottom quark and anti-quark). This corresponds to $m = 4.18$ GeV. The top quark would have larger coupling but is in fact too massive ($m = 172$ GeV). (The lepton with largest mass is the tauon τ with $m = 1.78$ GeV.)

Higgs decay into photons. A Higgs particle can also decay into photons and this is in fact how it was discovered. How is this possible? If we try to write down a diagram in the theory introduced above we realize that there is no tree diagram. However, there are loop diagrams!

Consider

$$(10.210)$$

These terms arise from the expansion of the partition function if the fermion propagator appears 3 times and there are 2 fermion-photon and one fermion-scalar vertices. Schematically, the vertices are derivatives

$$\left[(-e\gamma) \left(\frac{1}{i} \frac{\delta}{\delta J^\mu} \right) \left(i \frac{\delta}{\delta \eta} \right) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}} \right) \right] \quad \text{or} \quad \left[g \left(\frac{1}{i} \frac{\delta}{\delta J} \right) \left(i \frac{\delta}{\delta \eta} \right) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}} \right) \right] \quad (10.211)$$

and they act here on a chain like

$$\left[(i\bar{\eta}) \left(\frac{1}{i} S \right) (i\eta) \right] \left[(i\bar{\eta}) \left(\frac{1}{i} S \right) (i\eta) \right] \left[(i\bar{\eta}) \left(\frac{1}{i} S \right) (i\eta) \right]. \quad (10.212)$$

Note that the derivative with respect to $\bar{\eta}$ can be commuted through the square brackets and acts on $\bar{\eta}$ from the left. Factors $1/i$ and i cancel. The derivative with respect to η receives an additional minus sign from commuting and this cancels against i^2 . In this way the vertices can connect the elements of the chain. However, for a closed loop also the beginning and end of the chain must be connected. To make this work, one can first bring the $(i\eta)$ from the end of the chain to its beginning. This leads to one additional minus sign from anti-commuting Grassmann fields. This shows that *closed fermion lines have one more minus sign*.

In position space and including sources, the first diagram is

$$(10.213)$$

$$g(-1) \int_{x,y,z} \text{tr} \left\{ \left[\frac{1}{i} S(x-y) \right] (-e\gamma^\mu) \left[\frac{1}{i} S(y-z) \right] (-e\gamma^\nu) \left[\frac{1}{i} S(z-x) \right] \right\} \\ \times \int_{u,v,w} \left[\frac{1}{i} \Delta_{\mu\alpha}(y-u) \right] (iJ^\alpha(u)) \left[\frac{1}{i} \Delta_{\nu\beta}(z-v) \right] (iJ^\beta(v)) \left[\frac{1}{i} \Delta(x-w) \right] (iJ(w)) \quad (10.214)$$

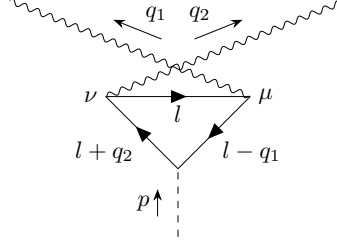
The trace is for the Dirac matrix indices. If one translates this now to momentum space and considers the amputated diagram for an S-matrix element, one finds that momentum conservation constrains momenta only up to one free integration momentum or loop momentum.

In fact, more generally, there is one integration momentum for every closed loop. The first diagram is then

$$(10.215)$$

$$\begin{aligned}
& (-1)g e^2 \epsilon_\mu^*(q_1) \epsilon_\nu^*(q_2) \int_l \frac{1}{[l+q_1]^2 + m^2 - i\epsilon} [l^2 + m^2 - i\epsilon] [l-q_2]^2 + m^2 + i\epsilon} \\
& \times \text{tr} \left\{ \left[-i(l + \not{q}_1) + m \right] \gamma^\mu \left[-i\not{l} + m \right] \gamma^\nu \left[-i(l - \not{q}_2) + m \right] \right\}
\end{aligned} \tag{10.216}$$

For the second diagram we can write



$$\tag{10.217}$$

$$(-1)g e^2 \epsilon_\mu^*(q_1) \epsilon_\nu^*(q_2) \int_l \dots \tag{10.218}$$

where the integrand is the same up to the interchange $q_1 \leftrightarrow q_2$ and $\mu \leftrightarrow \nu$. We can therefore concentrate on evaluating the first diagram. We use there the abbreviation

$$\int_l = \int \frac{d^4 l}{(2\pi)^4}. \tag{10.219}$$

The Feynman $i\epsilon$ terms allow to perform a Wick rotation to Euclidean space $l^0 = i\tilde{l}_E^0$ so that l^2 is then positive. Let us count powers of l . First, in the Dirac trace we have terms with up to 5 gamma matrices. However, only traces of an even number of gamma matrices are non-zero.

With a bit of algebra one finds for the Dirac trace

$$\begin{aligned}
& \text{tr} \left\{ \left[-i(l + \not{q}_1) + m \right] \gamma^\mu \left[-i\not{l} + m \right] \gamma^\nu \left[-i(l - \not{q}_2) + m \right] \right\} \\
& = -m \text{tr} \left\{ (l + \not{q}_1) \gamma^\mu \not{l} \gamma^\nu + (l + \not{q}_1) \gamma^\mu \gamma^\nu (l - \not{q}_2) + \gamma^\mu \not{l} \gamma^\nu (l - \not{q}_2) \right\} + m^3 \text{tr} \{ \gamma^\mu \gamma^\nu \} \\
& = -4m \left[(l + q_1)^\mu l^\nu + (l + q_1)^\nu l^\mu - (l + q_1) \cdot l \eta^{\mu\nu} \right. \\
& \quad \left. + (l + q_1)^\mu (l - q_2)^\nu + (l + q_1) \cdot (l - q_2) \eta^{\mu\nu} - (l + q_1)^\nu (l - q_2)^\mu \right. \\
& \quad \left. + l^\mu (l - q_2)^\nu + (l - q_2)^\mu l^\nu - \eta^{\mu\nu} l \cdot (l - q_2) \right] + 4\eta^{\mu\nu} m^3 \\
& = -4m \left[4l^\mu l^\nu - l^2 \eta^{\mu\nu} - l^2 \eta^{\mu\nu} + 2q_1^\mu l^\nu - 2q_2^\nu l^\mu - q_1^\mu q_2^\mu + q_1^\nu q_2^\nu - (q_1 \cdot q_2) \eta^{\mu\nu} \right] + 4\eta^{\mu\nu} m^3
\end{aligned} \tag{10.220}$$

Let us now consider the denominator. One can introduce so-called Feynman parameters to write

$$\begin{aligned}
& \frac{1}{[(l+q_1)^2 + m^2][l^2 + m^2][(l-q_2)^2 + m^2]} \\
& = 2! \int_0^1 du_1 \dots du_3 \delta(u_1 + u_2 + u_3 - 1) \frac{1}{[u_1[(l+q_1)^2 + m^2] + u_2[l^2 + m^2] + u_3[(l-q_2)^2 + m^2]]^3} \\
& = 2 \int_0^1 du_1 \dots du_3 \frac{\delta(u_1 + u_2 + u_3 - 1)}{[l^2 + 2l(u_1 q_1 - u_3 q_2) + u_1 q_1^2 + u_3 q_2^2 + m^2]^3}.
\end{aligned} \tag{10.221}$$

We have used here the identity (will be proven in the second part of the course QFT 2)

$$\frac{1}{p_1 \dots p_n} = (n-1)! \int_0^1 du_1 \dots du_n \frac{\delta(u_1 + \dots + u_n - 1)}{[u_1 A_1 + \dots + u_n A_n]^n} \tag{10.222}$$

In a next step one commutes the integral over $u_1 \dots u_3$ with the integral over l . It is useful to change integration variables according to

$$\begin{aligned} l + u_1 q_1 - u_3 q_2 &\rightarrow k, \\ l &= k - u_1 q_1 + u_3 q_2. \end{aligned} \quad (10.223)$$

Collecting terms we find for the first diagram

$$(-1)g e^2 \epsilon_\mu^*(q_1) \epsilon^*(q_2) 2 \int_0^1 du_1 \dots du_3 \delta(u_1 + u_2 + u_3 - 1) \int \frac{d^4 k}{(2\pi)^4} \frac{A^{\mu\nu}}{[k^2 + u_1 q_1^2 + u_3 q_2^2 - (u_1 q_1 - u_3 q_2)^2 + m^2]^3} \quad (10.224)$$

where

$$\begin{aligned} A^{\mu\nu} &= -4m \left[4k^\mu k^\nu - k^2 \eta^{\mu\nu} + \text{terms linear in } k \right. \\ &\quad + 4(u_1 q_1 - u_3 q_2)^\mu (u_1 q_1 - u_3 q_2)^\nu - (u_1 q_1 - u_3 q_2)^2 \eta^{\mu\nu} \\ &\quad \left. - q_1^\mu q_2^\nu + q_1^\nu q_2^\mu - (q_1 \cdot q_2) \eta^{\mu\nu} - \eta^{\mu\nu} m^2 \right]. \end{aligned} \quad (10.225)$$

The integral over k is now symmetric around the origin. There is no contribution from linear terms in k and also the quadratic terms cancels. In fact, one can prove that

$$\lim_{d \rightarrow 4} \int \frac{d^d k}{(2\pi)^d} \frac{4k^\mu k^\nu - (k^2 + A) \eta^{\mu\nu}}{(k^2 + A)^3} = 0. \quad (10.226)$$

We will develop the techniques to prove this in QFT2.

Taking this as well as $\epsilon_\mu^*(q_1) q_1^\mu = \epsilon_\nu^*(q_2) q_2^\nu = 0$ and $q_1^2 = q_2^2 = 0$ into account leads to

$$A^{\mu\nu} = -4m [1 - 4u_1 u_2] [q_1^\mu q_2^\nu - (q_1 \cdot q_2) \eta^{\mu\nu}]. \quad (10.227)$$

Note that this is symmetric with respect to $(q_1, \mu) \leftrightarrow (q_2, \nu)$, so we can add the second diagram by multiplying with 2. We obtain

$$\begin{aligned} i\mathcal{T} &= 8g e^2 m \epsilon_\mu^*(q_1) \epsilon_\nu^*(q_2) [q_1^\nu q_2^\mu - (q_1 \cdot q_2) \eta^{\mu\nu}] \\ &\quad \times 2 \int_0^1 du_1 \dots du_3 \delta(u_1 + u_2 + u_3 - 1) [1 - 4u_1 u_3] \int \frac{d^4 k}{(2\pi)^4} \frac{1}{[k^2 + 2u_1 u_3 q_1 \cdot q_2 + m^2]^3} \end{aligned} \quad (10.228)$$

To evaluate the integral over k we note that in the rest frame of the decaying scalar boson $p = q_1 + q_2 = (M, 0, 0, 0)$ such that $p^2 = 2q_1 \cdot q_2 = -M^2$. If we concentrate on fermions that are very heavy such that $m \gg M$ we can expand in the term $u_1 u_3 q_1 \cdot q_2$ in the integral over k . One finds to lowest order

$$\int \frac{d^4 k}{(2\pi)^4} \frac{1}{[k^2 + m^2]^3} = i \frac{1}{(4\pi)^2} \frac{1}{2m^2}. \quad (10.229)$$

This i is due to the Wick rotation $k^0 = i k_E^0$. Also the integral over Feynman parameters can now easily be performed

$$\begin{aligned} &2 \int_0^1 du_1 \dots du_3 \delta(u_1 + u_2 + u_3 - 1) [1 - 4u_1 u_3] \\ &= 2 \int_0^1 du_1 du_3 \theta(1 - u_1 - u_3) [1 - 4u_1 u_3] \\ &= 2 \int_0^1 du_1 \int_0^{1-u_1} du_3 [1 - 4u_1 u_3] \\ &= 2 \int_0^1 du_1 [(1 - u_1) - 4u_1 \frac{1}{2} (1 - u_1)^2] \\ &= 2 - 3 + \frac{8}{3} - 1 = \frac{2}{3}. \end{aligned} \quad (10.230)$$

Collecting terms we find

$$i\mathcal{T} = i \frac{8ge^2}{3(4\pi)^2 m} \epsilon_\mu^*(q_1) \epsilon_\nu^*(q_2) [q_1^\nu q_2^\mu - (q_1 \cdot q_2) \eta^{\mu\nu}]. \quad (10.231)$$

Photon polarization sums and Ward identity. Before we continue we need to develop a method to perform the spin sums for photons. In the squared amplitude expressions like the following appear

$$\sum_{\text{polarizations}} |\mathcal{T}|^2 = \sum_{\text{polarizations}} \epsilon_\mu^*(q) \epsilon_\nu(q) M^\mu(q) M^{\nu*}(q) \quad (10.232)$$

We have extended here the polarization vector of a photon from the amplitude by decomposing

$$\tau = \epsilon_\mu^*(q) M^\mu(q). \quad (10.233)$$

Let us choose without loss of generality $q^\mu = (E, 0, 0, E)$ and use the polarization vector introduced previously.

$$\epsilon_\mu^{(1)} = (0, \frac{1}{\sqrt{2}}, -\frac{i}{\sqrt{2}}, 0) \quad (10.234)$$

$$\epsilon_\mu^{(2)} = (0, \frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0) \quad (10.235)$$

such that

$$\epsilon_\mu^{*(1)} \epsilon_\nu^{(1)} + \epsilon_\mu^{*(2)} \epsilon_\nu^{(2)} = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix} \quad (10.236)$$

This would give

$$\sum_{j=1}^2 \epsilon_\mu^{*(j)} \epsilon_\nu^{(j)} M^\mu M^{\nu*} = |M^1|^2 + |M^2|^2. \quad (10.237)$$

To simplify this one can use an identity we will prove later,

$$q_\mu M^\mu(q) = 0. \quad (10.238)$$

This is in fact a consequence of gauge symmetry known as *Ward identity*. For the above choice of q^μ it follows

$$-M^0 + M^3 = 0 \quad (10.239)$$

Accordingly, one can add $0 = -|M^0|^2 + |M^3|^2$ to the spin sum

$$\sum_{j=1}^2 \epsilon_\mu^{*(j)} \epsilon_\nu^{(j)} M^\mu M^{\nu*} = -|M^0|^2 + |M^1|^2 + |M^2|^2 + |M^3|^2 = \eta_{\mu\nu} M^\mu M^{\nu*}. \quad (10.240)$$

In this sense we can use for external photons

$$\sum_{j=1}^2 \epsilon_\mu^{*(j)} \epsilon_\nu^{(j)} \rightarrow \eta_{\mu\nu} \quad (10.241)$$

With this we can now calculate the sums over final state photon polarizations

$$\begin{aligned} \sum_{\text{pol.}} |\tau|^2 &= \left(\frac{8ge^2}{3(4\pi)^2 m} \right)^2 [q_1^\nu q_2^\mu - (q_1 \cdot q_2) \eta^{\mu\nu}] [q_1^\beta q_2^\alpha - (q_1 \cdot q_2) \eta^{\alpha\beta}] \\ &\quad \sum_{\text{pol.}} \epsilon_\mu^*(q_1) \epsilon_\alpha(q_1) \sum_{\text{pol.}} \epsilon_\nu^*(q_2) \epsilon_\beta(q_2) \\ &= \left(\frac{8ge^2}{3(4\pi)^2 m} \right)^2 2(q_1 \cdot q_2)^2 = \frac{2g^2 \alpha^2}{9\pi^2 m^2} \pi^4 \end{aligned} \quad (10.242)$$

For the particle decay rate $\varphi \rightarrow \gamma\gamma$ this gives with $|\vec{q}_1| = \frac{M}{2}$

$$\frac{d\Gamma}{d\Omega} = \frac{|\vec{q}_1|}{32\pi^2 M^2} \sum_{\text{pol.}} |\tau|^2 = \frac{g^2 \alpha^2}{9.32\pi^4 m^2} M^3 \quad (10.243)$$

Finally, we integrate over solid angle $\Omega = \frac{1}{2}4\pi$ where the factor $\frac{1}{2}$ is due to the fact that the photons in the final state are indistinguishable. The decay rate for $\varphi \rightarrow \gamma\gamma$ through a heavy fermion loop is finally

$$\Gamma = \frac{g^2 \alpha^2}{144\pi^3 m^2} M^3 \quad (10.244)$$

Note that because of $g = \frac{m}{V}$ this is in fact independent of the heavy fermion mass m .