Lectures on quantum field theory 1

Stefan Floerchinger

Theoretisch-Physikalisches Institut, Max-Wien-Platz 1, 07743 Jena, Germany

E-mail: stefan.floerchinger@uni-jena.de

Abstract: Notes for lectures that introduce students of physics to quantum field theory with applications to high energy physics, condensed matter and statistical physics. Prepared for a course at Friedrich-Schiller-University Jena in the summer term 2024.

Contents

Literature

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

Relativistic quantum field theory

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

Statistical field theory / renormalization group

- Jean Zinn-Justin, *Quantum field theory and critical phenomena*
- Andreas Wipf, *Statistical approach to quantum field theory*
- John Cardy, *Scaling and renormalization in statistical physics*
- Giorgio Parisi, *Statistical field theory*

Non-relativistic quantum field theory / condensed matter

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

Typos

Please send typos to stefan.floerchinger@uni-jena.de.

1 Introduction

1.1 What is quantum field theory?

Quantum field theory is the modern theoretical framework to describe almost all phenomena in fundamental physics. This includes the standard model of elementary particle physics with the electromagnetic, the weak and the strong force and most likely, in one way or another, also dark matter and gravitation.

There are close connections to quantum mechanics and historically quantum field theory was developed as quantum theory for infinetly many degrees of freedom when it became clear that a relativistic version of quantum mechanics is not consistent. In the modern understanding quantum field theory in actually underlying non-relativistic quantum mechanics and the latter follows from the former in a limit. There is also a non-relativistic version of quantum field theory which can describe few-body physics of non-relativistic particles, but can also be used favorably to describe many-body physics and condensed matter.

Another very interesting connection is between quantum field theory and statistical field theory. Many of the concepts needed for relativistic quantum field theory can only be properly understood from the point of view of statistical physics and moreover, the same concepts can be used to describe stochastic theories where fluctuations are not of a quantum origin but have different reasons. This goes even beyond physics and the natural sciences.

Relativistic quantum field theories have also an interesting intersection with group theory, the theory of symmetries. Specifically Lie groups of various kinds play an important role to understand the phenomena of the standard model of elementary particle physics. Also consequences of space-time symmetries like conservation laws or the basic notion of what a particle actually is can be mentioned here.

There is also a very interesting relation to (quantum) information theory that is currently being explored in more detail. It is well possible that further insights into quantum field dynamics arise here in the coming years.

1.2 What concepts are needed to understand it?

- Quantum theory
- Symmetries and Lie group theory
- Concepts from statistical physics

2 Basics of Lie groups

2.1 Symmetries and groups

Symmetry transformations

Studying symmetries and their consequences is one of the most fruitful ideas in physics. This holds especially in high energy and particle physics but by far not only there. To get started, we first define the notion of a symmetry transformation and relate it to the mathematical concept of a group.

It is natural to characterize a symmetry transformation by the following properties

- One symmetry transformation followed by another should be a symmetry transformation itself.
- There should be a unique (trivial) symmetry transformation doing nothing.
- For each symmetry transformation there needs to be a unique symmetry transformation reversing it.

With these properties, the set of all symmetry transformations G forms a group in the mathematical sense.

Properties of groups

More formally, a group G has the following properties.

(i) *Closure*: For all elements $f, g \in G$ the composition $g \cdot f$ is in G. (We use here transformations acting to the right so that $g \cdot f$ should be read as a transformation where we apply first f and then g.)

(ii) *Associativity*: $h \cdot (g \cdot f) = (h \cdot g) \cdot f$.

(iii) *Identity element*: There exists a unique unit element $\mathbb{1}$ in the group, $\mathbb{1} \in G$, such that $\mathbb{1} \cdot f = f \cdot \mathbb{1} = f$ for all $f \in G$.

(iv) *Inverse element*: For all elements $f \in G$ there is a unique inverse $f^{-1} \in G$ such that $f \cdot f^{-1} = f^{-1} \cdot f = \mathbb{1}$.

These basic properties define groups of many kinds, both finite and infinite, discrete and continuous.

Representations

One distinguishes between groups as abstract entities and concrete *representations*. The abstract abstract group is defined through the set of its elements and composition law. A representation is a concrete realization of the group elements and their composition law, for example as matrices acting on a vector space or transformations of some type.

For example, a very simple group is $\mathbb{Z}_2$. It has two elements, the unit element $\mathbb{1}$ and an element R with $R^2 = \mathbb{1}$. A representation of R on the space of functions $f(x)$ of a single variable x could be given by the parity transform $f(x) = f(-x)$. The unit element is represented by the identity transform $f(x) = f(x)$, and we thus have a representation of the group $\mathbb{Z}_2$.

Abelian and non-abelian groups

A group is called *abelian* if the group product is commutative, $f \cdot g = g \cdot f$ for all $f, g \in G$. Otherwise the group is called *non-abelian*.

2.2 Examples for Lie groups

Lie groups can be defined as *differentiable manifolds* with a group structure. They have an infinite number of elements. Let us start with a few examples.

- $G = \mathbb{R}$, the *additive* group of real numbers. The group "product" is here the addition, the inverse of an element is its negative and the neutral or unit element is zero. This is clearly an abelian group.
- $G = \mathbb{R}_+^*$, the *multiplicative* group of positive real numbers. Also an abelian group.

- $G = \mathrm{GL}(n, \mathbb{R})$, the general linear group of real $n \times n$ matrices g with $\det(g) \neq 0$ (such that they are invertible). Similarly, $G = \mathrm{GL}(n, \mathbb{C})$, the general linear group of complex $n \times n$ matrices. These groups are non-abelian for $n > 1$.

- $G = \mathrm{SL}(n, \mathbb{R})$ the special linear group is a subgroup of $\mathrm{GL}(n, \mathbb{R})$ with $\det(g) = 1$. This is a more general notion, the S for *special* usually means $\det(g) = 1$.

- $G = \mathrm{O}(n)$, the *orthogonal* group of real $n \times n$ matrices R with $R^T R = \mathbb{1}$. This immediately implies $\det(R) = \pm 1$. Again this is a subgroup of $\mathrm{GL}(n, \mathbb{R})$. As a manifold, $\mathrm{O}(n)$ is not connected. One component is the subgroup $\mathrm{SO}(n)$ with $\det(R) = 1$, the other is a separate submanifold where $\det(R) = -1$. One can understand $\mathrm{O}(n)$ as the group of *rotations* and *reflections* in the n-dimensional Euclidean space. The simplest non-trivial case is for $n = 2$ where SO(2) consists of elements of the form

  $$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

  This is clearly isomorphic to the group U(1) of complex phases $e^{i\theta}$. $\mathrm{SO}(n)$ is non-abelian for $n > 2$.

- $G = \mathrm{U}(n)$, the *unitary* group of complex $n \times n$ matrices U with $U^\dagger U = \mathbb{1}$. Now we immediately infer that $\det(U)$ is a complex number with absolute value 1. $\mathrm{U}(n)$ is non-abelian for $n > 1$.

- $G = \mathrm{SU}(n)$, the special unitary group with unit determinant. Plays an important role in physics, most importantly SU(2) and SU(3).

- $G = \mathrm{O}(r, n-r)$ the indefinite orthogonal group of $n \times n$ matrices R that leaves the metric $\eta = \mathrm{diag}(-1, \ldots, -1, +1, \ldots, +1)$ with r entries -1 and $n-r$ entries $+1$ invariant, in the sense that $R^T \eta R = \eta$. An example is $\mathrm{O}(1, 3)$, the group of Lorentz transformations in $d = 1 + 3$ dimensions.

- $G = \mathrm{Sp}(2n, \mathbb{R})$ is the symplectic group of $2n \times 2n$ matrices M that leaves a symplectic bilinear form

  $$\Omega = \begin{pmatrix} 0 & +\mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix} \tag{2.1}$$

  invariant in the sense that $M^T \Omega M = \Omega$. Here $\mathbb{1}_n$ is the n dimensional unit matrix and similarly 0. This is obviously a subgroup of $\mathrm{GL}(2n, \mathbb{R})$. There is also a complex version $\mathrm{Sp}(2n, \mathbb{C})$.

Lie groups have very nice features and a rich mathematical structure because they are both, *groups* and *manifolds*. We will now first introduce Lie groups and Lie algebras from an algebraic point of view, and subsequently also briefly introduce a differential-geometric characterization.

LECTURE 02

2.3 Algebraic approach to Lie groups and Lie algebras

Lie algebra and generators

Because a Lie group is also a manifold, group elements can be labeled by a (usually multi-dimensional) parameter or coordinate $\xi = (\xi^1, \ldots, \xi^m)$, i.e. we can write them as $g(\xi)$. Without

loss of generality we can assume that $\xi = 0$ corresponds to the unit element, $g(0) = \mathbb{1}$. Let us now consider infinitesimal transformations. We can write them as

$$g(d\xi) = \mathbb{1} + id\xi^j T_j + \dots, \tag{2.2}$$

where we use Einsteins summation convention implying a sum over the index j, and the ellipses stand for terms of quadratic and higher order in the infinitesimal $d\xi$. Note that we can write

$$iT_j = \frac{\partial}{\partial \xi^j} g(\xi)\Big|_{\xi=0}. \tag{2.3}$$

Formally, the objects iT_j constitute a basis of the tangent space of the Lie group manifold at the position of the unit element $g(\xi) = \mathbb{1}$, which is at $\xi = 0$. The factor i is conventional and used by physicists, while mathematicians usually work in a convention without it. The T_j are also known as the generators of the *Lie algebra*, to which we turn in a moment. The generators constitute a basis such that any element of the Lie algebra can be written as a linear superposition $v^j T_j$.

Exponetial map

A very important idea is now that one can compose finite group elements, at least in some region around the unit element, out of very many infinitesimal transformations. In other words one writes

$$g(\xi) = \lim_{N\to\infty} \left(\mathbb{1} + \frac{i\xi^j T_j}{N}\right)^N = \exp\left(i\xi^j T_j\right). \tag{2.4}$$

One recognizes here that the limit in (2.4) would give the exponential if T_j were just numbers, and one can essentially use this limit to define also the exponential of Lie algebra elements. Alternatively, the exponential may also be evaluated as the usual power series

$$\exp\left(i\xi^j T_j\right) = \mathbb{1} + i\xi^j T_j + \frac{1}{2}\left(i\xi^j T_j\right)^2 + \frac{1}{3!}\left(i\xi^j T_j\right)^3 + \dots$$

Note that for $\alpha, \beta \in \mathbb{R}$ one can combine

$$\exp\left(i\alpha\xi^j T_j\right)\exp\left(i\beta\xi^j T_j\right) = \exp\left(i(\alpha+\beta)\xi^j T_j\right). \tag{2.5}$$

Such transformations (for fixed ξ) form a one-parameter subgroup.

Combining transformations

It is more difficult to combine transformations $\exp\left(i\xi^j T_j\right)$ and $\exp\left(i\zeta^j T_j\right)$ when ξ is not parallel to ζ. The reason is that $\xi^j T_j$ and $\zeta^j T_j$ can not be assumed to commute. To combine two transformations, one needs to use the Baker-Campbell-Hausdorff formula

$$\exp(X)\exp(Y) = \exp(Z(X,Y)), \tag{2.6}$$

with

$$Z(X,Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \dots \tag{2.7}$$

This shows that it is crucial to know how to calculate commutators between the Lie algebra elements.

Commutator and structure constants

For transformations close to the identity element we can write using (2.6) and (2.7)

$$\exp\left(i\xi^j T_j\right)\exp\left(i\zeta^j T_j\right) = \exp\left(i\omega^j T_j\right), \tag{2.8}$$

with

$$\omega^l = \xi^l + \zeta^l - \frac{1}{2}\xi^j \zeta^k f_{jk}{}^l + \ldots,$$

where the ellipses stand now for terms of quadratic and higher order in ξ and ζ. We are using here the *structure constants* $f_{jk}{}^l$ of the Lie algebra defined through the commutator

$$[T_j, T_k] = i\, f_{jk}{}^l\, T_l. \tag{2.9}$$

The structure constants are obviously anti-symmetric,

$$f_{jk}{}^l = -f_{kj}{}^l.$$

Equation (2.9) tells that the commutator of two generators can itself be expressed as a linear combination of generators. Together with eqs. (2.6) and (2.7) this makes sure that the group elements (2.4) can be multiplied and indeed form a group. In other words, if eq. (2.9) holds, we can multiply group elements as in eq. (2.8) to yield another term of the same structure such that they form a group. On the other side, one could also start from the group multiplication law and demand that the left hand side of (2.8) can be written as on the right hand side. At order $\sim \xi\zeta$ this implies then a relation of the form (2.9).

Generators and stucture constants for unitary groups

For unitary Lie groups where $g^\dagger = (1 - id\xi^j T_j^\dagger) = g^{-1} = (1 - id\xi^j T_j)$ the generators must be hermitian,

$$T_j = T_j^\dagger.$$

Also, in that case the structure constants are real,

$$f_{jk}{}^l = f^*_{jk}{}^l.$$

This follows from

$$-if^*_{jk}{}^l T_l = [T_j, T_k]^\dagger = [T_k, T_j] = if_{kj}{}^l T_l = -if_{jk}{}^l T_l.$$

Lie algebra

The definition (2.9) also makes sure that linear combinations of generators, which obviously form a vector space, constitute a *Lie algebra*. To that end, the Lie bracket $[\cdot, \cdot]$ must have the properties

(i) *Bilinearity*: $[\lambda A + \mu B, C] = \lambda[A, C] + \mu[B, C]$,

(ii) *Antisymmetry*: $[A, B] = -[B, A]$,

(iii) *Jacobi identity*: $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.

From the Jacobi identity for the generators

$$[T_j, [T_k, T_l]] + [T_k, [T_l, T_j]] + [T_l, [T_j, T_k]] = 0, \tag{2.10}$$

one infers for the structure constants

$$f_{jn}{}^m f_{kl}{}^n + f_{kn}{}^m f_{lj}{}^n + f_{ln}{}^m f_{jk}{}^n = 0. \tag{2.11}$$

Representations

The commutation relation (2.9), expressed also in terms of the structure constants, define a Lie algebra, similar as the multiplication rules do for a group. One distinguishes between a particular Lie algebra as an abstract entity and a concrete incarnation or *representation* of it.

Fundamental representation

A first example is the *fundamental representation*

$$(T_j^{(F)})^m{}_n = (t_j)^m{}_n. \tag{2.12}$$

For SU(N), the generators in the fundamental representation t_j are hermitian and traceless $N \times N$ matrices. This is a real vector space of dimension $N^2 - 1$.

For SU(2) we can write $t_j = \sigma_j/2$ with the three Pauli matrices,

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

For SU(3) one takes $t_j = \lambda_j/2$ with the eight Gell-Mann matrices

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix},$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

The restriction to the purely imaginary matrices λ_2, λ_5 and λ_7 generates the Lie algebra of the orthogonal group SO(3) in its fundamental representation.

Adjoint representation

From the Jacobi identity (2.11), one can see that the structure constants can actually be used to construct another representation, the so-called *adjoint representation*. Here one sets the matrices to

$$(T_j^{(A)})^m{}_l = i f_{jl}{}^m. \tag{2.13}$$

Indeed one has for the commutator of two generators

$$[T_j^{(A)}, T_k^{(A)}]^m{}_l = -f_{jn}{}^m f_{kl}{}^n + f_{kn}{}^m f_{jl}{}^n = -f_{jk}{}^n f_{nl}{}^m = i f_{jk}{}^n (T_n^{(A)})^m{}_l.$$

The dimension of the adjoint representation equals the number of generators of the Lie algebra. For example, the Lie algebra of SU(3) has 8 generators and accordingly the adjoint representation is given by 8×8 matrices.

The fundamental and the adjoint representation are the most important representations. The adjoint representation always exists and can be used to classify Lie algebras.

However, there are many more representations of *Lie algebras* and they all induce corresponding representations of the *Lie group* through the exponential map (2.4).

2.4 Differential geometric approach to Lie groups and Lie algebras

Left translation

Now that we understand already some of the properties of Lie groups and Lie algebras, let us discuss them also from a geometric point of view. It is interesting to consider the group multiplication as a map on the group manifold,

$$L_h : G \to G, \qquad L_h(g) = hg. \tag{2.14}$$

This is the so-called *left translation*.

Tangent space at unit element

Recall that in (2.3) we have introduced the generators T_j as a basis for the *tangent space* of the group manifold G at the identity $g = \mathbb{1}$ or $\xi = 0$. More formally, one can construct the tangent space of a manifold as a basis for vectors, which are in turn defined through curves. For this construction we first consider a curve in the group manifold parametrized by some parameter $\alpha \in \mathbb{R}$ and we assume that it goes through the unit element $g(\alpha_0) = \mathbb{1}$. We can write the curve as $g(\alpha)$, or in terms of coordinates ξ on the group manifold as $\xi(\alpha)$ such that $g(\alpha) = g(\xi(\alpha))$ and $\xi(\alpha_0) = 0$. Now consider the derivative

$$\frac{d}{d\alpha} g(\alpha)\Big|_{\alpha=\alpha_0} = \frac{\partial}{\partial \xi^j} g(\xi)\Big|_{\xi=0} \frac{d\xi^j}{d\alpha} = iT_j \frac{d\xi^j}{d\alpha}.$$

This is now an element of the tangent space $T_{\mathbb{1}}(G)$ of the group manifold at the point where $g(\xi) = \mathbb{1}$. Any element of this vector space can be written as a linear combination of the basis elements

$$iT_j = \frac{\partial}{\partial \xi^j} g(\xi)\Big|_{\xi=0}.$$

Induced basis for tangent spaces at other points

Interestingly, this basis for $T_{\mathbb{1}}(G)$ can be extended to a basis for the tangent spaces at other positions of the group manifold. To that end we can use the left translation (2.14). Specifically, from the curve $g(\alpha)$ we can construct another curve through the left translation (2.14)

$$\tilde{g}(\alpha) = L_h(g(\alpha)) = hg(\alpha).$$

The derivative at the point α_0 is now

$$\frac{d}{d\alpha} \tilde{g}(\alpha)\Big|_{\alpha=\alpha_0} = \frac{\partial}{\partial \xi^j} hg(\xi)\Big|_{\xi=0} \frac{d\xi}{d\alpha}.$$

One observes that a basis for the tangent space $T_h(G)$ is given by

$$iT_j(h) = \frac{\partial}{\partial \xi^j} hg(\xi)\Big|_{\xi=0} = h \frac{\partial}{\partial \xi^j} g(\xi)\Big|_{\xi=0} = i\, hT_j. \tag{2.15}$$

In this way we can actually get a basis for the tangent spaces everywhere in the entire group manifold. It is quite non-trivial that the tangent spaces can be parametrized by a single set of basis functions $iT_j(h)$. One says that the manifold G is *parallelizable*.

Vector fields on group manifolds

Formally, the map (2.15) between the tangent spaces $T_{\mathbb{1}}(G)$ and $T_h(G)$ is an example for a *push-forward*, induced by the map (2.14) on the manifold itself. One also writes this as

$$T_j(h) = L_{h*} T_j(\mathbb{1}) = L_{h*} T_j.$$

One may now construct *vector fields* on the entire manifold as linear combinations,

$$V(h) = v^j(h) T_j(h). \tag{2.16}$$

Left-invariant vector fields

Such a vector field is called *left invariant* if

$$L_{g*} V(h) = V(gh).$$

Because the basis $T_j(h)$ is left-invariant by construction, the vector field (2.16) is left-invariant when the coefficients $v^j(h)$ are independent of the position on the manifold, i. e. independent of h.

In summary, we may say that the generators of the Lie algebra T_j induce actually a *left-invariant* basis for *vector fields* on the entire group manifolds. One may even understand the Lie algebra itself as an algebra of left-invariant vector fields. The Lie bracket is then introduced as the *Lie derivative* of vector fields.

2.5 Examples for matrix Lie algebras

Let us end this section with a few examples for Lie algebras corresponding to matrix Lie groups introduced previously.

- $\mathfrak{su}(n)$ is the Lie algebra corresponding to the group $\mathrm{SU}(n)$. We write the group elements as $U = \exp(it)$. From $U^\dagger U = \mathbb{1}$ one infers $t^\dagger = t$. Writing this in components, the real part is symmetric, $\mathrm{Re}(t_{nm}) = \mathrm{Re}(t_{mn})$, and the imaginary part is anti-symmetric, $\mathrm{Im}(t_{nm}) = -\mathrm{Im}(t_{mn})$. Moreover, we have the condition $\det(U) = 1$. The latter can be rewritten as

$$\det(U) = \exp(\ln(\det(U))) = \exp(\mathrm{Tr}\{\ln(U)\}) = \exp(i\mathrm{Tr}\{t\}) = 1, \tag{2.17}$$

  so that we need $\mathrm{Tr}\{t\} = 0$. These arguments show that the Lie algebra $\mathfrak{su}(n)$ as a real vector space has $n^2 - 1$ linearly independent generators T_j.

- $\mathfrak{so}(n)$ is the Lie algebra corresponding to the group $\mathrm{SO}(n)$. Here we write the group elements as $R = \exp(it)$ and they are real matrices such that $R^T R = \mathbb{1}$. For the Lie algebra elements we have again $t = t^\dagger$. In order for an infinitesimal transformation $R = \mathbb{1} + it$ to be real, the components t_{mn} must be imaginary, and therefore also anti-symmetric. The condition $\mathrm{Tr}\{t\} = 0$ is then automatically fulfilled. These arguments show that the Lie algebra $\mathfrak{so}(n)$ has $n(n-1)/2$ linearly independent generators T_j.

- $\mathfrak{sp}(2n)$ is the Lie algebra corresponding to the group $\mathrm{Sp}(2n)$. The group elements $R = \exp(it)$ are real matrices that satisfy $R^T \Omega R = \Omega$ with $\Omega = -\Omega^T$ given in (2.1). For an infinitesimal transformation $R = \mathbb{1} + it$ one finds the condition

$$\Omega t + t^T \Omega = \Omega t - t^T \Omega^T = \Omega t - (\Omega t)^T = 0. \tag{2.18}$$

  In other words, Ωt must be symmetric. These arguments show that the Lie algebra $\mathfrak{sp}(2n)$ has $n(2n+1)$ linearly independent generators T_j.

LECTURE 03

3 Classical field theory

3.1 Relativistic scalar field theory

Classical action

Let us consider the classical action for a real scalar field

$$\begin{aligned} S[\phi] &= \int dt \int d^3x\, \mathscr{L}(\phi, \partial_\mu \phi) = \int dt \int d^3x \left\{ \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 \right\} \\ &= \int dt \int d^3x \left\{ -\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 \right\} \end{aligned} \tag{3.1}$$

We have suppressed here the argument, but the field ϕ is to be understood as a (real) function of time and space

$$\phi = \phi(x) = \phi(t, \mathbf{x}).$$

The integrand $\mathscr{L}$ is known as the *Lagrange density*. We represent time derivatives by dots, i. e.

$$\dot{\phi} = \frac{\partial}{\partial t}\phi(t, \mathbf{x}),$$

and the spatial gradient by $\boldsymbol{\nabla}\phi(t, \mathbf{x})$. In the following we use the abbreviation

$$\int_x = \int d^4x = \int dt \int d^3x.$$

It is sometimes convenient to restrict the integrals to some intervals in time, for example $t_\mathrm{i} \leq t \leq t_\mathrm{f}$, and the spatial integral could be restricted to some volume V. In that case one must specify what boundary conditions the field $\phi(t, \mathbf{x})$ is supposed to satisfy on the boundary ∂V. Some common choices are Dirichlet boundary conditions, where the field is fixed to some value on the boundary, or Neumann boundary conditions, where the normal derivative of the field is fixed, or periodic boundary conditions.

The square brackets in $S[\Phi]$ indicate 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 $\phi(t, \mathbf{x})$, with $(t, \mathbf{x}) \in \mathbb{R}^d$ where we usually take $d = 1 + 3$. We also use the Minkowski metric with mainly plus signature, $g_{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1)$.

Blackboard video

Fields as vectors

One can consider $\phi(t, \mathbf{x})$ as a vector in a (real) vector space of infinite dimension where components are labeled by time t and the spatial position $\mathbf{x}$. In particular, linear superpositions of field configurations $\phi(t, \mathbf{x}) = \lambda_1\phi_1(t, \mathbf{x}) + \lambda_2\phi_2(t, \mathbf{x})$ are again field configurations. It is sometimes useful to think about a field theory as limit of a discrete lattice model where the positions $\mathbf{x}$ and times t are restricted to discrete points on some lattice. When space and time are furthermore constrained to a finite region of spacetime, the number of spacetime positions $(t, \mathbf{x})$ becomes finite. A field configuration is then specified by a set of real numbers, one per spacetime lattice site. This would be a finite set of numbers in the lattice scheme but it becomes infinitely large in the continuum and infinite volume limits.

Functional spaces

How regular should one assume field configurations $\phi(t, \mathbf{x})$ to be? It is tempting to assume that they should be continuous and at least differentiable once so that the action in eq. (3.1) is well defined. Assumptions of this kind are sometimes justified in classical physics, but in the statistical and quantum field theoretic formalism we discuss below one also needs to work with very irregular field configurations that are not even continuous.

3.2 Variational principle and equations of motion

Variation of the action

One can obtain the classical equation of motion by the principle of stationary action

$$\delta S = 0.$$

For the action in (3.1) this yields

$$\begin{aligned}\delta S &= \int_x \left\{ \dot\phi(x)\, \delta\dot\phi(x) - \boldsymbol{\nabla}\phi(x)\, \delta\boldsymbol{\nabla}\phi(x) - m^2\phi(x)\, \delta\phi(x) - \frac{\lambda}{3!}\phi(x)^3\, \delta\phi(x) \right\} \\ &= \int_x \left\{ -g^{\mu\nu}\partial_\mu\phi(x)\, \delta\partial_\nu\phi(x) - m^2\phi(x)\, \delta\phi(x) - \frac{\lambda}{3!}\phi(x)^3\, \delta\phi(x) \right\}.\end{aligned}$$

Note that on the right hand side the variation of the field $\delta\phi(x)$ occurs, but also the variation of the derivatives $\delta\partial_\mu\phi(x)$.

Partial integration

To deal with the time derivative of the variation $\delta\dot\phi(x)$ we integrate by parts,

$$\begin{aligned}\int_{t_\mathrm{i}}^{t_\mathrm{f}} dt \int_V d^3x \left\{ \dot\phi(x)\, \delta\dot\phi(x) \right\} &= \int_V d^3x \left\{ \dot\phi(t_\mathrm{f},\mathbf{x})\, \delta\phi(t_\mathrm{f},\mathbf{x}) - \dot\phi(t_\mathrm{i},\mathbf{x})\, \delta\phi(t_\mathrm{i},\mathbf{x}) \right\} \\ &\quad - \int_{t_\mathrm{i}}^{t_\mathrm{f}} dt \int_V d^3x \left\{ \ddot\phi(x)\, \delta\phi(x) \right\}\end{aligned}$$

The first term on the right hand side is a boundary term at the final time t_f and initial time t_i. Similarly one can deal with the spatial gradient term,

$$\begin{aligned}\int_{t_\mathrm{i}}^{t_\mathrm{f}} dt \int_V d^3x \left\{ -\boldsymbol{\nabla}\phi(x)\, \delta\boldsymbol{\nabla}\phi(x) \right\} &= -\int_{t_\mathrm{i}}^{t_\mathrm{f}} dt \int_{\partial V} d^2x\, \mathbf{n} \left\{ \boldsymbol{\nabla}\phi(x)\, \delta\phi(x) \right\} \\ &\quad + \int_{t_\mathrm{i}}^{t_\mathrm{f}} dt \int_V d^3x \left\{ \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\phi(x)\, \delta\phi(x) \right\}.\end{aligned}$$

The first term on the right hand side is a surface integral with outward-pointing normal vector $\mathbf{n}$ on the boundary ∂V.

Fixed boundary conditions

We consider now fully fixed or constrained field configurations on the boundaries of the spacetime volume at t_i, t_f and the spatial boundary ∂V. This means that the variation $\delta\phi(x)$ is supposed to vanish there, $\delta\phi(x) = 0$. For example we could demand at initial and final time

$$\phi(t_\mathrm{i},\mathbf{x}) = \phi_\mathrm{i}(\mathbf{x}), \qquad \phi(t_\mathrm{f},\mathbf{x}) = \phi_\mathrm{f}(\mathbf{x}), \tag{3.2}$$

and take the spatial volume V to be all of $\mathbb{R}^3$ so that there is actually no spatial boundary.

Combining terms yields then

$$\delta S = \int_x \left\{ -\ddot\phi(x) + \boldsymbol{\nabla}^2\phi(x) - m^2\phi(x) - \frac{\lambda}{3!}\phi(x)^3 \right\} \delta\phi(x).$$

For this to vanish for arbitrary variation $\delta\phi(x)$ inside the spacetime volume we need

$$-\ddot\phi(x) + \boldsymbol{\nabla}^2\phi(x) - m^2\phi(x) - \frac{\lambda}{3!}\phi(x)^3 = 0, \tag{3.3}$$

which is the classical equation of motion. Together with the boundary conditions (3.2) the classical field $\phi(t,\mathbf{x})$ is actually fully fixed through this differential equation. Indeed, this equation of motion is a quasi-linear hyperbolic partial differential equation of second order. A solution is determined by initial data in the form of a configuration of the field $\phi(t_\mathrm{i},\mathbf{x}) = \phi_\mathrm{i}(\mathbf{x})$ and its first time derivative, on some Cauchy hypersurface like $t = t_\mathrm{i}$, or alternatively by field configurations at initial and final time as in eq. (3.2).

Blackboard video

3.3 Maxwell theory

Gauge field and field strength tensor

Another example for a classical field theory is Maxwell theory. The field is here the four-vector potential, or gauge field $A_\mu(x) = (-\Phi(x), \mathbf{A}(x))$. We work in unites where $c = 1$ and with Minkowski space metric $\eta_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$. The electric and magnetic fields are given by

$$\mathbf{E}(x) = -\frac{\partial}{\partial t}\mathbf{A}(x) - \boldsymbol{\nabla}\Phi(x), \qquad \mathbf{B}(x) = \boldsymbol{\nabla} \times \mathbf{A}(x),$$

and can be combined into the field strength tensor

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}. \tag{3.4}$$

Note that this immediately implies the homogeneous Maxwell equations,

$$\varepsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma}(x) = \varepsilon^{\mu\nu\rho\sigma}\partial_\nu[\partial_\rho A_\sigma(x) - \partial_\sigma A_\rho(x)] = 0,$$

following from the fact that $\partial_\nu\partial_\rho A_\sigma(x) = \partial_\rho\partial_\mu A_\sigma(x)$. Here we are using the completely antisymmetric Levi-Civita symbol $\varepsilon^{\mu\nu\rho\sigma}$ in $d = 1 + 3$ dimensions.

One can also write the homogeneous Maxwell equations in the familiar form $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$ and $\partial_t\mathbf{B} + \boldsymbol{\nabla} \times \mathbf{E} = 0$.

Gauge invariance

We note immediately that the field strength tensor is antisymmetric, $F_{\mu\nu} = -F_{\nu\mu}$. It is invariant under gauge transformations

$$A_\mu(x) \to A_\mu(x) + \partial_\mu\alpha(x), \tag{3.5}$$

where $\alpha(x)$ is an arbitrary scalar function with $\partial_\mu\partial_\nu\alpha(x) = \partial_\nu\partial_\mu\alpha(x)$. This also implies that $A_\mu(x)$ is not fully fixed by the measurable electric and magnetic fields, but only up to gauge transformations.

Action

The action for classical Maxwell theory in in the presence of some electromagnetic current $J^\mu(x) = (\rho(x), \mathbf{J}(x))$ given by

$$\begin{aligned} S[A_\mu] &= \int d^dx \left\{ -\frac{1}{4}F^{\mu\nu}(x)F_{\mu\nu}(x) - J^\mu(x)A_\mu(x) \right\} \\ &= \int d^dx \left\{ -\frac{1}{2}\partial^\mu A^\nu(x)\partial_\mu A_\nu(x) + \frac{1}{2}\partial^\mu A^\nu(x)\partial_\nu A_\mu(x) - J^\mu(x)A_\mu(x) \right\}. \end{aligned}$$

Note that this is invariant under gauge transformations (3.5), up to a boundary term, when the current is conserved, $\partial_\mu J^\mu(x) = 0$. For the first term this is clear because the field strength tensor is invariant, and for the second term it follows through partial integration.

One can expand the action (3.3) in the form

$$S = \int d^dx \left\{ \frac{1}{2}\partial_k A_0\partial_k A_0 - \partial_0 A_k\partial_k A_0 + \frac{1}{2}\partial_0 A_k\partial_0 A_k - \frac{1}{2}\partial_j A_k\partial_j A_k + \frac{1}{2}\partial_j A_k\partial_k A_j - J^0 A_0 - J^k A_k \right\},$$

where Latin indices $j, k = 1, 2, 3$ run over spatial components.

Variation

To vary the action we need the variation of the field strength tensor, expressed in terms of the variation of the gauge field,

$$\delta F_{\mu\nu}(x) = \partial_\mu \delta A_\nu(x) - \partial_\nu \delta A_\mu(x).$$

Performing then partial integrations and using the anti-symmetry property of $F_{\mu\nu}$ yields

$$\begin{aligned}\delta S &= \int d^d x \left\{ -\frac{1}{2} F^{\mu\nu}(x) \delta F_{\mu\nu}(x) - J^\mu(x) \delta A_\mu(x) \right\} \\ &= \int d^d x \left\{ \left[\partial_\mu F^{\mu\nu}(x) - J^\nu(x) \right] \delta A_\nu(x) \right\}.\end{aligned}$$

Accordingly, the principle of stationary action implies the inhomogeneous Maxwell equations

$$\partial_\mu F^{\mu\nu}(x) = J^\nu(x).$$

One can also write these in the familiar form $\boldsymbol{\nabla} \cdot \mathbf{E} = \rho$ and $-\partial_t \mathbf{E} + \boldsymbol{\nabla} \times \mathbf{B} = \mathbf{J}$.

3.4 Gross-Pitaevskii theory

Action

We also discuss an example for a non-relativistic classical field theory. It describes a Bose-Einstein condensate of bosons with repulsive contact interaction. The action is

$$S[\phi] = \int dt \int d^3x \left\{ \frac{i\hbar}{2} [\phi^* \partial_t \phi - \phi \partial_t \phi^*] - \frac{\hbar^2}{2m} \boldsymbol{\nabla}\phi^* \boldsymbol{\nabla}\phi - V\phi^*\phi - \frac{\lambda}{2}(\phi^*\phi)^2 \right\}. \tag{3.6}$$

The field $\phi(t, \mathbf{x})$ is here complex. We have included an external potential $V(t, \mathbf{x})$ which could be an optical trap, for example, and $\lambda \geq 0$ is the interaction parameter. Note that despite the presence of the imaginary unit i the action is real. This becomes more explicit when it is rewritten in terms of real fields φ_1 and φ_2 defined through $\phi = [\varphi_1 + i\varphi_2]/\sqrt{2}$.

Variation and equation of motion

For a complex field one can either consider the real and imaginary part as independent fields, and vary with respect to them, or one can vary with respect to ϕ and ϕ^*. In the latter case one obtains two complex conjugate equations of motion, so it is enough to do one of these variations. Specifically, variation with respect to ϕ^* yields, up to boundary terms,

$$\delta S = \int dt \int d^3x \left\{ \left[i\hbar \partial_t \phi + \frac{\hbar^2}{2m} \boldsymbol{\nabla}^2 \phi - V\phi - \lambda \phi^* \phi \phi \right] \delta\phi^* \right\}.$$

The corresponding equation of motion is a kind of non-linear Schrödinger equation, the Gross-Pitaevskii equation

$$i\hbar \partial_t \phi(t, \mathbf{x}) = -\frac{\hbar^2}{2m} \boldsymbol{\nabla}^2 \phi(t, \mathbf{x}) + V(t, \mathbf{x}) \phi(t, \mathbf{x}) + \lambda \phi^*(t, \mathbf{x}) \phi(t, \mathbf{x}) \phi(t, \mathbf{x}). \tag{3.7}$$

Because this is a differential equation of first order in time, initial conditions can be posed in the form $\phi(t_\mathrm{i}, \mathbf{x}) = \phi_\mathrm{i}(\mathbf{x})$. As a complex equation this fixes both the real and imaginary part $\varphi_1(t_\mathrm{i}, \mathbf{x})$ and $\varphi_2(t_\mathrm{i}, \mathbf{x})$, so the number of initial conditions is the same as for the relativistic real scalar field.

3.5 Hamiltonian formalism

Conjugate momentum field

As familiar from classical mechanics, one may also introduce a Hamiltonian description which is connected to the Lagrangian description through a Legendre transform. As a first step in that direction one defines a conjugate momentum field. One starts here from the Lagrange *function* at some given time t,

$$L[\Phi(t), \dot{\Phi}(t)] = \int d^3x \, \mathscr{L}(\Phi(t, \mathbf{x}), \partial_\mu \Phi(t, \mathbf{x})).$$

For a field theory, the Lagrange function is a *functional* of the field $\Phi(t, \mathbf{x})$ at fixed time t, with spatial position $\mathbf{x}$ treated similar to an index in mechanics. In other words, we can see the field theory as a mechanical system in the continuum limit with one degree of freedom per spatial position $\mathbf{x}$. The difference is that $\mathbf{x}$ is being integrated over instead of the sum over different mechanical degrees of freedom familiar from mechanics.

The field $\Phi(t, \mathbf{x})$ has a canonical conjugate momentum field $\Pi(t, \mathbf{x})$, which is defined through

$$\Pi(t, \mathbf{x}) = \frac{\partial \mathscr{L}(\Phi(t, \mathbf{x}), \partial_\mu \Phi(t, \mathbf{x}))}{\partial \dot{\Phi}(t, \mathbf{x})}.$$

Examples

Let us consider the actions we have introduced already.

- For the real relativistic scalar field with action as in (3.1) one finds the momentum field conjugate to the real scalar field $\phi(t, \mathbf{x})$ to be

$$\pi(t, \mathbf{x}) = \dot{\phi}(t, \mathbf{x}).$$

- For the Maxwell theory with action as in (3.3) one finds the momentum field conjugate to $A_k(t, \mathbf{x})$ to be minus the electric field,

$$\pi_k(t, \mathbf{x}) = \frac{\partial \mathscr{L}}{\partial \dot{A}_k(t, \mathbf{x})} = \partial_0 A_k(t, \mathbf{x}) - \partial_k A_0(t, \mathbf{x}) = F_{0k}(t, \mathbf{x}) = -E_k(t, \mathbf{x}).$$

  In contrast, the field $A_0(t, \mathbf{x})$ has a vanishing conjugate momentum field,

$$\pi_0(t, \mathbf{x}) = \frac{\partial \mathscr{L}}{\partial \dot{A}_0(t, \mathbf{x})} = 0.$$

- Finally, for the Gross-Pitaevskii theory with action as in (3.6) one finds the momentum field conjugate to the complex scalar field $\phi(t, \mathbf{x})$ to be

$$\pi_\phi(t, \mathbf{x}) = \frac{\partial \mathscr{L}}{\partial \dot{\phi}(t, \mathbf{x})} = \frac{i\hbar}{2} \phi^*(t, \mathbf{x}),$$

  while the conjugate momentum field to the complex conjugate field $\phi^*(t, \mathbf{x})$ is

$$\pi_{\phi^*}(t, \mathbf{x}) = \frac{\partial \mathscr{L}}{\partial \dot{\phi}^*(t, \mathbf{x})} = \frac{-i\hbar}{2} \phi(t, \mathbf{x}).$$

Hamiltonian as Legendre transform

The Hamiltonian (at some time t) is now given by the functional Legendre transform

$$\begin{aligned} H[\Phi(t),\Pi(t)] &= \int d^3x\, \mathscr{H}(\Phi(t,\mathbf{x}),\boldsymbol{\nabla}\Phi(t,\mathbf{x}),\Pi(t,\mathbf{x})) \\ &= \int d^3x \sup_{\dot\Phi(t,\mathbf{x})} \left\{ \Pi(t,\mathbf{x})\dot\Phi(t,\mathbf{x}) - \mathscr{L}(\Phi(t,\mathbf{x}),\partial_\mu\Phi(t,\mathbf{x})) \right\}. \end{aligned}$$

It is to be understood as a functional of $\Phi(t,\mathbf{x})$ and $\Pi(t,\mathbf{x})$ at fixed time t but with $\mathbf{x}$ playing the role of an index. Through the Legendre transform the time derivative of the field $\dot\Phi(t,\mathbf{x})$ is replaced by the conjugate momentum field $\Pi(t,\mathbf{x})$.

Blackboard video

Examples

For the three theories introduced above we find

- For the real relativistic scalar field with action as in (3.1) the Hamiltonian is

$$H = \int d^3x \left\{ \frac{1}{2}\pi(t,\mathbf{x})^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi(t,\mathbf{x}))^2 + V(\phi(t,\mathbf{x})) \right\}, \tag{3.8}$$

  where we have introduced the potential

$$V(\phi) = \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}\phi^4.$$

Blackboard video

- For the Gross-Pitaevskii theory with action as in (3.6) one finds the Hamiltonian to be

$$H = \int d^3x \left\{ \frac{\hbar^2}{2m}\boldsymbol{\nabla}\phi^*(t,\mathbf{x})\cdot\boldsymbol{\nabla}\phi(t,\mathbf{x}) + V(t,\mathbf{x})\phi^*(t,\mathbf{x})\phi(t,\mathbf{x}) + \frac{\lambda}{2}\phi^*(t,\mathbf{x})^2\phi(t,\mathbf{x})^2 \right\}. \tag{3.9}$$

Note that in both theories the Hamiltonian density $\mathscr{H}$ is bounded from below for $\lambda > 0$. For the Maxwell theory the Hamiltonian description is more involved, essentially because A_0 has vanishing conjugate momentum field. We will not discuss this here.

Functional differentiation

In the following we need the notion of a functional derivative. For a functional $I[\phi]$ with variation

$$\delta I[\phi] = \int d^dx \left\{ f(x)\delta\phi(x) \right\},$$

one defines

$$\frac{\delta I[\phi]}{\delta\phi(x)} = f(x).$$

More generally, when the variation is of the form

$$\delta I[\phi] = \int d^d x \left\{ f(x)\delta\phi(x) + g^j(x)\frac{\partial}{\partial x^j}\delta\phi(x) + h^{jk}(x)\frac{\partial^2}{\partial x^j \partial x^k}\delta\phi(x) + \dots \right\}$$

one needs to first perform partial integrations. Usually the series on the right hand side terminates after one, two or three terms.When $\delta\phi(x)$ is assumed to vanish on the boundaries one can drop the boundary terms arising from the partial integration. This yields

$$\delta I[\phi] = \int d^d x \left\{ \left[f(x) - \frac{\partial}{\partial x^j} g^j(x) + \frac{\partial^2}{\partial x^j \partial x^k} h^{jk}(x) - \dots \right] \delta\phi(x) \right\},$$

and the functional derivative is thus

$$\frac{\delta I[\phi]}{\delta\phi(x)} = f(x) - \frac{\partial}{\partial x^j} g^j(x) + \frac{\partial^2}{\partial x^j \partial x^k} h^{jk}(x) - \dots .$$

As an example consider the functional $I[\phi] = \phi(y)$. One can write this as

$$I[\phi] = \int d^d x \delta^{(d)}(x-y)\phi(x),$$

and thus the functional derivative is here

$$\frac{\delta I[\phi]}{\delta\phi(x)} = \frac{\delta\phi(y)}{\delta\phi(x)} = \delta^{(d)}(x-y).$$

The definition of the functional derivative depends slightly on the context, in particular it is sometimes used for integrals over space and sometimes for integrals over space and time. With a bit of care it gets clear from the context what is the right definition in a given context.

Poisson brackets

In classical mechanics one introduces *Poisson brackets* to describe time evolution or symmetry transformations in the Hamiltonian formalism. This can also be done in a classical field theory. The Poisson bracket between two functionals $A[\Phi(t), \Pi(t)]$ and $B[\Phi(t), \Pi(t)]$ of the fields and conjugate momenta at some given fixed time t is defined as

$$\{A, B\} = \int d^3 x \left\{ \frac{\delta A}{\delta\Phi(t,\mathbf{x})}\frac{\delta B}{\delta\Pi(t,\mathbf{x})} - \frac{\delta A}{\delta\Pi(t,\mathbf{x})}\frac{\delta B}{\delta\Phi(t,\mathbf{x})} \right\}.$$

The functional derivatives are here defined for three-dimensional integrals over space.

In particular, by taking the Poisson bracket with the Hamiltonian, one can obtain the time derivative of a functional $A[\Phi(t), \Pi(t)]$ without explicit time dependence, along the solution to the equation of motion,

$$\frac{d}{dt} A[\Phi(t), \Pi(t)] = \{A, H\}. \tag{3.10}$$

For a functional $A[\Phi(t), \Pi(t)]$ with explicit time dependence one has to add the partial time derivative of A to the right hand side.

Relativistic scalar theory as example

For the real relativistic scalar field theory the Hamiltonian is given in eq. (3.8). To calculate the Poisson brackets we need the functional derivatives

$$\frac{\delta H[\pi(t), \phi(t)]}{\delta\pi(t,\mathbf{x})} = \pi(t,\mathbf{x}),$$

and

$$\frac{\delta H[\pi(t),\phi(t)]}{\delta\phi(t,\mathbf{x})} = -\boldsymbol{\nabla}^2\phi(t,\mathbf{x}) + m^2\phi(t,\mathbf{x}) + \frac{\lambda}{3!}\phi(t,\mathbf{x})^3.$$

As a first example let us take $A[\pi(t),\phi(t)] = \phi(t,\mathbf{y})$. Here we find $\delta A/\delta\phi(t,\mathbf{x}) = \delta^{(3)}(\mathbf{x}-\mathbf{y})$ and $\delta A/\delta\pi(t,\mathbf{x}) = 0$ such that eq. (3.10) gives

$$\frac{d}{dt}A[\phi(t),\pi(t)] = \dot{\phi}(t,\mathbf{y}) = \pi(t,\mathbf{x}).$$

This is consistent with the definition of the conjugate momentum field for this theory.

Similarly, for the choice $A[\phi(t),\pi(t)] = \pi(t,\mathbf{y})$ we have $\delta A/\delta\phi(t,\mathbf{x}) = 0$ and $\delta A/\delta\pi(t,\mathbf{x}) = \delta^{(3)}(\mathbf{x}-\mathbf{y})$ such that eq. (3.10) gives

$$\frac{d}{dt}A[\phi(t),\pi(t)] = \dot{\pi}(t,\mathbf{y}) = \boldsymbol{\nabla}^2\phi(t,\mathbf{y}) - m^2\phi(t,\mathbf{y}) - \frac{\lambda}{3!}\phi(t,\mathbf{y})^3.$$

This is the equation of motion previously obtained from variation of the action.

Fundamental Poisson brackets

Based on the definitions one may easily check the fundamental Poisson brackets

$$\{\phi(t,\mathbf{x}),\phi(t,\mathbf{y})\} = \{\pi(t,\mathbf{x}),\pi(t,\mathbf{y})\} = 0, \qquad \{\phi(t,\mathbf{x}),\pi(t,\mathbf{y})\} = \delta^{(3)}(\mathbf{x}-\mathbf{y}).$$

These can be taken as a starting point for "canonical quantization", which is a heuristic transition from a classical field theory to a quantum field theory.

Blackboard video

LECTURE 04

3.6 Noether's theorems

There is a fundamental relation between symmetries and conservation laws, first formulated by Emmy Noether (1882 – 1935). We discuss this here in the context of classical field theory.

Global symmetries

Consider again the action for a classical field theory, which we write in the form

$$S[\Phi] = \int_\Omega d^4x \mathscr{L}(\Phi(x),\partial_\mu\Phi(x)).$$

Here Ω denotes some region in spacetime, for example bounded by an initial time t_i and final time t_f. We now study continuous transformations of the fields parametrized by some real number ξ such that $\xi = 0$ is the identity transformation. It is enough to study infinitesimal transformations out of which we can build also finite transformations. We write them as

$$\Phi(x) \to \Phi(x) + \delta_\xi\Phi(x).$$

We assume that the action is invariant up to a total derivative or boundary term,

$$S[\Phi] \to S[\Phi] + \delta S[\Phi] = S[\Phi] + \int_\Omega d^4x\,\{\partial_\mu\Lambda^\mu(x)\} = S[\Phi] + \int_{\partial\Omega} d\Sigma_\mu\Lambda^\mu(x).$$

In such a situation we speak of a continuous global symmetry of the action.

Local conservation law

Noether's first theorem states that for every continuous global symmetry of the action there exists a conserved current. So see this we consider the change in the action, written in the form

$$\begin{aligned}\delta S[\Phi] &= \int_\Omega d^4x \left\{ \frac{\partial \mathscr{L}}{\partial \Phi(x)} \delta_\xi \Phi(x) + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi(x))} \partial_\mu \delta_\xi \Phi(x) \right\} \\ &= \int_\Omega d^4x \left\{ \left[\frac{\partial \mathscr{L}}{\partial \Phi(x)} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi(x))} \right] \delta_\xi \Phi(x) + \partial_\mu \left[\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi(x))} \delta_\xi \Phi(x) \right] \right\} = \int_\Omega d^4x \partial_\mu \left\{ \Lambda^\mu(x) \right\}.\end{aligned}$$

We now use the equations of motion,

$$\frac{\delta S}{\delta \Phi(x)} = \frac{\partial \mathscr{L}}{\partial \Phi(x)} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi(x))} = 0,$$

and obtain the local conservation law

$$\partial_\mu \left[\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi(x))} \delta_\xi \Phi(x) - \Lambda^\mu(x) \right] = 0.$$

Global U(1) symmetry

As a first example we consider a global U(1) transformation for the complex scalar field,

$$\phi(x) \to e^{i\alpha} \phi(x), \qquad \phi^*(x) \to e^{-i\alpha} \phi^*(x).$$

In infinitesimal form this reads

$$\delta_\alpha \phi(x) = i\alpha \phi(x), \qquad \delta_\alpha \phi^*(x) = -i\alpha \phi^*(x). \tag{3.11}$$

The Gross-Pitaevskii action (3.6) is invariant under this transformation with $\Lambda^\mu(x) = 0$. We find here

$$\frac{\partial \mathscr{L}}{\partial(\partial_0 \phi)} = \frac{i\hbar}{2} \phi^*, \qquad \frac{\partial \mathscr{L}}{\partial(\partial_0 \phi^*)} = -\frac{i\hbar}{2} \phi,$$

and

$$\frac{\partial \mathscr{L}}{\partial(\partial_j \phi)} = -\frac{\hbar^2}{m} \partial_j \phi^*, \qquad \frac{\partial \mathscr{L}}{\partial(\partial_j \phi^*)} = -\frac{\hbar^2}{m} \partial_j \phi.$$

Noethers theorem implies the conservation law

$$\partial_0 \left[-\alpha \hbar \phi^*(x) \phi(x) \right] + \partial_j \left[i\alpha \frac{\hbar^2}{m} \phi^* \partial_j \phi - i\alpha \frac{\hbar^2}{m} \phi \partial_j \phi^* \right] = 0,$$

or equivalently $\partial_\mu N^\mu(x) = 0$ with the current composed out of the particle number density

$$N^0(t, \mathbf{x}) = \phi^*(t, \mathbf{x}) \phi(t, \mathbf{x})$$

and the particle number current

$$\mathbf{N}(t, \mathbf{x}) = -\frac{i\hbar}{m} \left[\phi^*(t, \mathbf{x}) \boldsymbol{\nabla} \phi(t, \mathbf{x}) - \phi(t, \mathbf{x}) \boldsymbol{\nabla} \phi^*(t, \mathbf{x}) \right].$$

These expressions agree formally with the probability density and probability current in single-particle quantum mechanics, but have gained a different significance in the context of the classical Gross-Pitaevskii field theory.

Translations

Another interesting symmetry transformation is given by translations in spacetime, $x^\mu \to x^\mu + a^\mu$. This should actually be seen as a set of four linearly independent transformations for the choices $\mu = 0, 1, 2, 3$, which we discuss here together. Fields change according to

$$\Phi(x) \to \Phi(x - a),$$

such that for example a maximum at a position x_0^μ before the transformation is moved to $x_0^\mu + a^\mu$ after the transformation. When the transformation is infinitesimal we can write

$$\Phi(x) \to \Phi(x) + \delta_a \Phi(x) = \Phi(x) - a^\mu \partial_\mu \Phi(x).$$

The Lagrangian density transforms itself like a scalar field, $\mathscr{L}(x) \to \mathscr{L}(x) - a^\mu \partial_\mu \mathscr{L}(x)$.

Energy-momentum conservation law

Accordingly we find a symmetry of the action with $\Lambda^\mu(x) = -a^\mu \mathscr{L}(x)$, and the local conservation law is according to Noethers first theorem

$$\partial_\mu \left[\frac{\partial \mathscr{L}}{\partial \partial_\mu \Phi} a^\nu \partial_\nu \Phi - a^\mu \mathscr{L} \right] = 0.$$

Because a^μ is arbitrary we obtain the conservation law for the canonical energy-momentum tensor,

$$\partial_\mu \mathscr{T}^{\mu\nu}(x) = 0,$$

with

$$\mathscr{T}^{\mu\nu}(x) = -\frac{\partial \mathscr{L}}{\partial \partial_\mu \Phi(x)} g^{\nu\rho} \partial_\rho \Phi(x) + g^{\mu\nu} \mathscr{L}(x).$$

Energy-momentum tensor for real relativistic scalar field

As an example we consider the real relativistic scalar field theory with action in eq. (3.1). The energy-momentum tensor is given by

$$\mathscr{T}^{\mu\nu}(x) = g^{\mu\rho} g^{\nu\sigma} \partial_\rho \phi(x) \partial_\sigma \phi(x) + g^{\mu\nu} \left[-\frac{1}{2} g^{\rho\sigma} \partial_\rho \phi(x) \partial_\sigma \phi(x) - V(\phi(x)) \right].$$

The zero-zero component is again the Hamiltonian density already derived by different means in equation (3.8).

LECTURE 05

3.7 Scalar field theory in general coordinates

Action for complex scalar field in general coordinates and with external gauge field

There exists are more elegant formulation of symmetries and the associated conservation laws wich is the subject of Noethers second theorem. The idea it to promote the global symmetry to a local transformation, i. e. one that depends on spacetime coordinates. Consider the following action for a complex relativistic scalar field

$$S[\phi] = \int d^4x \sqrt{g(x)} \left\{ -g^{\mu\nu}(x) \left[\partial_\mu \phi^*(x) + i A_\mu(x) \phi^*(x) \right] \left[\partial_\nu \phi(x) - i A_\mu(x) \phi(x) \right] - V(\phi^*(x)\phi(x)) \right\}, \tag{3.12}$$

with $\sqrt{g(x)} = \sqrt{-\det(g_{\mu\nu}(x))}$. We have written the theory in general (not necessarily Cartesian) coordinates which leads to the appearance of the spacetime-dependent metric $g_{\mu\nu}(x)$ with inverse $g^{\mu\nu}(x)$. We have also introduced an external gauge field $A_\mu(x)$ and replaced partial derivatives $\partial_\mu \phi(x)$ by covariant derivatives $\partial_\mu \phi(x) - i A_\mu(x) \phi(x)$.

Local U(1) symmetry

Due to the presence of the gauge field $A_\mu(x)$ we can now consider local U(1) transformations of the form

$$\phi(x) \to e^{i\alpha(x)}\phi(x), \qquad \phi^*(x) \to e^{-i\alpha(x)}\phi^*(x), \qquad A_\mu(x) \to A_\mu(x) + \partial_\mu\alpha(x). \tag{3.13}$$

This leaves the action (3.12) invariant. This is immediately clear for the potential term $V(\phi^*\phi)$ when it is taken to depend only on the U(1) invariant combination $\phi^*\phi$. For the derivative terms we have

$$\partial_\mu\phi(x) \to \partial_\mu\left[e^{i\alpha(x)}\phi(x)\right] = e^{i\alpha(x)}\partial_\mu\phi(x) + e^{i\alpha(x)}\phi(x)i\partial_\mu\alpha(x).$$

However, the inhomogeneous term on the right-hand side gets canceled by the transformation of the gauge field term in covariant derivatives.

Conservation law from gauge invariance

Noether's second theorem is concerned with the conservation laws that arise from local symmetries. Let us consider the local U(1) symmetry in eq. (3.13) in the infinitesimal form $\delta_\alpha\phi(x) = i\alpha(x)\phi(x)$, $\delta_\alpha\phi^*(x) = -i\alpha(x)\phi^*(x)$ and $\delta A_\mu(x) = \partial_\mu\alpha(x)$. The change in the action can be written in the form

$$\delta S = \int d^4x \left\{ i\alpha\frac{\delta S}{\delta\phi(x)}\phi(x) - i\alpha\frac{\delta S}{\delta\phi^*(x)}\phi^*(x) + \frac{\delta S}{\delta A_\mu(x)}\partial_\mu\alpha(x)\right\} = 0.$$

By the principle of stationary action the functional derivatives with respect to the fields $\phi(x)$ and $\phi^*(x)$ vanish. For the third term we define the current $J^\mu(x)$ through

$$\frac{\delta S}{\delta A_\mu(x)} = \sqrt{g(x)}J^\mu(x). \tag{3.14}$$

Using partial integration we obtain from local U(1) gauge invariance the conservation law

$$\frac{1}{\sqrt{g(x)}}\partial_\mu\left[\sqrt{g(x)}J^\mu(x)\right] = 0. \tag{3.15}$$

This can be seen as electromagnetic current conservation. To make things concrete we give the corresponding expression for the action in eq. (3.12), as obtained by variation of $A_\mu(x)$,

$$J^\mu(x) = -ig^{\mu\nu}(x)\left[\phi^*(x)\partial_\nu\phi(x) - \phi(x)\partial_\nu\phi^*(x)\right].$$

General coordinate invariance

The action in eq. (3.12) is also invariant under invertible general coordinate transformations, $x^\mu \to x'^\mu(x)$. The scalar fields transform like

$$\phi(x) \to \phi'(x') = \phi(x(x')),$$

which implies for its derivatives

$$\frac{\partial}{\partial x^\mu}\phi(x) \to \frac{\partial}{\partial x'^\mu}\phi'(x') = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu}\phi(x(x')).$$

A similar transformation behavior is needed for the external gauge field,

$$A_\mu(x) \to A'_\mu(x') = \frac{\partial x^\nu}{\partial x'^\mu}A_\nu(x(x')),$$

and for the metric,

$$g_{\mu\nu}(x) \to g'_{\mu\nu}(x') = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu}g_{\rho\sigma}(x(x')).$$

Using the Jacobi determinant one finds that $d^4x\sqrt{g}$ is a covariant spacetime volume element. Also, for the inverse metric this implies

$$g^{\mu\nu}(x) \to g'^{\mu\nu}(x') = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma} g^{\rho\sigma}(x(x')).$$

Combining terms we find that the action (3.12) is indeed invariant under general coordinate transformations.

Infinitesimal general coordinate transformations

In an action as in eq. (3.12) the coordinates x^μ are just integration variables. One may therefore label them from x'^μ back to x^μ after the coordinate transformation. For the metric this leads to the transformation rule

$$g_{\mu\nu}(x) \to g'_{\mu\nu}(x) = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma}(x) - \left[g'_{\mu\nu}(x') - g'_{\mu\nu}(x)\right].$$

We now also specialize to infinitesimal coordinate transformations, $x'^\mu = x^\mu - \varepsilon^\mu(x)$. This implies for the metric

$$g_{\mu\nu}(x) \to g_{\mu\nu}(x) + \mathcal{L}_\varepsilon g_{\mu\nu}(x),$$

where the Lie derivative of the metric is defined as

$$\mathcal{L}_\varepsilon g_{\mu\nu}(x) = \varepsilon^\rho(x)\partial_\rho g_{\mu\nu}(x) + g_{\rho\nu}(x)\partial_\mu\varepsilon^\rho(x) + g_{\mu\rho}(x)\partial_\nu\varepsilon^\rho(x).$$

Similarly, the external gauge field transforms $A_\mu(x) \to A_\mu(x) + \mathcal{L}_\varepsilon A_\mu(x)$ where the Lie derivative of a one-form field is given by

$$\mathcal{L}_\varepsilon A_\mu(x) = \varepsilon^\nu(x)\partial_\nu A_\mu(x) + A_\nu(x)\partial_\mu\varepsilon^\nu(x).$$

Finally, the scalar field transforms as $\phi(x) \to \phi(x) + \mathcal{L}_\varepsilon\phi(x)$ with the Lie derivative of a scalar field defined as

$$\mathcal{L}_\varepsilon\phi(x) = \varepsilon^\mu(x)\partial_\mu\phi(x).$$

Formulated in this way, general coordinate transformations resemble closely other local symmetry transformations like the local U(1) gauge transformations discussed above.

Energy-momentum conservation from general coordinate invariance

We also know that the action must be invariant under general coordinate transformations. For an infinitesimal transformation we can write

$$\delta S = \int d^4x \left\{ \frac{\delta S}{\delta\phi(x)}\varepsilon^\mu(x)\phi(x) + \frac{\delta S}{\delta\phi^*(x)}\varepsilon^\mu(x)\phi^*(x) + \frac{\delta S}{\delta A_\mu(x)}\left[\varepsilon^\nu(x)\partial_\nu A_\mu(x) + A_\nu(x)\partial_\mu\varepsilon^\nu(x)\right] \right.$$
$$\left. + \frac{\delta S}{\delta g_{\mu\nu}(x)}\left[\varepsilon^\rho(x)\partial_\rho g_{\mu\nu}(x) + g_{\rho\nu}(x)\partial_\mu\varepsilon^\rho(x) + g_{\mu\rho}(x)\partial_\nu\varepsilon^\rho(x)\right] \right\} = 0.$$

Again the first two terms on the right hand side vanish when the equation of motion for $\phi(x)$ is fulfilled. We use now the definition (3.14) and also define the energy-momentum tensor $T^{\mu\nu}(x)$ through

$$\frac{\delta S}{\delta g_{\mu\nu}(x)} = \frac{1}{2}\sqrt{g(x)}T^{\mu\nu}(x).$$

It is symmetric by definition, $T^{\mu\nu}(x) = T^{\nu\mu}(x)$. The change in action can now be written as

$$\delta S = \int d^4x\sqrt{g(x)}\Big\{ J^\mu(x)\left[\varepsilon^\nu(x)\partial_\nu A_\mu(x) + A_\nu(x)\partial_\mu\varepsilon^\nu(x)\right]$$
$$+ \frac{1}{2}T^{\mu\nu}(x)\left[\varepsilon^\rho(x)\partial_\rho g_{\mu\nu}(x) + g_{\rho\nu}(x)\partial_\mu\varepsilon^\rho(x) + g_{\mu\rho}(x)\partial_\nu\varepsilon^\rho(x)\right] \Big\} = 0.$$

For the last term in the first line we perform a partial integration and use the current conservation law (3.15). This leads to a term involving the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. For the second and third term in the last line we also perform partial integrations and use the symmetry of the energy-momentum tensor. This yields, up to boundary terms,

$$\delta S = \int d^4x \sqrt{g(x)}\, \varepsilon^\rho(x) \left\{ F_{\rho\mu}(x) J^\mu(x) - \frac{1}{\sqrt{g(x)}} \partial_\mu \left[\sqrt{g(x)} T^\mu{}_\rho(x) \right] + \frac{1}{2} T^{\mu\nu}(x) \partial_\rho g_{\mu\nu}(x) \right\} = 0.$$

Because $\varepsilon^\rho(x)$ is arbitrary the term in curly brackets must vanish. For vanishing field strength of the external gauge field, $F_{\mu\nu}(x) = 0$, and constant Minkowski space metric, $g_{\mu\nu}(x) = g_{\mu\nu}$, this reduces to the standard energy-momentum conservation law in Cartesian coordinates, $\partial_\mu T^{\mu\nu}(x) = 0$. In summary, energy-momentum conservation can be seen as a consequence of general coordinate invariance of the action in terms of Noether's second theorem.

Energy-momentum tensor for complex scalar field in general coordinates

Let us finally calculate the energy-momentum tensor for the action in eq. (3.12) by variation of the metric. Varying only the metric but keeping the scalar field $\phi(x)$ fixed, and setting the gauge field $A_\mu(x)$ to zero for simplicity, we find

$$\delta S = \int d^4x \left[\delta\sqrt{g}\right] \left\{ -g^{\mu\nu} \partial_\mu \phi^* \partial_\nu \phi - V(\phi^*\phi) \right\} + \int d^4x \sqrt{g} \left\{ - \left[\delta g^{\mu\nu}\right] \partial_\mu \phi^* \partial_\nu \phi \right\}$$

We need the formulas

$$\delta\sqrt{g} = \frac{1}{2} \sqrt{g} g^{\mu\nu} \delta g_{\mu\nu}, \qquad \delta g^{\mu\nu} = -g^{\mu\rho} g^{\nu\sigma} \delta g_{\rho\sigma}.$$

Writing then

$$\delta S = \int d^4x \sqrt{g} \left\{ \frac{1}{2} T^{\mu\nu} \delta g_{\mu\nu} \right\},$$

and comparing terms, leads to the energy-momentum tensor of a complex scalar field

$$T^{\mu\nu} = 2 g^{\mu\rho} g^{\nu\sigma} \partial_\rho \phi^* \partial_\sigma \phi - g^{\mu\nu} \left[g^{\rho\sigma} \partial_\rho \phi^* \partial_\sigma \phi + V(\phi^*\phi) \right]. \tag{3.16}$$

Specializing to the zero-zero component we find the energy density

$$T^{00} = \mathscr{H} = \dot{\phi}^* \dot{\phi} + \boldsymbol{\nabla} \phi^* \boldsymbol{\nabla} \phi + V(\phi^*\phi).$$

LECTURE 06

4 Classical statistical field theory

Classical field theories are deterministic in the sense that one may fix initial data in the form of the field $\phi(t, \mathbf{x})$ and its first time derivative or conjugate momentum $\pi(t, \mathbf{x}) = \dot{\phi}(t, \mathbf{x})$ on some Cauchy surface, and the equation of motion (3.3) then fixes it everywhere. However, there are situations where the field configuration is not known precisely, but only stochastic information is available. We will now discuss such situations in some detail. This is interesting by itself but also serves as an excellent technical training for quantum field theory.

4.1 Static probabilistic description

Boltzmann probability weights

Recall from statistical mechanics that for a classical system in the canonical ensemble the probability density for a microstate to be in a phase space region $d^N q d^N p$ with coordinates $\mathbf{q}$ and conjugate momenta $\mathbf{p}$ is given by

$$Q(\mathbf{q},\mathbf{p}) = \frac{1}{Z}\exp\left(-\beta H(\mathbf{q},\mathbf{p})\right),$$

where $H(\mathbf{q},\mathbf{p})$ is the Hamiltonian, $\beta = 1/(k_{\mathrm{B}}T)$ is the inverse of temperature times the Boltzmann constant k_{B}, and Z is the partition function. The latter has to be fixed such that the probability distribution is properly normalized,

$$Z = \int d^N q d^N p \exp\left(-\beta H(\mathbf{q},\mathbf{p})\right).$$

In thermal equilibrium one can determine many different observables by taking averages with respect to the probability distribution $Q(\mathbf{q},\mathbf{p})$. For example, the expectation value of energy would be given by

$$E = \langle H\rangle = \int d^N q d^N p\, Q(\mathbf{q},\mathbf{p})\, H(\mathbf{q},\mathbf{p}).$$

We are here interested in generalizing this probabilistic description to field theories.

Infinite number of degrees of freedom

When considered from a mechanics point of view, a field theory has one degree of freedom per space point, so formally infinetly many. The infinity arises here in fact for two reasons:

- Points are dense in space / space is a contiuum.
- The space $\mathbb{R}^3$ we consider has infinite volume.

As we will see in due course, these two kinds of infinities lead to all kind of interesting consequences and differences to quantum mechanics for a finite number of degrees of freedom. In order to make progress it is oftentimes needed to regularize the theory. We introduce now a first regularization scheme, although there are many more.

Lattice regularization

A lattice regularization, which is also often used for numerical calculations, consists of two steps:

- Space is being discretized by considering points on a lattice.
- The volume is made finite by restricting it to a box, typically with periodic boundary conditions.

To recover the original theory from this regularized theory we need to study two limits:

- The contiuum limit where the lattice spacing goes to zero.
- The infinite volume limit where the box size becomes large.

For the technical steps we restric ourselves for simplicity to a single space dimension, the generalization to three spatial dimensions is straightforward.

Consider the chain of points

$$x_j \in \{0, \varepsilon, 2\varepsilon, \ldots, (N-1)\varepsilon\},$$

with the periodicity condition that $x_N = N\varepsilon$ is again the point $x_0 = 0$. One may visualize this as a ring with circumference $L = N\varepsilon$. The length ε corresponding to the distance between neighboring points is known as the *lattice spacing*. The contiuum limit corresponds to $\varepsilon \to 0$, while the infinite volume limit corresponds to $L \to \infty$.

Let us now consider the real scalar field on this discretized space. The "mechanical" degrees of freedom are essentially the N field values

$$\phi(t, x_j) = \phi(t, j\varepsilon).$$

We also need spatial derivatives, which get discretized according to

$$\frac{\partial}{\partial x}\phi(t,x) \to \frac{\phi(t,x_{j+1}) - \phi(t,x_j)}{\varepsilon}.$$

Blackboard video

Discretized Hamiltonian

We leave the details as an excercise and give here just the regularized form of the Hamiltonian for the real scalar field in one spatial dimension,

$$H = \sum_{j=1}^{N} \varepsilon \left\{ \frac{1}{2}\pi(t,x_j)^2 + \frac{1}{2}\left(\frac{\phi(t,x_{j+1}) - \phi(t,x_j)}{\varepsilon} \right)^2 + V(\phi(t,x_j)) \right\}.$$

Blackboard video

Probability distribiution and partition function

For the discretized theory we can immediately write down the thermodynamic equilibrium probability distribution for a field configuration specified by the $2N$ numbers $\phi(t, x_j)$, $\pi(t, x_j)$ with $j = 1, \ldots, N$ at some given time t. In other words, this is a probabilty density for a given field configuration to be in the $2N$ dimensional infinitesimal phase-space volume element

$$\prod_{j=1}^{N} \{d\phi(x_j) d\pi(x_j)\},$$

and it is given by

$$p[\phi, \pi] = \frac{1}{Z} e^{-\beta H}. \tag{4.1}$$

The normalization factor is here the partition function, given as a $2N$ dimensional integral,

$$Z = \prod_{j=1}^{N} \left\{ \int_{-\infty}^{\infty} d\phi(x_j) \int_{-\infty}^{\infty} d\pi(x_j) \right\} e^{-\beta H}.$$

We introduced here the Boltzmann weight $e^{-\beta H}$ with (discretized) Hamiltonian H and inverse temperature $\beta = 1/T$ (in units where $k_B = 1$). The partition function is a "sum", or actually an integral, over the possible field configurations at the given time t weighted with the Boltzmann factor.

Blackboard video

4.2 Functional integrals

Functional integral

The functional integral over fields is formally defined from the continuum and infinite volume limit of the phase-space integral

$$\int D\phi = \lim_{\varepsilon\to 0, L\to\infty} \prod_{j=1}^{N} \int_{-\infty}^{\infty} d\phi(x_j),$$

and similar for the conjugate momenta. With this we can write the partition function as

$$Z = \int D\phi D\pi \, e^{-\beta H[\phi,\pi]}.$$

At this point we can at least formally again work with the contiuum version of the field theoretic Hamiltonian H in (3.8).

A remark is in order at this point: for situations where the Hamiltonian contains terms of higher order than quadratic in the fields (which is the case for the Hamiltonian in (3.8) when $\lambda > 0$) the continuum limit needed to define the functional integral is more involved than we have described here. The short distance regularization can only be removed (by letting ε go to zero) if the theory is at the same time *renormalized*. We will discuss renormalization later on. For the time being take the above to be a formal definition of the functional integral.

Blackboard video

Expectation values

Of particular interest are observables $A[\phi]$ that depend on the field ϕ but not the conjugate momenta π, for example products of field values at different positions. Such expectation values can then be calculated as

$$\langle A[\phi]\rangle = \frac{1}{Z}\int D\phi D\pi A[\phi]\, e^{-\beta H[\phi,\pi]}.$$

A first example would be the field expectation value $\langle\phi(t,\mathbf{x})\rangle$, another the correlation function of fields at different spatial positions $\langle\phi(t,\mathbf{x})\phi(t,\mathbf{y})\rangle$.

Blackboard video

Scaling the partition function

Consider an additative change in the Hamiltonian of the form

$$\beta H[\phi,\pi] \to \beta H[\phi,\pi] + C,$$

where C is independent of the fields. This changes the partition function by a factor,

$$Z \to e^{-C} Z,$$

but does not change expectation values like $\langle A[\phi]\rangle$ because the factor cancels in the ratio! It can even happen that terms like C diverge such that formally $Z \to \infty$ or $Z \to 0$, but this is not a problem because the absolute value of Z is irrelvant. The probability density in (4.1) is not modified by this transformation.

Blackboard video

Integrating out the conjugate momenta

Note that the partition function separates into two factors, one involving the conjugate momenta, and one the actual fields. In the discrete version, the functional integral over the conjugate momenta is simply an N-dimensiona product of Gaussian integrals,

$$\prod_{j=1}^{N} \left\{ \int_{-\infty}^{\infty} d\pi(x_j) e^{-\beta \frac{\varepsilon}{2} \pi(x_j)^2} \right\}.$$

These integrals are easily performed and we obtain just a factor

$$\left(\frac{2\pi}{\beta \varepsilon} \right)^{N/2}.$$

This is in particular independent of the field ϕ and can therefore be dropped according to the argument above. It remains to work with the functioal integral over the actual fields ϕ and a reduced Boltzmann weight where the Hamiltonian involves just the potential energy $H_{\rm pot}[\phi]$.

Blackboard video

Euclidean action

The exponent of the Boltzmann weight factor is, mainly for historic reasons, also often called *Euclidean action* and denoted by $S[\varphi]$. For example we have for the real scalar field

$$S[\varphi] = \beta H_{\rm pot} = \int d^3x \left\{ \frac{1}{2} \partial_j \varphi \partial_j \varphi + V(\varphi) \right\}.$$

We have rescaled the fields by a factor, $\varphi = \sqrt{\beta}\phi$, and adapted the definition of the potential $V(\varphi)$ accordingy, such that the coefficient of the spatial derivative term becomes $1/2$. This is a common convention. We are also using Einsteins summation convention where j is summed from 1 to 3.

As all fields are evaluted at a single instance in time t we can drop this time argument as long as we are interested in classical thermal equilibrium situations, and work with fields $\varphi(\mathbf{x})$. In the following we will introduce a somewhat larger class of field theoretic models.

Blackboard video

4.3 O(N) models

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.

Blackboard video

Scalar O(N) models in d dimensions

Let us consider models of the form

$$S[\phi] = \int d^d x \left\{ \frac{1}{2} \partial_j \phi_n \partial_j \phi_n + \frac{1}{2} m^2 \phi_n \phi_n + \frac{1}{8} \lambda \left(\phi_n \phi_n \right)^2 \right\}. \tag{4.2}$$

Here, $\phi_n = \phi_n(\mathbf{x})$ with $n = 1, \ldots, N$ 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 and $d = 4$ will correspond to a quantum field theory after Wick rotation to Euclidean space). The index j is accordingly summed in the range $j = 1, \ldots, d$.

Blackboard video

Applications

Models of the type (4.2) have many applications. For $N = 1$ they correspond 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. Scalar fields are also used in cosmology, for example for the inflaton, or in nuclear physics, for example to describe pions.

Blackboard video

Engineering dimensions

In equation (4.2) 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 (4.2). 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.

Blackboard video

Symmetries

It is useful to discuss the symmetries of the model (4.2). 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 (4.2) 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

$$x^j \to x'^j = R^{jk} x^k.$$

The matrices R fulfill the condition $R^T R = \mathbb{1}$ and we demand that they connect continuously 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

$$R^{jk} = \delta^{jk} + \frac{i}{2}\delta\omega_{mn}\, J^{jk}_{(mn)},$$

where $J^{jk}_{(mn)} = -i(\delta_{mj}\delta_{nk} - \delta_{mk}\delta_{nj})$ 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

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

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 $\mathbf{x}$ before the transformation will have a maximum at $R\mathbf{x}$ afterwards. The field must transform as

$$\phi_n(\mathbf{x}) \to \phi'_n(\mathbf{x}) = \phi_n(R^{-1}\mathbf{x}).$$

Note that derivatives transform as

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

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 (4.2) is invariant under rotations acting on the fields, as one can confirm easily. Colloquially speaking, no direction in space is singled out.

Blackboard video

Translations

Another useful symmetry transformations are translations $\mathbf{x} \to \mathbf{x} + \mathbf{a}$. The fields get transformed as

$$\phi_n(\mathbf{x}) \to \phi'_n(\mathbf{x}) = \phi_n(\mathbf{x} - \mathbf{a}).$$

One easily confirms that the action (4.2) is also invariant under translations. Colloquially speaking, this implies that no point in space is singled out.

Blackboard video

Global internal $O(N)$ transformations

There is another useful symmetry of the action (4.2) given by rotations (and mirror reflections) in the "internal" space of fields,

$$\phi_n(\mathbf{x}) \to O_{nm}\phi_m(\mathbf{x}).$$

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 (4.2) is indeed invariant under these transformations.

Blackboard video

LECTURE 07

Partition function and functional derivatives

The partition function for the model (4.2) reads

$$Z[J] = \int D\phi \; e^{-S[\phi] + \int d^d x \{J_n(\mathbf{x})\phi_n(\mathbf{x})\}} \tag{4.3}$$

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

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

and correlation functions, e. g.

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

Blackboard video

Classical field equation

In the the functional integral the contribution of field configurations $\phi(\mathbf{x})$ is suppressed if the corresponding action $S[\phi]$ is large. In the partition function (4.3), 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

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

This equation is the field equation or equation of motion of a classical field theory. For the model (4.2) one has concretely

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

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). This field equation is the correspondence of Maxwells equations in electrodynamics for our scalar theory. The source J corresponds to the electromagnetic current in Maxwell's equations.

Blackboard video

The $O(N)$ symmetric potential

The model in (4.2) can be generalized somewhat to the action

$$S[\phi] = \int d^d x \left\{ \frac{1}{2}\partial_j \phi_n \partial_j \phi_n + V(\rho) \right\}, \tag{4.4}$$

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. The previous case (4.2) 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,

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

If the minimum is positive, $\rho_0 > 0$, the linear term vanishes of course, and one takes $m^2 = 0$. In contrast, if the minimum is at $\rho_0 = 0$ one has in general $m^2 > 0$. In practice, one uses either ρ_0 or m^2 for a parametrization of $V(\rho)$. 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.

Blackboard video

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 $J_n(\mathbf{x}) = 0$, and the model (4.4) we need to solve

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

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.

Blackboard video

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 $\lambda \to \infty$ with fixed ρ_0 and can be implemented as a constraint

$$\phi_n(\mathbf{x})\phi_n(\mathbf{x}) = 2\rho_0. \tag{4.5}$$

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 (the naming conventions are historic, one should not confuse the fields π_j with conjugate momentum fields)

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

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

$$\sigma = \sqrt{2\rho_0 - \vec{\pi}^2}.$$

Blackboard video

Linearly and non-linearly realized 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

$$\pi_n \to O^{(N-1)}_{nm}\pi_m, \qquad \sigma \to \sigma.$$

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

$$\delta\pi_n = \delta\alpha_n\sigma = \delta\alpha_n\sqrt{2\rho_0 - \vec{\pi}^2}, \qquad \delta\sigma = -\delta\alpha_n\pi_n,$$

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*.

Blackboard video

Action

Let us now write an action for the non-linear sigma model. Because of the constraint (4.5), the effective potential term in (4.4) becomes irrelevant and only the kinetic term remains,

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

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

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

The second term originates from

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

Blackboard video

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

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

Only in the presence of the determinant $\det(G(\vec{\pi}))$ the functional measure preserves the $O(N)$ symmetry. Accordingly, the functional integral for the non-linear sigma model must be adapted to contain the factor $\det(G(\vec{\pi}))$.

Ising model

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

4.4 Gaussian functional integrals and perturbation theory

Gaussian integrals

We now want to develop methods to actually eveluate functional integrals and to calculate correlation functions. We digress for a moment and consider Gaussian integrals of the type

$$\int_{\mathbb{R}^N} d^N \varphi \left\{ \exp\left(-\frac{1}{2} \varphi_j K_{jk} \varphi_k + J_k \varphi_k \right) \right\},$$

where indices j and k are summed in the range $1, \ldots, N$. The integral is here an infinite volume integral in N real dimensions weighted by a Gaussian function. We need to assume that the real part of the (symmetric) matrix K is positive definite, in the sense that the eigenvalues of $\mathrm{Re}(K_{jk})$ are positive. With some eigenvector v_k this implies

$$v_j^* \mathrm{Re}(K_{jk}) v_k = \mathrm{Re}(v_j^* K_{jk} v_k) = \mathrm{Re}(\lambda) v_k^* v_k > 0.$$

The source J_k is not restricted and can be complex. We want to show

$$\int_{\mathbb{R}^N} d^N\varphi \left\{ \exp\left(-\frac{1}{2}\varphi_j K_{jk}\varphi_k + J_k\varphi_k \right) \right\} = \frac{(2\pi)^{N/2}}{\sqrt{\det K}} \exp\left(\frac{1}{2} J_j (K^{-1})_{jk} J_k \right). \tag{4.6}$$

Blackboard video

Proof

The proof is done in three steps:

1. Assume first that K is real and $J = 0$. Then one can find an orthogonal Matrix O with unit determinant such that
$$K = O^T \Lambda O,$$
with $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$ a diagomal matrix with real positive entries. One can substitute integration variables $d^N\varphi \to d^N y$, where $y_j = O_{jk}\varphi_k$ because the Jacobi determinant is unity here. That implies
$$\int d^N\varphi \exp\left(-\frac{1}{2} x_j K_{jk}\varphi_k \right) = \int d^N y \exp\left(-\frac{1}{2}\sum_k \lambda_k y_k^2 \right) = \prod_{k=1}^N \left\{ \lambda_k^{-1/2} \int dx \exp\left(-\frac{1}{2}x^2 \right) \right\},$$
where we did another variable substitution $x = \sqrt{\lambda_k} y_k$ in the last step. Now one uses
$$\int_{-\infty}^{\infty} dx \exp\left(-\frac{1}{2}x^2 \right) = \sqrt{2\pi},$$
and
$$\det(K) = \det(\Lambda) = \prod_{k=1}^N \lambda_k,$$
which proves our formula in this special case.

2. Now consider real K and real J. Completing the square gives
$$\exp\left(-\frac{1}{2}\varphi^T K\varphi + J^T\varphi \right) = \exp\left(\frac{1}{2} J^T K^{-1} J \right) \exp\left(-\frac{1}{2}(\varphi - K^{-1}J)^T K (\varphi - K^{-1}J) \right),$$
and the integral over the second term gets reduced to what we have done before with a shift of integration variables.

3. Finally, the result can be extended to complex K and complex J (with the restriction that $\mathrm{Re}(K)$ has positive eigenvalues) by observing that the left and right hand sides of eq. (4.6) are holomorphic functions of K and J.

Gaussian integration can actually be extended to field theories and will be very useful for the following.

Blackboard video

Wick theorem

Now that we understand how to do Gaussian integrals we can also consider correlation functions of the type

$$G_{ij\cdots k} = \frac{1}{Z}\int\limits_{\mathbb{R}^N} d^N\varphi \left\{\varphi_i\varphi_j\cdots\varphi_k \exp\left(-\frac{1}{2}\varphi_j K_{jk}\varphi_k\right)\right\},$$

First we note that the number of φ_j insertions under the integral must be even, otherwise the integral must yield zero, as a result of the odd transformation behavior of the integrand with respect to reflections, $\varphi \to -\varphi$. To evaluate such integrals we can use the trick

$$G_{ij\cdots k} = \left(\frac{1}{Z[J]}\frac{\partial}{\partial J_i}\frac{\partial}{\partial J_j}\cdots\frac{\partial}{\partial J_k}Z[J]\right)_{J=0},$$

where

$$Z[J] = \int\limits_{\mathbb{R}^N} d^N\varphi \left\{\exp\left(-\frac{1}{2}\varphi_j K_{jk}\varphi_k + J_k\varphi_k\right)\right\} = \frac{(2\pi)^{N/2}}{\sqrt{\det K}}\exp\left(\frac{1}{2}J_j(K^{-1})_{jk}J_k\right)$$

is an extended version of the partition function. The prefactor

$$\frac{(2\pi)^{N/2}}{\sqrt{\det K}}$$

cancels out, so all we have to consider is the exponential

$$\exp\left(\frac{1}{2}J_j(K^{-1})_{jk}J_k\right).$$

Acting now with partial derivative operators brings down terms like $(K^{-1})_{ij}$. Recall that we need to set $J = 0$ at the end.

For example, the two point correlation function gives simply

$$G_{ij} = \langle\varphi_i\varphi_j\rangle = (K^{-1})_{ij},$$

and similarly, the four-point correlation function gives

$$\begin{aligned} G_{ijkl} = \langle\varphi_i\varphi_j\varphi_k\varphi_l\rangle =& (K^{-1})_{ij}(K^{-1})_{kl} + (K^{-1})_{ik}(K^{-1})_{jl} + (K^{-1})_{il}(K^{-1})_{jk} \\ =& \langle\varphi_i\varphi_j\rangle\langle\varphi_k\varphi_l\rangle + \langle\varphi_i\varphi_k\rangle\langle\varphi_j\varphi_l\rangle + \langle\varphi_i\varphi_l\rangle\langle\varphi_j\varphi_k\rangle. \end{aligned}$$

These are examples of a general relation: For a Gaussian (probability) weight one can calculate correlation functions by adding up all possible *contractions* which each contribute an term of the form of the two-point function or covariance matrix. This is known as *Wick's theorem*.

Blackboard video

Gaussian field theory

We now consider the $O(N)$ model in the quadratic regime, i. e. for $\lambda = 0$. We can write the partition function as

$$\begin{aligned} Z_2[J] &= \int D\phi\, \exp\left(-\frac{1}{2}\int_{\mathbf{x}}\{\phi_n(\mathbf{x})(-\partial_j\partial_j + m^2)\phi_n(\mathbf{x})\} + \int_{\mathbf{x}} J_n(\mathbf{x})\phi_n(\mathbf{x})\right) \\ &= \int D\phi\, \exp\left(-\frac{1}{2}\int_{\mathbf{x},\mathbf{y}}\{\phi_n(\mathbf{x})D_{nm}(\mathbf{x},\mathbf{y})\phi_m(\mathbf{y})\} + \int_{\mathbf{x}} J_n(\mathbf{x})\phi_n(\mathbf{x})\right) \\ &= \exp\left(\frac{1}{2}\int_{\mathbf{x},\mathbf{y}} J_n(\mathbf{x})G_{nm}(\mathbf{x},\mathbf{y})J_m(\mathbf{y})\right). \end{aligned}$$

Field expectation values follow as

$$\langle \phi_m(\mathbf{x}) \rangle = \frac{1}{Z_2} \frac{\delta}{\delta J_n(\mathbf{x})} Z_2[J] = \int_{\mathbf{y}} G_{nm}(\mathbf{x}, \mathbf{y}) J_m(\mathbf{y}).$$

Blackboard video

Two-point functions

Also two-point correlation functions of fields can be calculated within this Gaussian approximation easily,

$$\langle \phi_n(\mathbf{x}) \phi_m(\mathbf{y}) \rangle = \frac{1}{Z_2} \frac{\delta^2}{\delta J_n(\mathbf{x}) \delta J_m(\mathbf{y})} Z_2[J] = G_{nm}(\mathbf{x} - \mathbf{y}) + \langle \phi_n(\mathbf{x}) \rangle \langle \phi_m(\mathbf{y}) \rangle.$$

The two-point function decomposes into a product of expectation values and a connected correlation function or propagator

$$G_{mn}(\mathbf{x} - \mathbf{y}) = \langle \phi_n(\mathbf{x}) \phi_m(\mathbf{y}) \rangle_c = \langle \phi_n(\mathbf{x}) \phi_m(\mathbf{y}) \rangle - \langle \phi_n(\mathbf{x}) \rangle \langle \phi_m(\mathbf{y}) \rangle.$$

Note that $\langle \phi_m(\mathbf{x}) \rangle$ vanishes here when $J = 0$ but the correlation function contains then still the connected part.

Blackboard video

Greens function

The analog of the matrix K_{ij} in this field theoretic context is the operator (read as an infinite-dimensional matrix)

$$D_{nm}(\mathbf{x}, \mathbf{y}) = \delta_{nm} \left(-\frac{\partial}{\partial x^j} \frac{\partial}{\partial x^j} + m^2 \right) \delta^{(d)}(\mathbf{x} - \mathbf{y}).$$

This is also known as inverse propagator. We need to find its inverse, i. e. another integral operator $G_{nm}(\mathbf{x}, \mathbf{y})$ such that

$$\int_{\mathbf{y}} D_{mn}(\mathbf{x}, \mathbf{y}) G_{nk}(\mathbf{y}, \mathbf{z}) = \delta_{mk} \delta^{(d)}(\mathbf{x} - \mathbf{z}).$$

As a consequence of translational symmetry, G_{jk} is actually only a function of the difference of coordinates $\mathbf{x} - \mathbf{y}$. After partial integration we find the relation

$$\left(-\frac{\partial}{\partial x^j} \frac{\partial}{\partial x^j} + m^2 \right) G_{jk}(\mathbf{x} - \mathbf{y}) = \delta_{jk} \delta^{(3)}(\mathbf{x} - \mathbf{y}).$$

This shows that the so-called propagator $G_{jk}(\mathbf{x} - \mathbf{y})$ is actually a Greens function to the operator $(-\partial_j^2 + m^2)$. As usual, a Greens function can also depend on the boundary conditions which parametrize here the state of the theory in more detail.

Blackboard video

Solution in terms of Fourier transforms

For the ground state one can find the correct Greens function through Fourier transform. We write

$$G_{jk}(\mathbf{x}-\mathbf{y}) = \int \frac{d^d p}{(2\pi)^d} e^{i\mathbf{p}(\mathbf{x}-\mathbf{y})} G_{jk}(\mathbf{p}),$$

and obtain the simple relation for the Greens function in Fourier space,

$$G_{jk}(\mathbf{p}) = \frac{\delta_{jk}}{\mathbf{p}^2 + m^2}.$$

Blackboard video

LECTURE 08

Correlation function in position space

For $d = 3$ spatial dimensions, let us calculate the correlation function in position space. The integral can be written as

$$\begin{aligned} G_{jk}(\mathbf{x}-\mathbf{y}) =& \frac{1}{(2\pi)^3} \int d\Omega \int_0^\infty p^2 dp\, e^{ip|\mathbf{x}-\mathbf{y}|\cos(\vartheta)} \frac{\delta_{jk}}{p^2+m^2} \\ =& \frac{4\pi}{2(2\pi)^3} \int_{-1}^1 d\cos(\vartheta) \int_0^\infty p^2 dp\, e^{ip|\mathbf{x}-\mathbf{y}|\cos(\vartheta)} \frac{\delta_{jk}}{p^2+m^2} \\ =& \frac{1}{4i\pi^2|\mathbf{x}-\mathbf{y}|} \int_0^\infty dp\, p \left(e^{ip|\mathbf{x}-\mathbf{y}|} - e^{-ip|\mathbf{x}-\mathbf{y}|} \right) \frac{\delta_{jk}}{p^2+m^2}. \end{aligned}$$

The momentum integral can first be rewritten as an integral along the entire real line and one can then close the integration contour in the upper half of the complex plane,

$$\begin{aligned} G_{jk}(\mathbf{x}-\mathbf{y}) =& \frac{1}{4i\pi^2|\mathbf{x}-\mathbf{y}|} \int_{-\infty}^\infty dp\, p\, e^{ip|\mathbf{x}-\mathbf{y}|} \frac{\delta_{jk}}{p^2+m^2} \\ =& \frac{1}{4i\pi^2|\mathbf{x}-\mathbf{y}|} \oint dp\, p\, e^{ip|\mathbf{x}-\mathbf{y}|} \frac{\delta_{jk}}{(p+im)(p-im)}. \end{aligned}$$

Here one can use the residue theorem which tells that the integral is $2\pi i$ times the residue at $p = im$. The final result is

$$G_{jk}(\mathbf{x}-\mathbf{y}) = \frac{\delta_{jk}}{4\pi|\mathbf{x}-\mathbf{y}|} e^{-m|\mathbf{x}-\mathbf{y}|}.$$

We see that m has the effect of supressing correlations at large distances exponentially, in addition to an algebraic decay which is also there for $m = 0$. In fact,

$$\xi = \frac{1}{m},$$

is also known as the *correlation length* in the context of statistical field theory. We also note that $G_{jk}(\mathbf{x}-\mathbf{y})$ is divergent in the coincidence limit $|\mathbf{x}-\mathbf{y}| \to 0$, which corresponds to the region of large wavenumbers in Fourier space. This is known as an *ultraviolet* divergence. In concrete applications to condensed matter problems there is typically no such divergence but the model theory we have started with looses its physical significance for very high momenta or very short distances.

Blackboard video

Exercise

Determine the correlation function $G_{jk}(x-y)$ in $d=1$ spatial dimensions. Determine also the four-point correlation function $\langle\phi(x)\phi(y)\phi(z)\phi(w)\rangle$ for a single real scalar field $N=1$ in the absense of a source term, i. e. for $J=0$.

Complex scalar fields

Let us now consider a complex scalar field. We can understand it as two real fields combined in a special way,

$$\varphi(\mathbf{x}) = \frac{1}{\sqrt{2}}\left[\phi_1(\mathbf{x}) + i\phi_2(\mathbf{x})\right], \qquad \varphi^*(\mathbf{x}) = \frac{1}{\sqrt{2}}\left[\phi_1(\mathbf{x}) - i\phi_2(\mathbf{x})\right].$$

The prefactor is introduced just for convenience. One may now either take ϕ_1 and ϕ_2 as independent fields, or, equivalently, φ and φ^*. The functional integral can be written as

$$\int D\varphi D\varphi^* = \int D\phi_1 D\phi_2.$$

When working with complex fields it is convenient to work also with complex sources. A typical euclidean action for the complex fields reads

$$S[\varphi] = \int d^dx \left\{\partial_j\varphi^*\partial_j\varphi + m^2\varphi^*\varphi + \frac{\lambda}{2}(\varphi^*\varphi)^2\right\}.$$

The corresponding partition function is

$$Z[J] = \int D\varphi D\varphi^* e^{-S[\varphi]+\int_x\{J^*(\mathbf{x})\varphi(\mathbf{x})+J(\mathbf{x})\varphi^*(\mathbf{x})\}}.$$

Again, for a Gaussian approximation one can easily determine $Z_2[J]$.

Blackboard video

Perturbation theory

Let us now consider for simplicity the simple scalar theory ($N=1$) with an interaction term $(\lambda/4!)\phi^4$ and write the partition function formally in the form

$$\begin{aligned} Z[J] &= \int D\phi\, \exp\left(-\frac{\lambda}{4!}\int_{\mathbf{z}}\phi(\mathbf{z})\phi(\mathbf{z})\phi(\mathbf{z})\phi(\mathbf{z})\right)\exp\left(-S_2[\phi]+\int_{\mathbf{x}}J(\mathbf{x})\phi(\mathbf{x})\right) \\ &= \exp\left(-\frac{\lambda}{4!}\int_{\mathbf{z}}\frac{\delta}{\delta J(\mathbf{z})}\frac{\delta}{\delta J(\mathbf{z})}\frac{\delta}{\delta J(\mathbf{z})}\frac{\delta}{\delta J(\mathbf{z})}\right)\int D\phi\exp\left(-S_2[\phi]+\int_{\mathbf{x}}J(\mathbf{x})\phi(\mathbf{x})\right) \\ &= \text{const}\times\exp\left(-\frac{\lambda}{4!}\int_{\mathbf{z}}\frac{\delta}{\delta J(\mathbf{z})}\frac{\delta}{\delta J(\mathbf{z})}\frac{\delta}{\delta J(\mathbf{z})}\frac{\delta}{\delta J(\mathbf{z})}\right)\exp\left(\frac{1}{2}\int_{\mathbf{x},\mathbf{y}}J(\mathbf{x})G(\mathbf{x},\mathbf{y})J(\mathbf{y})\right). \end{aligned}$$

The two exponentials can be expanded into their Taylor series. Specifically the expansion of the first exponential leads to a perturbative series in the coupling constant λ. The constant term in the last line contains is independent of the source J but depends on temperature T so that we will have to partly take it into account for discussing thermodynamics.

Blackboard video

Feynman diagrams for the partition function

Terms in the perturbative series for the partition function $Z[J]$ and for correlation functions that are obtained as functional derivatives of $Z[J]$ have a nice graphical representation in terms of Feynman diagrams. This arises here in a purely classical statistical setup, but works very similar for quantum fields as we will see later on.

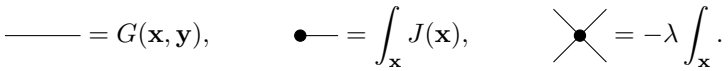
Expanding the two exponential functions in the partition function, one finds

$$Z[J] = \text{const} \times \sum_{V=0}^{\infty} \frac{1}{V!} \left(-\frac{\lambda}{4!} \int_{\mathbf{z}} \frac{\delta^4}{\delta J(\mathbf{z})^4} \right)^V \sum_{P=0}^{\infty} \frac{1}{P!} \left(\frac{1}{2} \int_{\mathbf{x},\mathbf{y}} J(\mathbf{x})\, G(\mathbf{x},\mathbf{y})\, J(\mathbf{y}) \right)^P , \tag{4.7}$$

where the index V can be understood as couting the number of four-vertices and P as counting the number of propagator lines. Once the functional derivatives associated with every vertex have been done, we have $2P - 4V$ powers of the source $J(\mathbf{x})$ left. When we later want to calculate correlation functions by taking functional derivatives of the partition functions, the functional derivatives act on these source terms.

Graphical representation

It is convenient to introduce a graphical representation for objects that appear in a systematic manner in the terms of the perturbative series (4.7). The three building blocks are composed of the propagator $G(\mathbf{x},\mathbf{y})$, the sources $J(\mathbf{x})$ and the four-vertex associated to the coupling constant λ. We introduce a graphical representation where propagators correspond to lines, sources to endpoints, and interaction terms to vertices,



The corresponding Feynman rules are rather simple for this model: each propagator gets attached with its two ends to a source or a vertex.

Blackboard video

Terms without vertices

As an exercise let us consider a few diagrams out of the infinite series generated by the expansion (4.7). Consider e.g. the term corresponding to $V = 0$ and $P = 1$,

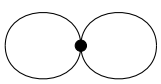
$$\frac{1}{2} \int_{\mathbf{x},\mathbf{y}} J(\mathbf{x})\, G(\mathbf{x},\mathbf{y})\, J(\mathbf{y}) \quad = \quad \frac{1}{2} \tag{4.8}$$

This has no vertices can can be seen as a contribution at lowest order λ^0 in the perturbative expansion. For $V = 0$ and larger values of P, we get products of these terms which give a factorized form of the partition function. The correlation functions following from this are obeying Wicks theorem.

Diagram with one vertex but no sources left

For $V = 1$ and $P = 2$ one obtains

$$-\frac{\lambda}{8} \int_{\mathbf{z}} G(\mathbf{z},\mathbf{z})^2 \quad = \quad \frac{1}{8} \tag{4.9}$$

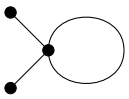
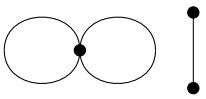


This is a diagram with two loops, but no source terms left, so it will not contribute to any correlation function. This is called a vacuum diagram.

Diagrams contributing to two-point function

For $V = 1$ and $P = 3$ there are two type of diagrams appearing,

$$-\frac{\lambda}{4}\int_{\mathbf{z}} G(\mathbf{z},\mathbf{z})\left[\int_{\mathbf{y}} G(\mathbf{z},\mathbf{y})\,J(\mathbf{y})\right]^2 - \frac{\lambda}{16}\int_{\mathbf{z}} G(\mathbf{z},\mathbf{z})^2 \int_{\mathbf{x},\mathbf{y}} J(\mathbf{x})\,G(\mathbf{x},\mathbf{y})\,J(\mathbf{y}) \tag{4.10}$$

$$= \frac{1}{4}$$  $$+ \frac{1}{16}$$ 

Here the first diagram gives a non-trivial and connected contribution to the two-point correlation function with one loop. In contrast, the second diagram falls into a product of two disconnected pieces with the first being a vacuum diagram and the second a contribution to the two-point function.

Different kinds of diagrams

In these three examples we already encountered two important classes of diagrams. There exist diagrams which do not include closed cycles or loops and are usually called *tree diagrams* whereas diagrams including closed cycles are call *loop diagrams*. Furthermore, there are diagrams as in (4.8), (4.9) and the first diagram in (4.10) which are *connected* whereas the second diagram in (4.10) is *disconnected*. A further distinction arises between *vacuum diagrams* and diagrams that actually contribute to correlation functions.

Blackboard video

Blackboard video

Divergences

At first sight, it might seem diagrams such as (4.9) contain divergencies associated to the coincidence limit of the propagator as well as an integral over the whole space. As we will realise later, these kind of diagrams (in QFT called 'vacuum bubbles') do not contribute to correlation functions. In contrast, loop diagrams such as the first diagram in (4.10) do contribute to correlation functions and the associated ultraviolet divergence from $G(\mathbf{z},\mathbf{z}) = \infty$ is handled later on by renormalisation.

Blackboard video

4.5 Generating functionals

Schwinger functional and connected diagrams

It is often useful to consider instead of the partition function $Z[J]$ the Schwinger functional $W[J]$ defined through

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

The expectation value is simply obtained through

$$\langle\phi(\mathbf{x})\rangle = \frac{\delta}{\delta J(\mathbf{x})}W[J] = \frac{1}{Z[J]}\frac{\delta}{\delta J(\mathbf{x})}Z[J].$$

The second functional derivative yields

$$\frac{\delta^2}{\delta J(\mathbf{x})\delta J(\mathbf{y})}W[J] = \frac{1}{Z[J]}\frac{\delta^2}{\delta J(\mathbf{x})\delta J(\mathbf{y})}Z[J] - \frac{1}{Z[J]^2}\frac{\delta}{\delta J(\mathbf{x})}Z[J]\frac{\delta}{\delta J(\mathbf{y})}Z[J]$$
$$= \langle\phi(\mathbf{x})\phi(\mathbf{y})\rangle - \langle\phi(\mathbf{x})\rangle\langle\phi(\mathbf{y})\rangle.$$

When $\mathbf{x}$ and $\mathbf{y}$ are far away from each other, the fields become typically uncorrelated such that the two terms on the right hand side agree. For this reason one calls the difference between them the connected part of the diagram and writes

$$\langle\phi(\mathbf{x})\phi(\mathbf{y})\rangle = \mathcal{G}(\mathbf{x},\mathbf{y}) + \langle\phi(\mathbf{x})\rangle\langle\phi(\mathbf{y})\rangle,$$

where the *connected correlation function* is given by

$$\langle\phi(\mathbf{x})\phi(\mathbf{y})\rangle_c = \frac{\delta^2}{\delta J(\mathbf{x})\delta J(\mathbf{y})}W[J] = \mathcal{G}(\mathbf{x},\mathbf{y}).$$

This connected correlation function the usually vanishes at large separation.

Blackboard video

Connected three-point correlation function

In a similar way one can now determine third functional derivatives,

$$\langle\phi(\mathbf{x})\phi(\mathbf{y})\phi(\mathbf{z})\rangle_c = \frac{\delta^3}{\delta J(\mathbf{x})\delta J(\mathbf{y})\delta J(\mathbf{z})}W[J] = \langle\phi(\mathbf{x})\phi(\mathbf{y})\phi(\mathbf{z})\rangle$$
$$- \langle\phi(\mathbf{x})\phi(\mathbf{y})\rangle\langle\phi(\mathbf{z})\rangle - \langle\phi(\mathbf{y})\phi(\mathbf{z})\rangle\langle\phi(\mathbf{x})\rangle - \langle\phi(\mathbf{z})\phi(\mathbf{x})\rangle\langle\phi(\mathbf{y})\rangle + 2\langle\phi(\mathbf{x})\rangle\langle\phi(\mathbf{y})\rangle\langle\phi(\mathbf{z})\rangle,$$

which is known as the connected three-point correlation function. Note that for vanishing expectation value, the connecetd three-point correlation function equals the full three-point correlation function.

Blackboard video

Connected four-point correlation function

The four point function can be worked out in a similar way. One finds

$$\langle\phi(\mathbf{x})\phi(\mathbf{y})\phi(\mathbf{z})\phi(\mathbf{w})\rangle_c = \frac{\delta^4 W}{\delta J(\mathbf{x})\delta J(\mathbf{y})\delta J(\mathbf{z})\delta J(\mathbf{w})} = \langle\phi(\mathbf{x})\phi(\mathbf{y})\phi(\mathbf{z})\phi(\mathbf{w})\rangle$$
$$- \langle\phi(\mathbf{x})\phi(\mathbf{y})\rangle\langle\phi(\mathbf{z})\phi(\mathbf{w})\rangle - \langle\phi(\mathbf{x})\phi(\mathbf{z})\rangle\langle\phi(\mathbf{y})\phi(\mathbf{w})\rangle$$
$$- \langle\phi(\mathbf{x})\phi(\mathbf{w})\rangle\langle\phi(\mathbf{y})\phi(\mathbf{z})\rangle + \text{terms involving } \langle\phi\rangle. \tag{4.11}$$

For vanishing expectation value, the connected four-point function subtracts from the four point function the "unconnected parts". In this way one can go on and decompose all correlation functions in connected and disconnected parts. It turns out that in many physics applications one actually is interested in connected terms only.

In the context of finite dimensional random variables, correlation functions are known as moments and connected correlation functions are known as cumulants.

Blackboard video

Thermodynamic significance

Recall that in our present context the partition function is for $J = 0$ and with $\beta = 1/T$

$$Z[0] = e^{W[0]} = e^{-\beta F(T)} = \mathrm{Tr}\left\{e^{-\beta H}\right\}.$$

Up to an additive constant the Schwinger functional at vanishing source is the free energy devided by the temperature,

$$W[0] = -F(T)/T = -(E - TS)/T.$$

From the free energy one can obtain for example the entropy according to $S = -\partial F/\partial T$ or the expectation value of energy as $E = F + TS$. Of course, to calculate this we need to follow carefully the dependence on temperature T. This can be generalized to situations with more conserved quantum numbers, such as some particle number N coupled to a chemical potential μ. We will later also study the generalization to quantum statistics.

Blackboard video

Partition function and Schwinger functional at vanishing source

We concentrate now first on the leading contribution at order λ^0 to the partition function and Schwinger functional at vanishing source. From the evaluation of the Gaussian integral we find (still neglecting the contribution of conjugate momenta here)

$$Z[0] = e^{W[0]} = \lim_{N\to\infty} (2\pi)^{N/2} \det(D)^{-1/2},$$

where D is the matrix inverse to the propagator function $G(\mathbf{x}, \mathbf{y})$. The factors of 2π can be dropped, they only lead to a numerical offset in $W[0]$ that is independent of temperature. To evaluate the determinant we use the identity

$$\ln(\det(D)) = \mathrm{tr}\{\ln(D)\}.$$

Blackboard video

Loop expansion for Schwinger functional

This leads to

$$W[0] = \mathrm{const} - \frac{1}{2}\mathrm{tr}\ln(D).$$

The second term can be graphically represented by a single closed loop. In a perturbative expansion we find that this gets supplemented by a two-loop term at order λ^1, two possible tree-loop terms at order λ^2 and so on,

$$W[0] = \mathrm{const} - \frac{1}{2}\mathrm{tr}\left\{\ln(D)\right\} + \mathcal{O}(\lambda) + \mathcal{O}(\lambda^2) + \ldots.$$

While it is straight forward to write down the connected vacuum diagrams and also to write down corresponding integral expressions, it will be more work to properly evaluate them.

In the partition function $Z[0] = \exp(W[0])$ we also get products of the connected diagrams through the expansion of the exponential and it is therefore given by a sum of all possible vacuum diagrams.

Blackboard video

Perturbative expansion for two- and four-point functions

We can now go ahead and consider terms in the Schwinger functional $W[J]$ that depend on the source. Because our theory is invariant under the Z_2 symmetry $\phi \to -\phi$, $J \to -J$, there are only even orders in J. We can write

$$W[J] = W[0] + \frac{1}{2}\int_{\mathbf{x},\mathbf{y}} J(\mathbf{x})\mathscr{G}(\mathbf{x},\mathbf{y})J(\mathbf{y}) + \frac{1}{4!}\int_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4} J(\mathbf{x}_1)J(\mathbf{x}_2)J(\mathbf{x}_3)J(\mathbf{x}_4)\mathscr{V}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4) + \dots,$$

where $\mathscr{G}(\mathbf{x},\mathbf{y})$ is a connected two-point correlation function including perturbative corrections, $\mathscr{V}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4)$ is a connected four-point correlation function including perturbative corrections and so on.

For the two-point function we find at order λ^0

$$\mathscr{G}(\mathbf{x},\mathbf{y}) = G(\mathbf{x}-\mathbf{y}),$$

and this gets corrected by a one-loop term at order λ^1, three different two-loop terms at order λ^2 and so on.

Similarly, the four-point function has the leading contribution (a tree diagram)

$$\mathscr{V}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4) = -\lambda\int d^dx\, G(\mathbf{x}_1-\mathbf{x})G(\mathbf{x}_2-\mathbf{x})G(\mathbf{x}_3-\mathbf{x})G(\mathbf{x}_4-\mathbf{x}),$$

and this gets supplemented by different one-loop terms at order λ, two-loop terms at order λ^2 and so on.

We will later introduce another generating functional, the one-particle irreducible effective action, which allows for an even more efficient organization of loop diagrams.

Blackboard video

4.6 Mean field theory and critical exponents

[More to come.]

4.7 Ultraviolet divergences and regularization

[More to come.]

4.8 Renormalization

[More to come.]

LECTURE 09

5 Quantum states

Let us now address quantum states in a quantum field theory. We start by recalling some notions and concepts from quantum mechanics of non-relativistic particles and will then transfer the different concepts to field theory.

5.1 Canonical quantization

A set of harmonic oscillators

[More to come.]

The non-relativistic scalar field

[More to come.]

Particles as quantum excitations

[More to come.]

Canonical commutation relations in field theory

In the continuum limit the real field $\phi(\mathbf{x})$ and the conjugate momentum field $\pi(\mathbf{x})$ represented by the operator

$$\pi(\mathbf{x}) = -i\frac{\delta}{\delta\phi(\mathbf{x})},$$

have the "canonical" commutation relation

$$[\phi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y}). \tag{5.1}$$

As usual this is expressed in units where $\hbar = 1$. This relation is the generalization of quantum mechanical commutation relation $[x_j, p_k] = i\delta_{jk}$. For free fields, which generalize in a sense the quantum mechanical harmonic oscillator, it is now interesting to introduce creation and annihilation operators by the relations

$$\begin{aligned}\phi(\mathbf{x}) &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_p}}\left(a_{\mathbf{p}}e^{i\mathbf{px}} + a^\dagger_{\mathbf{p}}e^{-i\mathbf{px}}\right),\\ \pi(\mathbf{x}) &= \int \frac{d^3p}{(2\pi)^3}\frac{\sqrt{E_{\mathbf{p}}}}{\sqrt{2}}\left(-ia_{\mathbf{p}}e^{i\mathbf{px}} + ia^\dagger_{\mathbf{p}}e^{-i\mathbf{px}}\right).\end{aligned} \tag{5.2}$$

Here we use the frequency of the momentum mode $\mathbf{p}$ given by

$$E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}.$$

In other words, $a_{\mathbf{p}}$ and $a^\dagger_{\mathbf{p}}$ are linear combinations of ϕ and π. The canonical commutation relations become then simply

$$[a_{\mathbf{p}}, a^\dagger_{\mathbf{q}}] = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q}), \qquad [a_{\mathbf{p}}, a_{\mathbf{q}}] = [a^\dagger_{\mathbf{p}}, a^\dagger_{\mathbf{q}}] = 0.$$

Blackboard video

Vacuum states

In the present formalism (for non-interacting fields) one can take the vaccum state $|0\rangle$ in the field theory to be such that

$$a_{\mathbf{p}}|0\rangle = 0,$$

for all momenta $\mathbf{p}$.

Exercise

(a) Check that the commutation relations for field and conjugate momenta imply those for creation and annhilation operators and *vice versa*. (b) For a non-interacting scalar quantum field theory, find the Schrödinger functional of the vaccum state.

Blackboard video

Single particle states

One can also construct now states for single particles in a momentum eigenstate by using the creation operator,

$$|\mathbf{p}\rangle \sim a^\dagger_{\mathbf{p}}|0\rangle.$$

We discuss the normalization and related issues later on. In a similar way one can construct two-particle states,

$$|\mathbf{p},\mathbf{q}\rangle \sim a^\dagger_{\mathbf{p}}a^\dagger_{\mathbf{q}}|0\rangle,$$

and as a consequence of the commutation relations it is autmotically symmetric,

$$|\mathbf{p},\mathbf{q}\rangle = |\mathbf{q},\mathbf{p}\rangle.$$

Blackboard video

Complex fields

It is also interesting to generalize these considerations to complex scalar fields where two sets of creation and annihilation operators naturally appear. These create and destroy particles and anti-particles, respectively (exercise).

Blackboard video

5.2 Pure and mixed states

The density matrix

Recall that in quantum mechanics for N particles, one can specify an arbitrary *pure state* at some time t in terms of a Schrödinger wave function

$$\psi_t(\mathbf{x}_1,\cdots,\mathbf{x}_N).$$

This corresponds to the *position space* representation of quantum mechanics. More abstractly one may work in terms of states in a Hilbert space,

$$|\psi_t\rangle.$$

A general mixed state needs to be described by a *density matrix* or a *density operator*. For a mixture of states $|\psi_j\rangle$ with probability p_j such that $\sum_j p_j = 1$, the density operator is given by

$$\rho_t = \sum_j p_j|\psi_j\rangle\langle\psi_j|.$$

The concept of a mixed state is needed if one does not know the state with certainty but has only a probailistic description available. Mixed states are also needed if one would like to describe degrees of freedom that are not fully isolated but *entangled* with other degrees of freedom. This is actually the general situation for the *local* description of a quantum field theory in some subvolume of space.

Blackboard video

Expectation values

From the density operator, one can calculate expectation values at time t as

$$\langle A(t)\rangle = \mathrm{Tr}\,\{\rho_t A\} = \sum_j p_j \mathrm{Tr}\,\{|\psi_j\rangle\langle\psi_j|A\} = \sum_j p_j \langle\psi_j|A|\psi_j\rangle.$$

For the concrete case of an N-particle state in quantum mechanics, one would have

$$\rho_t(\mathbf{x}_1,\cdots,\mathbf{x}_N;\mathbf{y}_1,\cdots,\mathbf{y}_N) = \sum_j p_j \psi_j(\mathbf{x}_1,\cdots,\mathbf{x}_N)\psi_j^*(\mathbf{y}_1,\cdots,\mathbf{y}_N).$$

An arbitrary operator can be written as

$$A(\mathbf{x}_1,\cdots,\mathbf{x}_N;\mathbf{y}_1,\cdots,\mathbf{y}_N)$$

in position space representation and the expectation value would be

$$\langle A(t)\rangle = \mathrm{Tr}\,\{\rho_t A\} = \int_{\mathbf{x}_1,\ldots,\mathbf{x}_N}\int_{\mathbf{y}_1,\ldots,\mathbf{y}_N} \rho_t(\mathbf{x}_1,\cdots,\mathbf{x}_N;\mathbf{y}_1,\cdots,\mathbf{y}_N)A(\mathbf{y}_1,\cdots,\mathbf{y}_N;\mathbf{x}_1,\cdots,\mathbf{x}_N).$$

Blackboard video

Momentum operator

As an example let us consider just a single particle, $N = 1$. We want to calculate the expectation value of the momentum component P_k. It corresponds to a derivative operator in the position space representation we use. In our notation it can be written as a distribution,

$$P_k(\mathbf{x},\mathbf{y}) = -i\frac{\partial}{\partial x^k}\delta^{(3)}(\mathbf{x}-\mathbf{y}).$$

With this one finds with a few steps involving partial integration

$$\langle P_k(t)\rangle = \int_{\mathbf{x},\mathbf{y}} \rho_t(\mathbf{x},\mathbf{y})\left[-i\frac{\partial}{\partial y^k}\delta^{(3)}(\mathbf{y}-\mathbf{x})\right] = \sum_j p_j \psi_j^*(\mathbf{x})\left[-i\frac{\partial}{\partial x^k}\psi_j(\mathbf{x})\right],$$

which is the expression familiar from quantum mechanics.

Blackboard video

Transition to field theory

Let us now go to quantum field theory. Instead of the positions $\mathbf{x}_1, \cdots, \mathbf{x}_N$, the defrees of freedom are now the field variables $\phi(\mathbf{x})$ at some fixed time t, for all possible spatial positons $\mathbf{x}$. The spatial position $\mathbf{x}$ now plays the role of the index $n = 1, \cdots, N$ and labels the different degrees of freedom (quantum fields).

Blackboard video

Pure and mixed states

A pure state at some time t is now specified by a so-called *Schrödinger functional*

$$\psi_t[\phi],$$

and a mixed state in a similar way by a *density matrix functional*

$$\rho_t[\phi_+, \phi_-].$$

The most general observable is also specified by a similar functional

$$A[\phi_1, \phi_2],$$

and an expectation value is given by

$$\langle A \rangle = \int D\phi_+ D\phi_- \, \rho_t[\phi_+, \phi_-] A[\phi_-, \phi_+].$$

The functional integrals $\int D\phi_+$ and $\int D\phi_-$ are here over fields at constant time t but for all spatial positions $\mathbf{x}$.

Blackboard video

Conjugate momentum field

In this "position space" representation of a field theoretic state, the conjugate momentum field corresponds to an operator,

$$\pi(t, \mathbf{x})[\phi_1, \phi_2] = \left[-i \frac{\delta}{\delta \phi_1(\mathbf{x})} \right] \delta[\phi_1 - \phi_2],$$

with the "functional Dirac delta" $\delta[\phi_1 - \phi_2]$ defined such that for some functional $f[\phi]$

$$\int D\phi_1 f[\phi_1] \delta[\phi_1 - \phi_2] = f[\phi_2].$$

With this one obtains the expection value

$$\langle \pi(t, \mathbf{x}) \rangle = \int D\phi \left[-i \frac{\delta}{\delta \phi_+(\mathbf{x})} \rho_t[\phi_+, \phi_-] \right]_{\phi_+ = \phi_- = \phi}.$$

In this way one can now calculate all kind of observables at some given time t.

Blackboard video

Field theory of a single mode

To get an intuition, let us consider the simple case of a field theory in $d = 1 + 0$ dimensions. This describes a single field mode and has applications for example to cavity-quantum electrodynamics. The Lagrangian for a real variable ϕ is

$$L = \frac{1}{2}(\partial_t \phi)^2 - \frac{1}{2}m^2\phi^2,$$

and it is equivalent to the harmonic oscillator. This means that we can easily take over some results known from quantum mechanics.

Blackboard video

5.3 Ground state and excited states

Ground state

The Schrödinger functional for the ground state is now simply

$$\psi_0[\phi] = ce^{-m\phi^2/2}$$

with a complex constant c. Accordingly, the density functional in that state

$$\rho_0[\phi_+, \phi_-] = \frac{1}{Z}e^{-m(\phi_+^2+\phi_-^2)/2}.$$

One can directly see that this is a pure state because it factorizes into a ket and a bra contribution.

Blackboard video

Excited states

Excited states with n particles or quanta are of the form

$$\rho_n[\phi_+, \phi_-] = \frac{1}{Z_n}H_n(\sqrt{m}\phi_+)H_n(\sqrt{m}\phi_-)e^{-m(\phi_+^2+\phi_-^2)/2}$$

where $H_n(x)$ are the Hermite polynomials

$$H_0(x) = 1, \qquad H_1(x) = 2x, \qquad H_2(x) = 4x^2 - 2, \cdots$$

These are still pure states. The corresponding Schrödinger functional would be

$$\psi_n[\phi] = \frac{1}{\sqrt{2^n n!}}H_n(\sqrt{m}\phi)ce^{-m\phi^2/2}.$$

Under time evolution, the Schrödinger functionals above would pick up a factor $e^{-im(n+1/2)t}$ which cancels, however, in the density functional.

Blackboard video

5.4 Coherent states

Coherent states

Another interesting class of states are coherent states. In quantum mechanics they are described by

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle = e^{-|\alpha|^2/2} e^{\alpha a^\dagger} |0\rangle,$$

with complex parameter α. Here they lead to the density matrix functional

$$\rho_\alpha[\phi_+, \phi_-] = \frac{1}{Z} \exp\left(-\frac{1}{2} m \left[\left(\phi_+ - \sqrt{\frac{2}{m}} \mathrm{Re}(\alpha) \right)^2 + \left(\phi_- - \sqrt{\frac{2}{m}} \mathrm{Re}(\alpha) \right)^2 \right] \right).$$

Again these are pure states. Under time eolution, one must replace $\alpha \to \alpha(t_0) e^{-im(t-t_0)}$ and one finds that $\mathrm{Re}(\alpha(t))$ describes the oscillatory behaviour of classical solutions to the equations of motion. The density matrix $\rho_{\alpha(t)}$ describes Gaussian fluctuations around this mean value.

Blackboard video

Themal states for a single mode

Finally, let us consider a thermal state. In the quantum mechanical formalism, it is described as

$$\rho = (1-b) \sum_{n=0}^{\infty} b^n |n\rangle\langle n|,$$

where $b = e^{-m/T}$ is the Boltzmann weight. Here this leads to the density matrix functional

$$\rho_T[\phi_+, \phi_-] = \frac{1}{Z}(1-b) \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{b}{2} \right)^n H_n(\sqrt{m}\phi_+) H_n(\sqrt{m}\phi_-) e^{-m(\phi_+^2 + \phi_-^2)/2}.$$

Here, one can use a property of the Hermite polynomials (Mehler's formula)

$$\sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) H_n(y) \left(\frac{b}{2} \right)^n = \frac{1}{\sqrt{1-b^2}} \exp\left[\frac{2b}{1+b} xy - \frac{b^2}{1-b^2} (x-y)^2 \right].$$

We thus find

$$\begin{aligned} \rho_T[\phi_+, \phi_-] &= \frac{1}{Z} \exp\left[-\frac{1}{2} m(\phi_+^2 + \phi_-^2) - \frac{b^2}{1-b^2} m(\phi_+ - \phi_-)^2 + \frac{2b}{1+b} m\phi_+\phi_- \right] \\ &= \frac{1}{Z} \exp\left[-\frac{1}{2} m \left(1 + \frac{2b^2}{1-b^2} \right) (\phi_+^2 + \phi_-^2) + \frac{2b}{1-b^2} m\phi_+\phi_- \right]. \end{aligned} \tag{5.3}$$

This does not factor into a ket and a bra part for $b > 0$. It is therefore not a pure state as expected.

Blackboard video

Gaussian versus non-Gaussian states

Let us summarize this discussion by remarking that the vacuum or ground state, the coherent states, as well as the thermal states all have density matrices $\rho[\phi_+, \phi_-]$ of Gaussian form. This is not the case for single or multiple particle excited states, though. For free quantum field theories, one can also expect Gaussian states in many circumstances. However, already with non-vanishing interaction this ceases to be the case.

Blackboard video

LECTURE 10

Higher dimensional Gaussian states

Let us now generalize the situation somewhat and consider a set of fields ϕ_n. The index n is here taken to be discrete and can run over a finite set of modes for example. However, it could be running over an infinite set. One may even consider n to be an abstract index that combines several indices such as momentum, flavor or spin.

We assume the Schrödinger functional to be of the form

$$\psi[\phi] = c \exp\left[-\frac{1}{2}(\phi - \bar{\phi})_m h_{mn} (\phi - \bar{\phi})_n + i j_n \phi_n\right],$$

with a symmetric and real matrix $h_{mn} = h_{nm}$. The density functional for this pure state is accordingly

$$\rho[\phi_+, \phi_-] = \frac{1}{Z} \exp\Big[-\frac{1}{2}(\phi_+ - \bar{\phi})_m h_{mn} (\phi_+ - \bar{\phi})_n \\ -\frac{1}{2}(\phi_- - \bar{\phi})_m h_{mn} (\phi_- - \bar{\phi})_n + i j_n (\phi_+ - \phi_-)_n \Big].$$

Blackboard video

Blackboard video

Characterization through one- and two-point functions

Let us characterize this state by its expectation values and correlation functions. Besides the field ϕ_n, another observable is its conjugate momentum field π_n. In the position space representation, we are working in here, it is represented by a derivative

$$\pi_n = -i \frac{\delta}{\delta \phi_n}.$$

This operator acts on the Schrödinger functional or density operator. The canonical commutation relations

$$[\phi_m, \pi_n] = i\delta_{mn}, \qquad [\phi_m, \phi_n] = [\pi_m, \pi_n] = 0,$$

are automatically fulfilled.

The field expectation value is given by

$$\langle \phi_m \rangle = \frac{1}{Z} \int D\phi \; \phi_m \rho[\phi, \phi] = \bar{\phi}_m.$$

In a similar way, the expectation value for the conjugate momentum can be obtained,

$$\langle \pi_m \rangle = \frac{1}{Z} \int D\phi \left(-i \frac{\delta}{\delta \phi_{+m}} \rho[\phi_+, \phi_-] \right)_{\phi_+ = \phi_- = \phi} = j_m.$$

An exercise in Gaussian integration yields the connected correlation functions

$$\langle \phi_m \phi_n \rangle_c = \langle \phi_m \phi_n \rangle - \langle \phi_m \rangle \langle \phi_n \rangle = \frac{1}{2} (h^{-1})_{mn}, \qquad \langle \pi_m, \pi_n \rangle_c = \frac{1}{2} h_{mn},$$
$$\langle \phi_m \pi_n + \pi_n \phi_m \rangle_c = 0, \qquad \langle \phi_m \pi_n - \pi_n \phi_m \rangle_c = [\phi_m, \pi_n] = i\delta_{mn}.$$

If the matrix h_{mn} is diagonal $h_{mn} = \tilde{h}_m \delta_{mn}$ (no sum convention), the different field modes are independent, otherwise they are correlated.

Blackboard video

Uncertainty relation

Imagine now that h_{mn} is diagonal. One then has

$$\langle \phi_m^2 \rangle \langle \pi_m^2 \rangle = \frac{1}{4}.$$

This is in fact the statement that Heisenberg's uncertainty bound is satisfied and saturated. Note that for a single mode in the ground state, we have

$$\langle \phi^2 \rangle = \frac{1}{2m}, \qquad \langle \pi^2 \rangle = \frac{m}{2}.$$

The energy $E = m$ here sets the quantum uncertainty. In quantum optics, it is possible, however, to prepare so-called squeezed states with

$$\langle \phi^2 \rangle = \frac{1}{2h}, \qquad \langle \pi^2 \rangle = \frac{h}{2},$$

where $h > m$ or $h < m$. These are still pure states and they are still Gaussian states. They also still satisfy the Heisenberg bound but, for $n > m$, have a reduced uncertainty of the field at the cost of an increased uncertainty of the conjugate momentum. For $n < m$, the uncertainty of π is reduced while the one of ϕ is increased.

For diagonal h_{mn}, the different modes ϕ_m are fully independent and the density matrix $\rho[\phi_+, \phi_-]$ decomposes into a product of independent factors. This indicates that these degrees of freedom are not entangled. The situation is different in the presence of off-diagonal terms in h_{mn}. In that case, there are non-vanishing correlations between fields and between conjugate momenta - but there is also *quantum entanglement*.

When quantum field theory is developed from a version of the theory with lattice regularization one finds that the field degrees of freedom in position space are strongly entangled by the spatial derivative term. For a quantum field theory entanglement is in this sense ubiquitous.

Blackboard video

5.5 Entanglement

Two-mode squeezed state

As the simplest example for an entangled Gaussian state consider the two-mode squeezed state with Schrödinger functional

$$\psi_r[\phi_1, \phi_2] = c \exp\left[-\frac{e^{2r}}{4} m(\phi_1 - \phi_2)^2 - \frac{e^{-2r}}{4} m(\phi_1 + \phi_2)^2\right].$$

For $r = 0$, this simply becomes the product state

$$\psi_0[\phi_1, \phi_2] = c \exp\left[-\frac{1}{2} m(\phi_1^2 + \phi_2^2)\right] = c \exp\left[-\frac{1}{2} m\phi_1^2\right] \exp\left[-\frac{1}{2} m\phi_2^2\right].$$

For $r > 0$, such a product decomposition is not possible, however. Generalizations of such two-mode squeezed states describe entangled states from inflation in the early universe or the entanglement of Hawking radiation emerging from a black hole with radiation falling into the horizon (for free bosonic theories). The density matrix for the two-mode system in the squeezed state is

$$\rho_{12}[\phi_{1+}, \phi_{2+}; \phi_{1-}, \phi_{2-}] = \frac{1}{Z} \exp\Bigg[- \frac{e^{2r}}{4} m(\phi_{1+} - \phi_{2+})^2 - \frac{e^{-2r}}{4} m(\phi_{1+} + \phi_{2+})^2 \\ - \frac{e^{2r}}{4} m(\phi_{1-} - \phi_{2-})^2 - \frac{e^{-2r}}{4} m(\phi_{1-} + \phi_{2-})^2 \Bigg].$$

Blackboard video

Reduced density matrix

It is instructive to calculate the reduced density matrix for the mode ϕ_1 by taking the partial trace of the density matrix. Quite generally, the reduced density matrix for a subsystem A of a larger system consisting of the parts A and B is given as the partial trace

$$\rho_A = \mathrm{Tr}_B \{\rho_{AB}\}.$$

If A and B are entangled and ρ_{AB} describes a pure state, the reduced density matrix is of a mixed state form.In contrast, for a pure product state $\rho_{AB} = \rho_A \otimes \rho_B$, the reduced density matrix ρ_A is

also pure. In the present case, taking the partial trace for the second mode corresponds to

$$\begin{aligned}
\rho_1[\phi_{1+},\phi_{1-}] =& \int D\phi\ \rho_{12}[\phi_{1+},\phi;\phi_{1-},\phi] \\
=& \frac{1}{Z}\int D\phi\ \exp\Big[-\frac{e^{2r}+e^{-2r}}{4}m(\phi_{1+}^2+\phi_{1-}^2) \\
&+2m\phi\left(\frac{e^{2r}-e^{-2r}}{4}m(\phi_{1+}+\phi_{1-})\right)-m\phi^2\frac{e^{2r}+e^{-2r}}{2}\Big] \\
=& \frac{1}{Z}\exp\Big[-\frac{1}{2}m\cosh(2r)(\phi_{1+}^2+\phi_{1-}^2)+\frac{1}{4}m\cosh(2r)\tanh^2(2r)(\phi_{1+}+\phi_{1-})^2\Big] \\
&\times\int D\phi\exp\Big[-m\cosh(2r)\left(\phi-\frac{1}{2}\tanh(2r)(\phi_{1+}+\phi_{1-})\right)^2\Big] \\
=& \frac{1}{Z}\exp\Big[-\frac{1}{2}m\cosh(2r)\left(1-\frac{1}{2}\tanh^2(2r)\right)(\phi_{1+}^2+\phi_{1-}^2) \\
&+\frac{1}{2}m\cosh(2r)\tanh^2(2r)\phi_{1+}\phi_{1-}\Big].
\end{aligned} \tag{5.4}$$

In the last step, we have performed the Gaussian integral over ϕ and dropped an irrelevant factor.

As expected, for $r>0$, the density matrix ρ_1 now is not of pure state form anymore. It does not factor into a ket and a bra because of the term $\propto\phi_{1+}\phi_{1-}$ in the exponent.
Note the resemblance of (5.4) and (5.3). This is an indication that entanglement can sometimes lead to a locally thermal looking state, albeit it is globally pure.

Blackboard video

LECTURE 11

LECTURE 12

6 Dynamics

So far we have been concerned with the description of states in the field theory, which can be specified for example at a globally fixed constant time $t=t_0$. Our next goal is to understand dynamics in time. Before going there it is worth to clarify an issue related to relativistic causality.

States on Cauchy surfaces

It seems a bit strange that time is singled out for the description of states, because, according to special relativity, observers that move with a velocity relative to each other have different notions of what equal time actually means. Indeed, states can be specified somewhat more generally on any Cauchy surface Σ. This is a $(d-1)$ dimensional submanifold of d-dimensional spacetime, a so-called hypersurface, with a normal vector that points in a time-like direction everywhere. A hypersurface $t=t_0$ with normal vector $n^\mu=(1,0,0,0)$ is then just a special case.

In the more general case, the density matrix on the hypersurface Σ is specified as a double functional of fields $\phi_+(x)$ and $\phi_-(x)$ where the coordinates are now on the hypersurface, that is $x \in \Sigma$, $\rho = \rho_\Sigma[\phi_+, \phi_-]$. In this formulation, a generalization of time evolution would correspond to an evolution between neighbouring Cauchy surfaces, e. g. $\Sigma_1 \to \Sigma_2 \to \cdots \to \Sigma_N$. Keeping this generalization in mind, we can still take evolution according to some globally defined time coordinate as a convenient special case in the following.

Blackboard video

Unitary time evolution

Similar as in quantum mechanics, the evolution in time, or between Cauchy surfaces, is realized by unitary evolution operators. For N-body quantum mechanics, this would be an operator of the type

$$U_{t_2 \leftarrow t_1}(\mathbf{x}_1, \cdots, \mathbf{x}_N; \mathbf{y}_1, \cdots, \mathbf{y}_N), \tag{6.1}$$

such that

$$\psi_{t_2}(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \int_{\mathbf{y}_1, \ldots, \mathbf{y}_N} U_{t_2 \leftarrow t_1}(\mathbf{x}_1, \cdots, \mathbf{x}_N; \mathbf{y}_1, \cdots, \mathbf{y}_N) \psi_{t_1}(\mathbf{y}_1, \cdots, \mathbf{y}_N). \tag{6.2}$$

The density matrix also needs the hermitian conjugate operator

$$U^\dagger_{t_1 \to t_2}(\mathbf{x}_1, \cdots, \mathbf{x}_N; \mathbf{y}_1, \cdots, \mathbf{y}_N) \tag{6.3}$$

so that the density matrix evolves as

$$\begin{aligned} \rho_{t_2}(\mathbf{x}_1, \cdots, \mathbf{x}_N; \mathbf{y}_1, \cdots, \mathbf{y}_N) = \int_{\mathbf{u}_1, \cdots, \mathbf{u}_N} \int_{\mathbf{v}_1, \cdots, \mathbf{v}_N} U_{t_2 \leftarrow t_1}(\mathbf{x}_1, \cdots, \mathbf{x}_N; \mathbf{u}_1, \cdots, \mathbf{u}_N) \cdot \\ \cdot \rho_{t_1}(\mathbf{u}_1, \cdots, \mathbf{u}_N; \mathbf{v}_1, \cdots, \mathbf{v}_N) U^\dagger_{t_1 \to t_2}(\mathbf{v}_1, \cdots, \mathbf{v}_N; \mathbf{y}_1, \cdots, \mathbf{y}_N) \end{aligned} \tag{6.4}$$

In a quantum field theory, one can specify in a similar way the unitary operator for evolution from one hypersurface to the next, e.g. $\Sigma_1 \to \Sigma_2$, or $t_1 \to t_2$

$$U_{t_2 \leftarrow t_1}[\phi_2, \phi_1], \tag{6.5}$$

such that the density matrix functional evolves as

$$\rho_{t_2}[\phi_{2+}, \phi_{2-}] = \int D\phi_{1+} \int D\phi_{1-} U_{t_2 \leftarrow t_1}[\phi_{2+}, \phi_{1+}] \rho_{t_1}[\phi_{1+}, \phi_{1-}] U^\dagger_{t_1 \to t_2}[\phi_{1-}, \phi_{2-}]. \tag{6.6}$$

This evolution equation of the density matrix functional is a special case of the general evolution equation for the density matrix in quantum mechanics

$$\rho_{t_2} = e^{-iH(t_2 - t_1)} \rho_{t_1} e^{iH(t_2 - t_1)}.$$

The left operator evolves the "ket" forward in time, while the right operator evolves the "bra" forward.

Blackboard video

Schwinger-Keldysh double time path

The evolution operator for the "bra" $e^{iH(t_2-t_1)}$ or

$$U^\dagger_{t_1\to t_2}[\phi_{1-},\phi_{2-}]$$

can also be understood as an operator that evolves backward in time. This is the idea beyond the Schwinger-Keldysh double time path that can be used to describe the time evolution for quantum field theories in general out-of-equilibrium situations. This is needed for example in cosmology or to describe non-equilibrium situations in condensed matter contexts. Note in particular that ρ_{t_1} can in principle be any density matrix. Because both the "ket" and the "bra" part of the density matrix are specified initially or as incoming one speaks of an in-in formalism. The outgoing state is not specified and must be calculated.

Blackboard video

Functional integral for time evolution

Let us now consider the time evolution operator

$$U_{t_\mathrm{f}\leftarrow t_0}[\phi_\mathrm{f},\phi_0].$$

We are free to insert here intermediate steps here and to write

$$U_{t_\mathrm{f}\leftarrow t_0}[\phi_\mathrm{f},\phi_0]=\int D\phi_N\cdots\int D\phi_2\int D\phi_1\,U_{t_\mathrm{f}\leftarrow t_N}[\phi_\mathrm{f},\phi_N]\cdots U_{t_2\leftarrow t_1}[\phi_2,\phi_1]U_{t_1\leftarrow t_0}[\phi_1,\phi_0].$$

By inserting many of these intermediate steps we can reduce everything to evolution operators over infinitesimal time steps $t_{j+1}=t_j+\varepsilon$. For these one can write

$$U_\varepsilon=e^{-i\varepsilon H}\approx\mathbb{1}-i\varepsilon H,$$

with terms of quadratic order and higher vanishing in the limit $\varepsilon\to 0$. The part $\mathbb{1}$ corresponds here to a functional Dirac delta, which we can write as

$$\delta[\phi_{j+1}-\phi_j]=\int D\pi_j\exp\left[i\int d^3x\,\pi_j(\mathbf{x})\left[\phi_{j+1}(\mathbf{x})-\phi_j(\mathbf{x})\right]\right]. \tag{6.7}$$

This is just the generalization of the familiar expression

$$\delta(x-y)=\int\frac{dp}{2\pi}e^{ip(x-y)},$$

to the functional formalism. The Hamiltonian is itself an operator that involves the fields ϕ and its spatial derivatives, and the conjugate momentum operators

$$-i\frac{\delta}{\delta\phi(\mathbf{x})}.$$

When acting on an expression as in (6.7) this functional derivative operator gives just $\pi_j(\mathbf{x})$ under the integral.

Collecting terms we find

$$U_{t_\mathrm{f}\leftarrow t_0}[\phi_\mathrm{f},\phi_0]=\int D\pi_N\int D\phi_N\cdots\int D\pi_1\int D\phi_1\int D\pi_0$$
$$\exp\left[i\sum_{j=0}^{N}\varepsilon\left\{\int d^3x\left\{\pi_j(\mathbf{x})\frac{\phi_{j+1}(\mathbf{x})-\phi_j(\mathbf{x})}{\varepsilon}\right\}-H[\phi_j,\pi_j]\right\}\right].$$

We have re-exponentiated here the term invoving the Hamiltonian and were a bit sloppy with the question how to order terms in the Hamiltonian. This is justified by the limit $\varepsilon \to 0$ we want to take next. We can replace with $t = t_0 + j\varepsilon$ the fields, $\phi_j(\mathbf{x}) \to \phi(t, \mathbf{x})$, conjugate momentum fields, $\pi_j(\mathbf{x}) \to \pi(t, \mathbf{x})$ and

$$\frac{\phi_{j+1}(\mathbf{x}) - \phi_j(\mathbf{x})}{\varepsilon} \to \frac{\partial}{\partial t}\phi(t, \mathbf{x}) = \dot{\phi}(t, \mathbf{x}).$$

Moreover, the sum over j in the exponent becomes an integral along time.

Blackboard video

Phase space functional integral

We find thus the functional integral expression

$$U_{t_\mathrm{f} \leftarrow t_0}[\phi_\mathrm{f}, \phi_0] = \int D\pi \int D\phi \exp\left[i \int_{t_0}^{t_\mathrm{f}} dt \int d^3x \left\{ \pi(x)\dot{\phi}(x) - \mathscr{H} \right\} \right],$$

where $\mathscr{H}$ is the Hamiltonian density. The functional integral includes now one integral over the field at each point in time and space between the initial and final time or Cauchy hypersurface. At the boundaries of the time interval (or on the bounding Cauchy surfaces) one must keep

$$\phi(t_0, \mathbf{x}) = \phi_0(\mathbf{x}), \qquad \phi(t_\mathrm{f}, \mathbf{x}) = \phi_\mathrm{f}(\mathbf{x}), \tag{6.8}$$

fixed. The integrals over the conjugate momentum fields are not constrained.

Blackboard video

Langrangian functional integral

Finally, for theories where the Hamiltonian is quadratic in the conjugate momentum fields $\pi(t, \mathbf{x})$ one can easily perform the functional integral over $\pi(t, \mathbf{x})$. Besides an irrelevant overall constant, this implies to extremize with respect to $\pi(x)$, which is effectively the Legendre transform to an integral over the Lagrangian in the exponential,

$$U_{t_\mathrm{f} \leftarrow t_0}[\phi_\mathrm{f}, \phi_0] = \int D\phi \exp\left[i \int_{t_0}^{t_f} dt \int d^3x \mathscr{L}(\phi, \partial_\mu \phi) \right], \tag{6.9}$$

where $\mathscr{L}(\phi, \partial_\mu \phi)$ is the Lagrangian density. Specifically, for the scalar field theory one has

$$\mathscr{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\nabla\phi^2 - V(\phi) = -\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi).$$

We use here the Minkowski metric with mainly plus convention, $g_{\mu\nu} = \mathrm{diag}(-, +, +, +)$, and the microscopic potential

$$V(\phi) = \frac{1}{2} m^2 \phi^2 + \frac{\lambda}{4!}\phi^4.$$

At the boundaries in time we still need to keep the fields fixed according to (6.8).

Blackboard video

LECTURE 13

Thermal density operators

Now that we know how to write evolution operators as functional integrals we can do many things with them. Let us discuss the density matrix of quantum fields in a thermal state. This is not to be confused with the classical fields in a thermal state we discussed previously. Thermal states are interesting by themself and the standard vacuum state is included in the limit $T \to 0$.

At fixed time t, the thermal density matrix is formally given by

$$\rho = \frac{1}{Z(\beta)} e^{-\beta H}, \qquad Z(\beta) = \mathrm{Tr}\left\{e^{-\beta H}\right\}, \tag{6.10}$$

where $\beta = 1/T$. By comparing this to the evolutionoperator e^{-itH} we see that the operator $e^{-\beta H}$ is in fact just an evolution operator into an imaginary time direction, for example from $t = t_0$ to $t = t_0 - i\beta$. Moreover, in the thermal partition function $Z(\beta)$ one needs to take the trace which means to identify the fields at $t = t_0$ and $t = t_0 - i\beta$. This leads to a torus geometry with peridicity in imaginary time direction where the fields satify the condition

$$\phi(t_0, \mathbf{x}) = \phi(t_0 - i\beta, \mathbf{x}).$$

This is called the Matsubara torus.

Blackboard video

Analytic continuation or Wick rotation

Let us analyse what happens to the action when we evaluate it along the imaginary time contour. We introduce the imaginary time coordinate τ through

$$t = t_0 - i\tau,$$

and integrate τ from 0 to $\beta = 1/T$. Note that

$$\frac{\partial}{\partial t} = i\frac{\partial}{\partial \tau},$$

and $dt = -id\tau$ The real-time action times the imaginary unit,

$$iS = i\int dt \int d^3x \left\{\frac{1}{2}\left(\frac{\partial}{\partial t}\phi\right)^2 - \frac{1}{2}\left(\boldsymbol{\nabla}\phi\right)^2 - V(\phi)\right\},$$

which is what enters the exponential in the transition operator U, becomes then, when evaluated along the Matsubara contour,

$$-S_{\mathrm{E}} = -\int_0^\beta d\tau \int d^3x \left\{\frac{1}{2}\left(\frac{\partial}{\partial \tau}\phi\right)^2 + \frac{1}{2}\left(\boldsymbol{\nabla}\phi\right)^2 + V(\phi)\right\}.$$

This is now an action in Euclidean space, where the metric is

$$ds^2 = d\tau^2 + d\mathbf{x}^2,$$

and the difference in sign between time and space coordinates has disapeared! The Euclidean action is of the same kind as the "actions" we have studied previously in the context of classical statistical field theories (which also explains why we called them "actions" even though formally these where parts of Hamiltonians divided by temperature). The difference is, however, that we now have one Euclidean dimension more! This additional dimension is periodic at non-zero temperature.

Blackboard video

Zero temperature or ground state

At this point it is interesting to consider the limit $T \to 0$ or $\beta \to \infty$. The circumference of the Matsubara torus becomes then infinite and τ is integrated from 0 to ∞, or, equivalently after a change of variables, from $-\infty$ to ∞. In that limit the theory is equivalent to what we discussed previously in the context of classical fields at finite temperature, but with one dimension more. The groud state of a quantum field theory in $d = 1 + 3$ dimensions can be represented by a statistical field theory with $d = 4$ dimensions.

Blackboard video

Matsubara frequencies

Taking the periodicity condition at non-zero temperature into account we can write

$$\phi(\tau, \mathbf{x}) = T \sum_{n=-\infty}^{\infty} \int \frac{d^3p}{(2\pi)^3} e^{i\omega_n \tau + i\mathbf{p}\mathbf{x}} \phi(-i\omega_n, \mathbf{p}),$$

where

$$\omega_n = 2\pi T n$$

is known as the Matsubara frequeny. While τ can be seen as an imaginary periodic time, ω_n can be seen as an imaginary discrete frequency. In the high temperature limit only the lowest Matsubara modes with $\omega_n = 0$ contribute effectively to thermodynamic observables, all others are strongly surpressed in the correlation function $1/[(2\pi T n)^2 + \mathbf{p}^2]$. Restricting to $n = 0$ leads to the classical limit of the theory we have discussed previously. In the opposite limit $T \to 0$ we obtain an integral,

$$T \sum_n \to \int \frac{d\omega}{2\pi},$$

over continuous Matsubara frequencies.

Blackboard video

Density matrix functional

The density matrix functional is given by

$$\rho[\phi_+, \phi_-] = \frac{1}{Z(\beta)} \int_{\phi_+, \phi_-} D\phi \; e^{-S_{\mathrm{E}}[\phi]}, \tag{6.11}$$

where the boundary conditions for the functional integral are

$$\phi(\tau = 0, \mathbf{x}) = \phi_+(\mathbf{x}), \qquad \phi(\tau = \beta, \mathbf{x}) = \phi_-(\mathbf{x}).$$

One easily confirms that the density matrix is normalized correctly,

$$\mathrm{Tr}\{\rho\} = \int D\phi_+ \; \rho[\phi_+, \phi_+] = \frac{1}{Z(\beta)} \int D\phi_+ \int_{\phi_+, \phi_+} D\phi \; e^{-S[\phi]} = \frac{1}{Z(\beta)} \int D\phi \; e^{-S[\phi]} = 1, \tag{6.12}$$

where $Z(\beta)$ is the thermodynamic partition function. The density matrix (6.11) can be combined with evolution operators as in eq. (6.9) and a similar representation for $U^\dagger$ to determine the density matrix at a later time. When the trace of such a density matrix is taken one obtains a *closed time path*, as a special case of the Schwinger-Keldysh double time path.

Blackboard video

In-out formalism

For many problems in quantum field theory one does not need the Schwinger-Keldysh or in-in formalism. Instead one can work in a situation where the ingoing as well as the out-going state are actually vacuum or ground states. This describes in particular situations with just a few particles in the initial and final state for which one can specify convenient creation and annihilation operators acting on the vaccum state. Much of the scattering physics needed to describe collider experiments can be described this way.

Blackboard video

Vacuum-to-vacuum transition amplitude

A contour where the incoming state and the outgoing state are the vacuum, but that nevertheless goes along real times can be achieved by rotating the real time slightly into the complex plane. By integratig fom $t_0 \to -\infty(1-i\epsilon)$ to the final time $t_\mathrm{f} \to \infty(1-i\epsilon)$ we have at both ends of the integration contour terms $e^{-i\infty(1-i\epsilon)H} \sim e^{-\epsilon\infty H}$, which effectively project to the ground state. For states above the minimal energy, the exponential supression is so strong that only the ground state remains.

The integration contour will play a role in deciding which Greens functions of a differential operator to take. Recall that Greens functions are non unique but depend on the boundary conditions. A simple prescription, equivalent to the above rotation in time integration contour, is to multipy the Hamiltonian with $(1-i\epsilon)$, or, even simpler and equivalently in practice, to replace m^2 with $m^2 - i\epsilon$. We will take this $i\epsilon$ prescription into account later on when calculating Greens functions.

Blackboard video

Feynman propagator

Let us consider now a two-point correlation function of the type

$$\frac{1}{i}G(x-y) = \frac{1}{Z}\int D\phi\, \phi(x)\phi(y)e^{iS_2[\phi]},$$

where time t is integrated along the vaccuum-to-vacuum or in-out contour and we work with the quadratic action

$$S_2[\phi] = \int d^4x \left\{-\frac{1}{2}g^{\mu\nu}\partial_\mu\phi(x)\partial_\nu\phi(x) - \frac{1}{2}m^2\phi(x)^2\right\} = -\frac{1}{2}\int \frac{d^4p}{(2\pi)^4}\left\{\phi^*(p)\left[p^2+m^2\right]\phi(p)\right\}.$$

We are using now a four-dimensional notation with $x=(t,\mathbf{x})$, $p=(\omega,\mathbf{p})$ and $p^2 = -\omega^2 + \mathbf{p}^2$. In the second equation we have introduced the fields in Fourier space through

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

where $px = -\omega t + \mathbf{px}$. For real fields one has $\phi^*(p) = \phi(-p)$. The two-point function follows now from the standard receipe of Gaussian integration and we obtain

$$\begin{aligned} G(x-y) &= \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \frac{1}{p^2+m^2-i\epsilon} \\ &= \int \frac{d^4p}{(2\pi)^4} e^{-ip^0(x^0-y^0)+i\mathbf{p}(\mathbf{x}-\mathbf{y})} \frac{-1}{(p^0-\sqrt{\mathbf{p}^2+m^2}+i\epsilon)(p^0+\sqrt{\mathbf{p}^2+m^2}-i\epsilon)}. \end{aligned} \tag{6.13}$$

Here we have inserted the $i\epsilon$ which will help us to pick the right integration contour. Note that $G(x-y)$ is a Greens function to the inverse propagator in the sense

$$\left[-g^{\mu\nu}\partial_\mu\partial_\nu + m^2\right] G(x-y) = \delta^{(4)}(x-y).$$

One may perform the integration over the frequency p^0 in eq. (6.13). Note first that there are poles at

$$p^0 = \sqrt{\mathbf{p}^2+m^2} - i\epsilon, \qquad p^0 = -\sqrt{\mathbf{p}^2+m^2} + i\epsilon.$$

For $x^0 - y^0 > 0$ we can close the p^0 integration contour in the lower half of the complex plane and get a contribution from the residue at $p^0 = \sqrt{\mathbf{p}^2+m^2} = E_\mathbf{p}$. In contrast, for $x^0 - y^0 < 0$ the contour must be closed in the upper half of the complex plane, and we pick up a contribbution from the residue at $p^0 = -\sqrt{\mathbf{p}^2+m^2} = -E_\mathbf{p}$. Taken together this yields

$$\begin{aligned} G(x-y) =& \theta(x^0-y^0) \int \frac{d^3p}{(2\pi)^3} \frac{i}{2E_\mathbf{p}} e^{-iE_\mathbf{p}(x^0-y^0)+i\mathbf{p}(\mathbf{x}-\mathbf{y})} \\ &+ \theta(y^0-x^0) \int \frac{d^3p}{(2\pi)^3} \frac{i}{2E_\mathbf{p}} e^{+iE_\mathbf{p}(x^0-y^0)+i\mathbf{p}(\mathbf{x}-\mathbf{y})} \end{aligned} \tag{6.14}$$

Depending on the time ordering we find either a term with positive frequency or one with negative frequency. The Feynman propagator $G(x-y)$ is also called time-ordered propagator.

Blackboard video

Exercise

By choosing different contours of the frequency integration, derive expressions for a retared propagtor that vanishes when $x^0 - y^0 < 0$ and an advanced propagator that vanishes when $x^0 - y^0 > 0$.

LECTURE 14

7 Non-relativistic bosonic fields

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

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

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

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

as

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

Blackboard video

Two zero crossings

One observes that the so-called inverse propagator has two zero-crossings, one at $p^0 = \sqrt{\mathbf{p}^2 + m^2}$ and one at $p^0 = -\sqrt{\mathbf{p}^2 + m^2}$. At these 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, $p^0 = \sqrt{\mathbf{p}^2 + m^2}$ gives the one for particles, $p^0 = -\sqrt{\mathbf{p}^2 + m^2}$ 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, $p^0 = \sqrt{\mathbf{p}^2 + m^2}$ which is for large m^2 rather far from the frequency of anti-particles. One can therefore replace in a first step

$$p^0 + \sqrt{\mathbf{p}^2 + m^2} \to 2\sqrt{\mathbf{p}^2 + m^2} \approx 2m.$$

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

$$p^0 = \sqrt{\mathbf{p}^2 + m^2} = m + \frac{\mathbf{p}^2}{2m} + \dots$$

This leads us to a quadratic action of the form

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

or for the full action in position space

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

Blackboard video

Rescaled fields and dispersion relation

It is now convenient to introduce rescaled fields by setting

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

The action becomes then

$$S = \int dt d^3x \left\{ \varphi^* \left(i\partial_t - V_0 + \frac{\boldsymbol{\nabla}^2}{2m} \right) \varphi - \frac{\lambda}{8m^2} (\varphi^*\varphi)^2 \right\}. \tag{7.1}$$

The dispersion relation is now with

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

given by

$$\omega = V_0 + \frac{\mathbf{p}^2}{2m}.$$

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 (7.1) 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.

Blackboard video

Symmetries of non-relativistic theory

The non-relativistic action in equation (7.1) 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,

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

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

Blackboard video

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

$$\varphi(x) \to e^{i(\alpha+\beta t)}\varphi(x), \qquad \varphi^*(x) \to e^{-i(\alpha+\beta t)}\varphi^*(x).$$

All terms in the action are invariant except for

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

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

$$\varphi^*(i\partial_t - V_0)\varphi \to \varphi^*(i\partial_t - \beta - V_0 + \beta)\varphi = \varphi^*(i\partial_t - V_0)\varphi.$$

This shows that

$$\varphi(x) \to e^{i(\alpha+\beta t)}\varphi, \qquad \varphi^* \to e^{-i(\alpha+\beta t)}\varphi^*, \qquad V_0 \to V_0 - \beta,$$

is in fact another symmetry of the action in equation (7.1). 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.

Blackboard video

Galilei transformation

Note that the action in equation (7.1) 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,

$$t \to t,$$

$$\mathbf{x} \to \mathbf{x} + \mathbf{v}t.$$

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 (7.1)? This is a little bit complicated and we directly give the transformation law,

$$\varphi(t,\mathbf{x}) \to \varphi'(t,\mathbf{x}) = e^{i\left(m\mathbf{v}\cdot\mathbf{x} - \frac{1}{2}m\mathbf{v}^2 t\right)}\varphi(t,\mathbf{x}-\mathbf{v}t).$$

Indeed one can confirm that

$$\left(i\partial_t + \frac{\boldsymbol{\nabla}^2}{2m}\right)\varphi(t,\mathbf{x}) \to e^{i\left(m\mathbf{v}\cdot\mathbf{x} - \frac{1}{2}m\mathbf{v}^2 t\right)}\left[\left(i\partial_t + \frac{\boldsymbol{\nabla}^2}{2m}\right)\varphi\right](t,\mathbf{x}-\mathbf{v}t),$$

so that the action (7.1) is invariant under Galilei transformations.

Blackboard video

Effective potential

One can write the action in (7.1) also as

$$S = \int dt d^3x \left\{\varphi^*\left(i\partial_t + \frac{\boldsymbol{\nabla}^2}{2m}\right)\varphi - V(\varphi^*\varphi)\right\}, \tag{7.2}$$

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

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

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

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

leads to

$$\phi_0 = \sqrt{\rho_0} = \sqrt{\frac{\mu}{\lambda}}.$$

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 $\mathbf{p} = 0$ 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. (7.1). 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

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

with real fields $\phi_1(x)$ and $\phi_2(x)$. The action in eq. (7.2) becomes (up to total derivatives)

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

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

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

In momentum space, the matrix between the fields becomes

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

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,

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

Here this leads to

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

or

$$\omega = \sqrt{\left(\frac{\mathbf{p}^2}{2m} + 2\lambda\phi_0^2\right)\frac{\mathbf{p}^2}{2m}}. \tag{7.3}$$

This is known as Bogoliubov dispersion relation.

Blackboard video

Linear and quadratic regimes

For small momenta, such that

$$\frac{\mathbf{p}^2}{2m} \ll 2\lambda\phi_0^2,$$

one finds

$$\omega \approx \sqrt{\frac{\lambda\phi_0^2}{m}}|\mathbf{p}|. \tag{7.4}$$

In contrast, for

$$\frac{\mathbf{p}^2}{2m} \gg 2\lambda\phi_0^2,$$

one recovers the usual dispersion relation for non-relativistic particles

$$\omega \approx \frac{\mathbf{p}^2}{2m}. \tag{7.5}$$

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

Blackboard video

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 friction. To understand this consider an interacting Bose-Einstein condensate flowing past some body of through a capillary. If the energy and momentum of the fluid are $E = E_0$ and $\mathbf{p} = 0$ in the fluid rest frame, they are

$$E' = E + \mathbf{p}\mathbf{v} + \frac{1}{2}M\mathbf{v}^2 = E_0 + \frac{1}{2}M\mathbf{v}^2, \qquad \mathbf{p}' = \mathbf{p} + M\mathbf{v} = M\mathbf{v},$$

in the rest frame of the body or capillary. We used here first the general transformation of energy E and momentum $\mathbf{p}$ under Galilei boost transformations and then the particular values for the homogeneous fluid state.

Imagine now that we can create an excitation or quasi-particle in the fluid with energy $\epsilon(\mathbf{p})$ and momentum $\mathbf{p}$. In the fluid rest frame we have now $E = E_0 + \epsilon(\mathbf{p})$ and $\mathbf{p} = \mathbf{p}$. The energy and momentum in the rest frame of the capillary are then

$$E' = E_0 + \epsilon(\mathbf{p}) + \mathbf{p}\cdot\mathbf{v} + \frac{1}{2}M\mathbf{v}^2, \qquad \mathbf{p}' = \mathbf{p} + M\mathbf{v}.$$

Comparison to the corresponding relation for the homogeneous state shows that the energy and momentum associated to the excitation are in the rest frame of the capiliary $\epsilon(\mathbf{p}) + \mathbf{p}\cdot\mathbf{v}$ and $\mathbf{p}$, respectively.

Blackboard video

Landau's criterion for superfluidity

Now the point is that at small temperature, excitations will only be created in the fluid in appreciable numbers when it is energetically favorable, i.e. for

$$\epsilon(\mathbf{p}) + \mathbf{p} \cdot \mathbf{v} < 0,$$

such that the energy of the fluid is lowered. If this relation is not fulfilled for any momentum $\mathbf{p}$, no excitations that could transport momentum out of a local fluid cell will be created. This means that there is no viscosity and the flow is superfluid. It follows that for friction to become possible, the fluid needs to have a fluid velocity larger than

$$v_c = \min_{\mathbf{p}} \frac{\epsilon(\mathbf{p})}{|\mathbf{p}|},$$

known as critical velocity. For the Bogoliubov dispersion relation (7.3) the critical velocity equals the velocity of sound.

Blackboard video

LECTURE 15

8 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 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 $\pi^+ \to \mu^+ + \nu_\mu$. 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.

Blackboard video

Mode function expansion

Let us write the non-relativistic bosonic fields as

$$\varphi(t,\mathbf{x}) = \int_{\mathbf{p}} v_{\mathbf{p}}(t,\mathbf{x})\, a_{\mathbf{p}}(t), \qquad \varphi^*(t,\mathbf{x}) = \int_{\mathbf{p}} v^*_{\mathbf{p}}(t,\mathbf{x})\, a^\dagger_{\mathbf{p}}(t),$$

with $\int_{\mathbf{p}} = \int \frac{d^3p}{(2\pi)^3}$ and the *mode functions*

$$v_{\mathbf{p}}(t,\mathbf{x}) = e^{-i\omega_{\mathbf{p}}t + i\mathbf{p}\mathbf{x}}.$$

While we plan to work in the in-out functional integral formalism, let us note that in an operator picture $a_{\mathbf{p}}(t)$ and $a^\dagger_{\mathbf{p}}(t)$ would be annihilation and creation operators for particles with momentum $\mathbf{p}$ and frequency

$$\omega_{\mathbf{p}} = \frac{\mathbf{p}^2}{2m} + V_0.$$

Note that in contrast to the relativistic case, the expansion of the non-relativistic field $\varphi(t,\mathbf{x})$ contains no creation operator and the one of $\varphi^*(t,\mathbf{x})$ no annihilation operator. This is a consequence of the absence of anti-particles in the non-relativistic theory.

Blackboard video

Scalar product

For the following discussion, it is useful to introduce a scalar product between two functions of space and time $f(t,\mathbf{x})$ and $g(t,\mathbf{x})$,

$$(f,g)_t = \int d^3x \left\{ f^*(t,\mathbf{x}) g(t,\mathbf{x}) \right\}.$$

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.

Blackboard video

Normalization of mode functions

The mode functions are normalized with respect to this scalar product as

$$(v_{\mathbf{p}}, v_{\mathbf{q}})_t = (2\pi)^3 \delta^{(3)}(\mathbf{p}-\mathbf{q}).$$

One can write

$$\begin{aligned} a_{\mathbf{p}}(t) =& (v_{\mathbf{p}}, \varphi)_t = \int d^3x e^{i\omega_{\mathbf{p}}t - i\mathbf{p}\mathbf{x}} \varphi(t,\mathbf{x}), \\ a^\dagger_{\mathbf{p}}(t) =& (v^*_{\mathbf{p}}, \varphi^*)_t = \int d^3x e^{-i\omega_{\mathbf{p}}t + i\mathbf{p}\mathbf{x}} \varphi^*(t,\mathbf{x}). \end{aligned}$$

Blackboard video

Time dependence of creation annihilation and creation operators

The right hand side depends on time t and it is instructive to take the time derivative,

$$\begin{aligned}\partial_t a_{\mathbf{p}}(t) &= \int d^3x\; e^{i\omega_{\mathbf{p}}t - i\mathbf{px}}[\partial_t + i\omega_{\mathbf{p}}]\varphi(t,\mathbf{x}) \\ &= \int d^3x\; e^{i\omega_{\mathbf{p}}t - i\mathbf{px}}\left[\partial_t + i\left(\frac{\mathbf{p}^2}{2m} + V_0\right)\right]\varphi(t,\mathbf{x}) \\ &= \int d^3x\; e^{i\omega_{\mathbf{p}}t - i\mathbf{px}}\left[\partial_t + i\left(-\frac{\boldsymbol{\nabla}^2}{2m} + V_0\right)\right]\varphi(t,\mathbf{x}).\end{aligned}$$

We used here first the dispersion relation and expressed then $\mathbf{p}^2$ as a derivative acting on the mode function (it acts here to the left). In a final step one can use partial integration to make the derivative operator act to the right,

$$\partial_t a_{\mathbf{p}}(t) = i\int d^3x\; e^{i\omega_{\mathbf{p}}t - i\mathbf{px}}\left[-i\partial_t - \frac{\boldsymbol{\nabla}^2}{2m} + V_0\right]\varphi(t,\mathbf{x}).$$

This expression confirms that $a_{\mathbf{p}}(t)$ were time-independent if $\varphi(t,\mathbf{x})$ were a solution of the one-particle Schrödinger equation. More general, it is a time-dependent, however. In a similar way one finds (exercise)

$$\partial_t a^\dagger_{\mathbf{p}}(t) = -i\int d^3x\; e^{-i\omega_{\mathbf{p}}t + i\mathbf{px}}\left[i\partial_t - \frac{\boldsymbol{\nabla}^2}{2m} + V_0\right]\varphi^*(t,\mathbf{x}).$$

Blackboard video

Incoming states

To construct the S-matrix, we first need incoming and out-going states. Incoming states can be constructed by the creation operator

$$a^\dagger_{\mathbf{p}}(-\infty) = \lim_{t\to-\infty} a^\dagger_{\mathbf{p}}(t).$$

For example, an incoming two-particle state would be

$$|\mathbf{p}_1,\mathbf{p}_2;\mathrm{in}\rangle = a^\dagger_{\mathbf{p}_1}(-\infty)a^\dagger_{\mathbf{p}_2}(-\infty)|0\rangle.$$

Blackboard video

Bosonic exchange symmetry

We note as an aside point that these state automatically obey bosonic exchange symmetry

$$|\mathbf{p}_1,\mathbf{p}_2;\mathrm{in}\rangle = |\mathbf{p}_2,\mathbf{p}_1;\mathrm{in}\rangle,$$

as a consequence of

$$a^\dagger_{\mathbf{p}_1}(-\infty)a^\dagger_{\mathbf{p}_2}(-\infty) = a^\dagger_{\mathbf{p}_2}(-\infty)a^\dagger_{\mathbf{p}_1}(-\infty).$$

Fock space

We note also general states of few particles can be constructed as

$$|\psi;\text{in}\rangle = C_0|0\rangle + \int_{\mathbf{p}} C_1(\mathbf{p})\ |\mathbf{p};\text{in}\rangle + \int_{\mathbf{p}_1,\mathbf{p}_2} C_2(\mathbf{p}_1,\mathbf{p}_2)|\mathbf{p}_1,\mathbf{p}_2;\text{in}\rangle + \dots$$

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

$$\sum_\alpha |\alpha;\text{in}\rangle\langle\alpha;\text{in}| = \mathbb{1},$$

and normalized such that $\langle\alpha;\text{in}|\beta;\text{in}\rangle = \delta_{\alpha\beta}$.

Blackboard video

Outgoing states

In a similar way to incoming states one can construct outgoing states with the operators

$$a^\dagger_{\mathbf{p}}(\infty) = \lim_{t\to\infty} a^\dagger_{\mathbf{p}}(t).$$

For example

$$|\mathbf{p}_1,\mathbf{p}_2;\text{out}\rangle = a^\dagger_{\mathbf{p}_1}(\infty)a^\dagger_{\mathbf{p}_2}(\infty)|0\rangle.$$

We consider usually transition amplitudes where outgoing states appear as a "bra", i. e. in the form

$$\langle\mathbf{p}_1,\mathbf{p}_2;\text{out}| = \langle 0|a_{\mathbf{p}_1}(\infty)a_{\mathbf{p}_2}(\infty).$$

One can read this in the sense that existing particles get annihlilated at asymptotically large times before the state becomes the vacuum again.

Blackboard video

8.1 The S-matrix

S-matrix

The S-matrix denotes now simply the transition amplitude between incoming and out-going general states $|\alpha;\text{in}\rangle$ and $|\beta;\text{out}\rangle$,

$$S_{\beta\alpha} = \langle\beta;\text{out}|\alpha;\text{in}\rangle.$$

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.

Blackboard video

Unitarity of the S-matrix

Let us first prove that the scattering matrix is unitary,

$$\begin{aligned}(S^\dagger S)_{\alpha\beta} &= \sum_\gamma (S^\dagger)_{\alpha\gamma} S_{\gamma\beta} \\ &= \sum_j \langle \gamma; \text{out}|\alpha; \text{in}\rangle^* \langle \gamma; \text{out}|\beta; \text{in}\rangle \\ &= \sum_j \langle \alpha; \text{in}|\gamma; \text{out}\rangle \langle \gamma; \text{out}|\beta; \text{in}\rangle \\ &= \langle \alpha; \text{in}|\beta; \text{in}\rangle \\ &= \delta_{\alpha\beta}.\end{aligned}$$

We have used here the completeness of the out states

$$\sum_j |\gamma; \text{out}\rangle\langle\gamma; \text{out}| = \mathbb{1}.$$

Blackboard video

Decmposition of S-matrix

It is useful to decompose the S-matrix as

$$S_{\alpha\beta} = \delta_{\alpha\beta} + \text{contributions from interactions.}$$

The first part $\delta_{\beta\alpha}$ is just the transition amplitude for the case that no scatering has occurred, i. e. the outgoing state is the same as the incoming state. For example, the S-matrix element for $2 \to 2$ scattering $\langle \mathbf{q}_1, \mathbf{q}_2; \text{out}|\mathbf{p}_1, \mathbf{p}_2; \text{in}\rangle$ has a contribution

$$(2\pi)^6 \left[\delta^{(3)}(\mathbf{p}_1 - \mathbf{q}_1)\, \delta^{(3)}(\mathbf{p}_2 - \mathbf{q}_2) + \delta^{(3)}(\mathbf{p}_1 - \mathbf{q}_2)\, \delta^{(3)}(\mathbf{p}_2 - \mathbf{q}_1)\right].$$

This is amplitude that momenta did not change, symmetrized in a way that respects bosonic exchange symmetry. The contribution from interactions (actual scattering) is more interesting and we concentrate on it in the following.

Blackboard video

Conservation laws, elaastic and inelastic collisions

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. $2 \to 2$, and inelastic collisions, such as $2 \to 4$. In a non-relativistic theory, such inelastic processes can occur for bound states, for example two H_2 - molecules can scatter into their constituents

$$H_2 + H_2 \to 4H.$$

Blackboard video

Connection between outgoing and incoming states

What is the connection between incoming and outgoing states? Let us write

$$\begin{aligned} a_{\mathbf{p}}(\infty) - a_{\mathbf{p}}(-\infty) &= \int_{-\infty}^{\infty} \partial_t a_{\mathbf{p}}(t) \\ &= i \int_{-\infty}^{\infty} dt \int d^3x \, e^{i\omega_{\mathbf{p}} t - i\mathbf{px}} \left[-i\partial_t - \frac{\boldsymbol{\nabla}^2}{2m} + V_0 \right] \varphi(t, \mathbf{x}). \end{aligned}$$

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 ($px = -p^0 x^0 + \mathbf{px} = -p^0 t + \mathbf{px}$),

$$\varphi(t, \mathbf{x}) = \int \frac{dp^0}{2\omega} \frac{d^3\mathbf{p}}{(2\pi)^3} e^{ipx} \varphi(p),$$

this would read

$$a_{\mathbf{p}}(\infty) - a_{\mathbf{p}}(-\infty) = i \left[-p^0 + \frac{\mathbf{p}^2}{2m} + V_0 \right] \varphi(p).$$

In a similar way one finds

$$\begin{aligned} a^\dagger_{\mathbf{p}}(\infty) - a^\dagger_{\mathbf{p}}(-\infty) &= -i \int_{-\infty}^{\infty} dt \int d^3x \, e^{-i\omega_{\mathbf{p}} t + i\mathbf{px}} \left[-i\partial_t - \frac{\boldsymbol{\nabla}^2}{2m} + V_0 \right] \varphi^*(t, \mathbf{x}) \\ &= -i \left[-p^0 + \frac{\mathbf{p}^2}{2m} + V_0 \right] \varphi^*(p). \end{aligned}$$

Blackboard video

Relation between S-matrix elements and correlation functions

To create particles in the initial state we can use $a^\dagger_{\mathbf{p}}(-\infty)$. In contrast, $a_{\mathbf{p}}(-\infty)$ gives a vanishing contribution when it acts on the incoming vacuum $|0\rangle$. For the final state we can similarly use $a_{\mathbf{p}}(\infty)$ to annihilate particles, while $a^\dagger_{\mathbf{p}}(\infty)$ gives a vanishing contribution when it acts on $\langle 0|$ from the right.

Blackboard video

So, effectively, one can replace

$$a_{\mathbf{p}}(\infty) \to i \left[-p^0 + \frac{\mathbf{p}^2}{2m} + V_0 \right] \varphi(p)$$

and similarly

$$a^\dagger_{\mathbf{p}}(-\infty) \to i \left[-p^0 + \frac{\mathbf{p}^2}{2m} + V_0 \right] \varphi^*(p).$$

This allows to reduce S-matrix elements to correlation functions in the in-out functional integral formalism.

Lehmann-Symanzik-Zimmermann (LSZ) reduction formula

As a concrete example, we obtain for the S-matrix element of $2 \to 2$ scattering

$$\begin{aligned}&\langle \mathbf{q}_1, \mathbf{q}_2; \text{out} | \mathbf{p}_1, \mathbf{p}_2; \text{in} \rangle \\ &= i^4 \left[-q_1^0 + \frac{\mathbf{q}_1^2}{2m} + V_0\right] \left[-q_2^0 + \frac{\mathbf{q}_2^2}{2m} + V_0\right] \left[-p_1^0 + \frac{\mathbf{p}_1^2}{2m} + V_0\right] \left[-p_2^0 + \frac{\mathbf{p}_2^2}{2m} + V_0\right] \\ &\quad \times \langle 0 | \varphi(q_1)\varphi(q_2)\varphi^*(p_1)\varphi^*(p_2) | 0 \rangle .\end{aligned}$$

This shows how S-matrix elements are connected to correlation functions. This relation is known as the *Lehmann-Symanzik-Zimmermann (LSZ) reduction formula*, here applied to non-relativistic quantum field theory.

Blackboard video

Relativistic scalar theories

Let us mention here that for a relativistic theory the LSZ formula is quite similar but one needs to replace

$$\left[-q^0 + \frac{\mathbf{q}^2}{2m} + V_0\right] \to \left[-(q^0)^2 + \mathbf{q}^2 + m^2\right],$$

and for particles $\varphi(q) \to \phi(q)$, $\varphi^*(q) \to \phi^*(q)$, while for anti-particles $\varphi(q) \to \phi^*(-q)$, $\varphi^*(q) \to \phi(-q)$.

Blackboard video

Correlation functions from functional integrals

The (formally time-ordered) correlation functions can be written as functional integrals in the in-out formalism,

$$\langle 0 | \varphi(q_1)\varphi(q_2)\varphi^*(p_1)\varphi^*(p_2) | 0 \rangle = \frac{1}{Z} \int D\varphi \; \varphi(q_1)\varphi(q_2)\varphi^*(p_1)\varphi^*(p_2) \; e^{iS[\varphi]}.$$

We can now calculate S-matrix elements from functional integrals!

Blackboard video

Partition function

Let us now consider a non-relativistic theory with the action

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

Compared to equation (7.1) we have rescaled the interaction parameter, $\frac{\lambda}{4m^2} \to \lambda$ and included the $i\epsilon$ term needed for the in-out formalism. We introduce now the partition function in the presence of source terms J as

$$Z[J] = \int D\varphi \; \exp\left[iS[\varphi] + i \int_x \{ J^*(x)\varphi(x) + J(x)\varphi^*(x) \} \right],$$

with $x = (t, \mathbf{x})$ and $\int_x = \int dt \int d^3x$.

Blackboard video

Perturbation theory for partition function

Let us write the partition function formally as

$$Z[J] = \int D\varphi \, \exp\left[-i\frac{\lambda}{2}\int_x \left(-i\frac{\delta}{\delta J(x)}\right)^2 \left(-i\frac{\delta}{\delta J^*(x)}\right)^2\right] \exp\left[iS_2[\varphi] + i\int \{J^*\varphi + \varphi^* J\}\right],$$

where the quadratic action is

$$S_2[\varphi] = \int_x \varphi^* \left(i\partial_t + \frac{\boldsymbol{\nabla}^2}{2m} - V_0 + i\epsilon\right)\varphi.$$

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

$$Z[J] = \exp\left[-i\frac{\lambda}{2}\int_x \left(-i\frac{\delta}{\delta J(x)}\right)^2 \left(-i\frac{\delta}{\delta J^*(x)}\right)^2\right] Z_2[J],$$

with the partition function for the quadratic theory

$$Z_2[J] = \int D\varphi \, e^{iS_2[\varphi] + i\int \{J^*\varphi + \varphi^* J\}}.$$

The latter is rather easy to evaluate this in momentum space, as we have seen previously. Gaussian integration yields

$$Z_2[J] = \exp\left[i\int_p J^*(p)\left(-p^0 + \tfrac{\mathbf{p}^2}{2m} + V_0 + i\epsilon\right)^{-1} J(p)\right].$$

Blackboard video

Relating functional derivatives in position and momentum space

In the following it will be useful to write also the interaction term in momentum space. One may use

$$\frac{\delta}{\delta J(x)} = \int d^4p \frac{\delta J(p)}{\delta J(x)}\frac{\delta}{\delta J(p)} = \int \frac{d^4p}{(2\pi)^4} e^{-ipx}(2\pi)^4 \frac{\delta}{\delta J(p)}$$
$$= \int \frac{d^4p}{(2\pi)^4} e^{-ipx}\delta_{J(p)} = \int_p e^{-ipx}\delta_{J(p)}.$$

Here we defined the abbreviation

$$\delta_{J(p)} = (2\pi)^4 \frac{\delta}{\delta J(p)}.$$

In a similar way

$$\frac{\delta}{\delta J^*(x)} = \int_p e^{ipx}\delta_{J^*(p)}.$$

We used also

$$\int_x e^{ipx} = (2\pi)^4 \delta^{(4)}(p).$$

Blackboard video

Perturbation series

One finds for the partition function

$$\begin{aligned} Z[J] &= \exp\left[-i\frac{\lambda}{2}\int_x \left(\tfrac{\delta}{\delta J(x)}\right)^2 \left(\tfrac{\delta}{\delta J^*(x)}\right)^2\right] Z_2[J] \\ &= \exp\left[-i\frac{\lambda}{2}\int_{k_1\ldots k_4} \left\{(2\pi)^4\delta^4(k_1+k_2-k_3-k_4)\delta_{J(k_1)}\delta_{J(k_2)}\delta_{J^*(k_3)}\delta_{J^*(k_4)}\right\}\right] \\ &\quad \times \exp\left[i\int_p J^*(p)\left(-p^0+\tfrac{\mathbf{p}^2}{2m}+V_0-i\epsilon\right)^{-1} J(p)\right]. \end{aligned} \tag{8.1}$$

One can now expand the exponential to obtain a formal perturbation series in λ, similar to what we have seen previously for statistical field theories.

Blackboard video

S-matrix element

Let us now come back to the S-matrix element for $2 \to 2$ scattering

$$\begin{aligned} &\langle \mathbf{q}_1, \mathbf{q}_2; \text{out} | \mathbf{p}_1, \mathbf{p}_2; \text{in}\rangle \\ &= i^4 \left[-q_1^0 + \tfrac{\mathbf{q}_1^2}{2m} + V_0\right]\left[-q_2^0 + \tfrac{\mathbf{q}_2^2}{2m} + V_0\right]\left[-p_1^0 + \tfrac{\mathbf{p}_1^2}{2m} + V_0\right]\left[-p_2^0 + \tfrac{\mathbf{p}_2^2}{2m} + V_0\right] \\ &\quad \times \left(\frac{1}{Z[J]}\delta_{J^*(q_1)}\delta_{J^*(q_2)}\delta_{J(p_1)}\delta_{J(p_2)} Z[J]\right)_{J=0}. \end{aligned}$$

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.

Order λ

At order λ we have different derivatives acting on $Z_2[J]$,

- $\delta_{J(p_1)}$ for incoming particles with momentum p_1
- $\delta_{J^*(q_1)}$ for outgoing particle with momentum q_1
- $\delta_{J(k)}$ and $\delta_{J^*(k)}$ for the interaction term.

Blackboard video

Propagator

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

$$\delta_{J(p_1)}\delta_{J^*(q_1)}\left[i\int_p J^*(p)\left(-p^0+\tfrac{\mathbf{p}^2}{2m}+V_0-i\epsilon\right)^{-1}J(p)\right] = iG(p)(2\pi)^4\delta^{(4)}(p_1-q_1).$$

with the non-relativistic propagator in momentum space

$$G(p) = \frac{1}{-p_1^0+\frac{\mathbf{p}_1^2}{2m}+V_0-i\epsilon}.$$

Blackboard video

Momentum conservation

If two derivatives representing external particles would hit the same integral in $Z_2[J]$, one would have no scattering because $\mathbf{p}_1 = \mathbf{q}_1$ and as a result of momentum conservation then also $\mathbf{p}_2 = \mathbf{q}_2$. This is no real scattering. Only if a derivative representing an incoming or outgoing particle is combined with a derivative from the interaction term, this is avoided.

Resulting contribution to S-matrix

By doing the algebra one finds at order λ the term for scattering

$$\langle \mathbf{q}_1, \mathbf{q}_2; \text{out}|\mathbf{p}_1, \mathbf{p}_2; \text{in}\rangle = -i\frac{\lambda}{2}4\,(2\pi)^4\delta^{(4)}(q_1+q_2-p_1-p_2).$$

The factor $4 = 2 \times 2$ comes from different ways to combine functional derivatives with sources.

Blackboard video

Momentum conservation

The overall Dirac function makes sure that the incoming four-momentum equals the out-going four-momentum,

$$p^{\text{in}} = p_1 + p_2 = q_1 + q_2 = p^{\text{out}}.$$

Transition amplitude

Quite generally, one can define for the non-trivial part of an S-matrix

$$\langle \beta; \text{out}|\alpha; \text{in}\rangle = (2\pi)^4\delta^{(4)}(p^{\text{out}} - p^{\text{in}})\, i\,\mathcal{T}_{\beta\alpha}.$$

Together with the trivial part from "no scattering", one can write

$$S_{\beta\alpha} = \delta_{\beta\alpha} + (2\pi)^4\delta^{(4)}(p^{\text{out}} - p^{\text{in}})\, i\,\mathcal{T}_{\beta\alpha}.$$

By comparison of expressions we find for the $2 \to 2$ scattering of non-relativistic bosons at lowest order in λ simply

$$\mathcal{T} = -2\lambda,$$

independant of momenta. More generally, the transition amplitude $\mathcal{T}$ is expected to depend on the momenta of incoming and outgoing particles.

Blackboard video

Diagrammatic representation

To keep the overview over a calculation it is sometimes useful to introduce a graphical representation. For the perturbation series discussed above we may represent incoming particles by and similarly outgoing particles by These functional derivatives are acting on the partition function $Z[J]$. The partition function in (8.1) can be written in a perturbative series with the interaction term represented by One can the let the functional derivatives act on the sources and at the end evaluate everything at $J = 0$. While the diagrammatic representation is useful, it is only an auxiliary tool to organize the algebra. With a bit of experience one can work well with it.

Blackboard video

Transition propability

Let us start from an S-matrix element in the form

$$\langle \beta; \text{out} | \alpha; \text{in} \rangle = (2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}})\, i\, \mathcal{T}$$

with transition amplitude $\mathcal{T}$ which may depend on the momenta itself. (For $2 \to 2$ 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

$$P = \frac{|\langle \beta; \text{out} | \alpha; \text{in} \rangle|^2}{\langle \beta; \text{out} | \beta; \text{out} \rangle \langle \alpha; \text{in} | \alpha; \text{in} \rangle}.$$

Blackboard video

Transition rate

The numerator contains a factor

$$\Big[(2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}})\Big]^2 = (2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}})(2\pi)^4 \delta^{(4)}(0).$$

This looks ill defined but becomes meaningful in a finite volume V and for finite time interval ΔT. In fact

$$(2\pi)^4 \delta^4(0) = \int d^4x\, e^{i0x} = V \Delta T.$$

For the transition rate $\dot{P} = \frac{P}{\Delta T}$ we can therefore write

$$\dot{P} = \frac{V (2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}}) |\mathcal{T}|^2}{\langle \beta; \text{out} | \beta; \text{out} \rangle \langle \alpha; \text{in} | \alpha; \text{in} \rangle}.$$

Blackboard video

Normalization of incoming and outgoing states

Moreover, for incoming and outgoing two-particle states, their normalization is obtained from

$$\begin{aligned}\langle \mathbf{p}_1, \mathbf{p}_2; \text{in} | \mathbf{p}_1, \mathbf{p}_2; \text{in} \rangle &= \lim_{\mathbf{q}_j \to \mathbf{p}_j} \langle \mathbf{p}_1, \mathbf{p}_2; \text{in} | \mathbf{q}_1, \mathbf{q}_2; \text{in} \rangle \\ &= \lim_{\mathbf{q}_j \to \mathbf{p}_j} \left[(2\pi)^6 \left(\delta^{(3)}(\mathbf{p}_1 - \mathbf{q}_1) \delta^{(3)}(\mathbf{p}_2 - \mathbf{q}_2) + \delta^{(3)}(\mathbf{p}_1 - \mathbf{q}_2) \delta^{(3)}(\mathbf{p}_2 - \mathbf{q}_1) \right) \right] \\ &= \left[(2\pi)^3 \delta^{(3)}(0) \right]^2 \\ &= V^2 .\end{aligned}$$

Blackboard video

Counting of momentum states

In a finite volume $V = L^3$, and with periodic boundary conditions, the final momenta are of the form

$$\mathbf{p} = \frac{2\pi}{L}(m, n, l),$$

with some integer numbers m, n, l. We can count final states according to

$$\sum_{m,n,l} = \sum_{m,n,l} \Delta m \Delta n \Delta l = L^3 \sum_{m,n,l} \frac{\Delta p_1 \Delta p_2 \Delta p_3}{(2\pi)^3}.$$

In the continuum limit this becomes

$$V \int \frac{d^3 p}{(2\pi)^3}.$$

The differential transition rate has one factor $V d^3 p/(2\pi)^3$ for each final state particle.

Blackboard video

Differential transition rate

For $2 \to 2$ scattering,

$$d\dot{P} = (2\pi)^4 \delta^{(4)}(p^{\text{out}} - p^{\text{in}}) |\mathcal{T}|^2 \frac{1}{V} \frac{d^3 q_1}{(2\pi)^3} \frac{d^3 q_2}{(2\pi)^3}.$$

This can be integrated to give the transition rate into a certain region of momentum states.

Blackboard video

Flux of incoming particles

We can go from the transition probability to a cross-section by dividing through the flux of incoming particles

$$\mathcal{F} = \frac{1}{V} v = \frac{2|\mathbf{p}_1|}{mV}.$$

Here we have a density of one particle per volume V and the relative velocity of the two particles is $v = 2|\mathbf{p}_1|/m$, in the center-of-mass frame where $|\mathbf{p}_1| = |\mathbf{p}_2|$, for identical particles with equal mass m.

Blackboard video

Differential cross-section

This cancels the last factor V and we find for the differential cross-section

$$d\sigma = \frac{|\mathcal{T}|^2 m}{2|\mathbf{p}_1|}(2\pi)^4\delta^{(4)}(p^{\text{out}} - p^{\text{in}})\frac{d^3q_1}{(2\pi)^3}\frac{d^3q_2}{(2\pi)^3}.$$

Blackboard video

Phase space integrals

In the center-of-mass frame one has also $\mathbf{p}^{\text{in}} = \mathbf{p}_1 + \mathbf{p}_2 = 0$ and accordingly

$$\delta^{(4)}(p^{\text{out}} - p^{\text{in}}) = \delta(E^{\text{out}} - E^{\text{in}})\,\delta^{(3)}(\mathbf{q}_1 + \mathbf{q}_2).$$

The three-dimensional part can be used to perform the integral over $\mathbf{q}_2$. 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 $|\mathbf{q}_1, \mathbf{q}_2; \text{out}\rangle$ equals the state $|\mathbf{q}_2, \mathbf{q}_1; \text{out}\rangle$. 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 $\mathbf{q}_2$,

$$d\sigma = \frac{|\mathcal{T}|^2 m}{2|\mathbf{p}_1|(2\pi)^2}\delta(E^{\text{out}} - E^{\text{in}})d^3q_1.$$

Magnitude and solid angle

We can now use

$$d^3\mathbf{q}_1 = |\mathbf{q}_1|^2 d|\mathbf{q}_1|\; d\Omega_{q_1}$$

where $d\Omega_{q_1}$ is the differential solid angle element. Moreover

$$E^{\text{out}} = \frac{\mathbf{q}_1^2}{2m} + \frac{\mathbf{q}_2^2}{2m} + 2V_0 = \frac{\mathbf{q}_1^2}{m} + 2V_0,$$

and

$$\frac{dE^{\text{out}}}{d|\mathbf{q}_1|} = 2\frac{|\mathbf{q}_1|}{m}.$$

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 $|\mathbf{q}_1|$ using the Dirac function $\delta(E^{\text{out}} - E^{\text{in}})$. This yields $|\mathbf{q}_1| = |\mathbf{p}_1|$ and

$$d\sigma = \frac{|\mathcal{T}|^2 m^2}{16\pi^2}\; d\Omega_{q_1}.$$

Blackboard video

Total cross-section

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-sections is therefore

$$\sigma = \frac{|\mathcal{T}|^2 m^2}{8\pi}.$$

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

$$\sigma = \frac{\lambda^2 m^2}{2\pi}.$$

Blackboard video

Dimensions

Let us check the dimensions. The action

$$S = \int dt\, d^3x \left\{ \varphi^* \left(i\partial_t + \tfrac{\boldsymbol{\nabla}^2}{2m} - V_0 \right) \varphi - \tfrac{\lambda}{2} (\varphi^* \varphi)^2 \right\}$$

must be dimensionless. The field φ must have dimension

$$[\varphi] = \text{length}^{-\frac{3}{2}}.$$

The interaction strength λ must accordingly have dimension

$$[\lambda] = \frac{\text{length}^3}{\text{time}}.$$

Because

$$\left[\tfrac{\boldsymbol{\nabla}^2}{2m}\right] = \frac{1}{\text{time}},$$

one has $[m] = \frac{\text{time}}{\text{length}^2}$ and therefore $[\lambda m] = \text{length}$. It follows that indeed

$$[\sigma] = \text{length}^2$$

as appropriate for a cross-section.

Blackboard video

LECTURE 16

9 Non-relativistic 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.

9.1 Pauli spinors

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

$$\Psi(t,\mathbf{x}) = \begin{pmatrix} \psi_\uparrow(t,\mathbf{x}) \\ \psi_\downarrow(t,\mathbf{x}) \end{pmatrix}$$

We also use the notation $\psi_a(t,\mathbf{x})$ where $a = 1, 2$ and

$$\psi_1(t,\mathbf{x}) = \psi_\uparrow(t,\mathbf{x}), \qquad \psi_2(t,\mathbf{x}) = \psi_\downarrow(t,\mathbf{x}).$$

Pauli equation

The Pauli equation is a generalisation of Schrödinger's equation (neglecting spin-orbit coupling),

$$\left[\left(-i\partial_t - \tfrac{\boldsymbol{\nabla}^2}{2m} + V_0\right)\mathbb{1} + \mu_B\,\boldsymbol{\sigma}\cdot\mathbf{B}\right]\Psi(t,\mathbf{x}) = 0,$$

or equivalently

$$\left[\left(-i\partial_t - \tfrac{\vec{\nabla}^2}{2m} + V_0\right)\delta_{ab} + \mu_B\,\boldsymbol{\sigma}_{ab}\cdot\mathbf{B}\right]\psi_b(t,\mathbf{x}) = 0.$$

Here we use the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and $\mathbf{B} = (B_1, B_2, B_3)$ is the magnetic field, while μ_B is the magneton that quantifies the magnetic moment.

Attempt for an action

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

$$S_2 \stackrel{?}{=} \int dt d^3x \left\{-\Psi^\dagger\left[\left(-i\partial_t - \frac{\boldsymbol{\nabla}^2}{2m} + V_0\right)\mathbb{1} + \mu_B\,\boldsymbol{\sigma}\cdot\mathbf{B}\right]\Psi\right\}$$

However, we also need to take care of fermionic (anti-symmetric) exchange symmetry, such that for fermionic states

$$|\mathbf{p}_1, \mathbf{p}_2; \mathrm{in}\rangle = -|\mathbf{p}_2, \mathbf{p}_1; \mathrm{in}\rangle.$$

To this aspect we turn next.

Blackboard video

9.2 Grassmann numbers and fields

Grassmann variables

So-called *Grassmann variables* are generators θ_i of an algebra, and they are anti-commuting such that

$$\theta_i \theta_j + \theta_j \theta_i = 0.$$

An immediate consequence is that ${\theta_j}^2 = 0$.

Basis

If there is a finite set of generators $\theta_1, \theta_2, \ldots, \theta_n$, 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

$$\begin{aligned}
&1, \\
&\theta_1, \theta_2, \ldots, \theta_n, \\
&\theta_1\theta_2, \theta_1\theta_3, \ldots, \theta_2\theta_3, \theta_2\theta_4, \ldots, \theta_{n-1}\theta_n, \\
&\ldots \\
&\theta_1\theta_2\theta_3 \cdots \theta_n.
\end{aligned}$$

There are 2^n such basis elements, because each Grassmann variable θ_j can be either present or absent.

Blackboard video

Grade of monomial

To a monomial $\theta_{j_1} \cdots \theta_{j_q}$ 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

$$A_p A_q = (-1)^{p \cdot q} A_q A_p.$$

In particular, the monomials of even grade

$$\begin{aligned}
&1, \\
&\theta_1\theta_2, \theta_1\theta_3, \ldots, \theta_2\theta_3, \ldots, \theta_{n-1}\theta_n, \\
&\ldots
\end{aligned}$$

commute with other monomials, be the latter of even or odd grade.

Blackboard video

Grassmann parity

One can define a Grassmann parity transformation P that acts on all generators according to

$$P(\theta_j) = -\theta_j, \qquad P^2 = \mathbb{1}.$$

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".

Blackboard video

Functions of Grassmann variables

Because of $\theta^2 = 0$, functions of a Grassmann variable θ are always linear,

$$f(\theta) = f_0 + \theta f_1.$$

Note that f_0 and f_1 could depend on other Grassmann variables but not θ.

Blackboard video

Differentiation for Grassmann variables

To define differentiation of $f(\theta)$ with respect to θ we first bring it to the form

$$f(\theta) = f_0 + \theta f_1$$

and set then

$$\frac{\partial}{\partial\theta} f(\theta) = f_1.$$

Note that similar to $\theta^2 = 0$ one has also $\left(\frac{\partial}{\partial\theta}\right)^2 = 0$. 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

$$\frac{\partial}{\partial\theta} f(\sigma(\theta), x(\theta)) = \frac{\partial\sigma}{\partial\theta}\frac{\partial f}{\partial\sigma} + \frac{\partial x}{\partial\theta}\frac{\partial f}{\partial x}.$$

The derivative we use here is a left derivative.

Consider for example

$$f = f_0 + \theta_1\theta_2.$$

One has then

$$\frac{\partial}{\partial\theta_1} f = \theta_2, \qquad \frac{\partial}{\partial\theta_2} f = -\theta_1,$$
$$\frac{\partial}{\partial\theta_2}\frac{\partial}{\partial\theta_1} f = 1, \qquad \frac{\partial}{\partial\theta_1}\frac{\partial}{\partial\theta_2} f = -1.$$

One could also define a right derivative such that

$$f\frac{\overleftarrow{\partial}}{\partial\theta_1} = -\theta_2, \qquad f\frac{\overleftarrow{\partial}}{\partial\theta_2} = \theta_1.$$

Blackboard video

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,

$$\int_{-\infty}^{\infty} dx\ c\ f(x) = c \int_{-\infty}^{\infty} dx\ f(x).$$

The other is invariance under shifts of the integration variable,

$$\int_{-\infty}^{\infty} dx\ f(x+a) = \int_{-\infty}^{\infty} dx\ f(x).$$

For a function of a Grassmann variable

$$f(\theta) = f_0 + \theta f_1$$

One sets therefore

$$\int d\theta\ f(\theta) = f_1.$$

In other words, we have defined

$$\int d\theta = 0, \qquad \int d\theta\ \theta = 1.$$

This is indeed linear and makes sure that

$$\int d\theta\ f(\theta+\sigma) = \int d\theta\,\{(f_0 + \sigma f_1) + \theta f_1\,\} = \int d\theta\ f(\theta) = f_1.$$

Note that one has formally

$$\int d\theta\ f(\theta) = \frac{\partial}{\partial\theta} f(\theta).$$

Blackboard video

Several variables

For functions of several variables one has

$$\int d\theta_1 \int d\theta_2 f(\theta_1, \theta_2) = \frac{\partial}{\partial\theta_1}\frac{\partial}{\partial\theta_2} f(\theta_1, \theta_2).$$

It is easy to see that derivatives with respect to Grassmann variables anti-commute

$$\frac{\partial}{\partial\theta_j}\frac{\partial}{\partial\theta_k} = -\frac{\partial}{\partial\theta_k}\frac{\partial}{\partial\theta_j},$$

and accordingly also the differentials anti-commute

$$d\theta_j d\theta_k = -d\theta_k d\theta_j.$$

Blackboard video

Functions of several Grassmann variables

A function that depends on a set of Grassmann variables $\theta_1, \ldots, \theta_n$ can be written as

$$f(\theta) = f_0 + \theta_j f_1^j + \frac{1}{2}\theta_{j_1}\theta_{j_2} f_2^{j_1\ j_2} + \ldots + \frac{1}{n!}\theta_{j_1}\cdots\theta_{j_n} f_n^{j_1\cdots j_n}.$$

We use here Einsteins summation convention with indices j_k being summed over. The coefficients $f_k^{j_1\cdots j_k}$ 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

$$f_n^{j_1\cdots j_n} = \tilde{f}_n \varepsilon_{j_1\cdots j_n},$$

where $\varepsilon_{j_1\cdots j_n}$ is the completely anti-symmetric Levi-Civita symbol in n dimensions with $\varepsilon_{12\ldots n} = 1$.

Blackboard video

Differentiation and integration

Let us now discuss what happens if we differentiate or integrate $f(\theta)$. One has

$$\frac{\partial}{\partial\theta_k} f(\theta) = f_1^k + \theta_{j_2} f_2^{kj_2} + \ldots + \frac{1}{(n-1)!}\theta_{j_2}\cdots\theta_{j_n} f_n^{kj_2\cdots j_n},$$

and similar for higher order derivatives. In particular

$$\frac{\partial}{\partial\theta_n}\cdots\frac{\partial}{\partial\theta_1} f(\theta) = f_n^{12\ldots n} = \tilde{f}_n.$$

This defines also the integral with respect to all n variables,

$$\int d\theta_n \cdots d\theta_1 f(\theta) = f_n^{12\ldots} = \tilde{f}_n$$
$$= \int d^n\theta f(\theta) = \int D\theta f(\theta).$$

Blackboard video

Linear change of Grassmann variables

Let us consider a linear change of the Grassmann variables in the form (summation over k is implied)

$$\theta_j = J_{jk}\theta'_k,$$

where J_{jk} is a matrix of commuting variables. We can write

$$f(\theta) = f_0 + \ldots + \frac{1}{n!}\left(J_{i_1 j_1}\theta'_{j_1}\right)\cdots\left(J_{i_n j_n}\theta'_{jn}\right)\varepsilon_{i_1\cdots i_n}\tilde{f}_n.$$

Now one can use the identity

$$\varepsilon_{i_1\ldots i_n} J_{i_1 j_1}\cdots J_{i_n j_n} = \det(J)\,\varepsilon_{j_1\ldots j_n}.$$

This can actually be seen as the definition of the determinant. One can therefore write

$$f(\theta) = f_0 + \ldots + \frac{1}{n!}\theta'_{j_1} \cdots \theta'_{j_n} \varepsilon_{j_1 \ldots j_n} \det(J) \tilde{f}_n.$$

The integral with respect to θ' is

$$\int d^n\theta' f(\theta) = \det(J)\tilde{f}_n.$$

In summary, one has

$$\int d^n\theta f(\theta) = \frac{1}{\det(J)} \int d^n\theta' f(\theta).$$

Linear change of ordinary variables

One should compare this to the corresponding relation for conventional integrals with $x_j = J_{jk}x'_k$. In that case one has

$$\int d^n x f(x) = \det(J) \int d^n x' f(x').$$

Note that the determinant appears in the denominator for Grassmann variables while it appears in the numerator for conventional integrals.

Blackboard video

Gaussian integrals of Grassmann variables

Consider a Gaussian integral of two Grassmann variables

$$\int d\theta d\xi \, e^{-\theta\xi b} = \int d\theta d\xi \, (1 - \theta\xi b) = \int d\theta d\xi \, (1 + \xi\theta b) = b.$$

For a Gaussian integral over conventional complex variables one has instead

$$\int d(\mathrm{Re}\, x) \; d(\mathrm{Im}\, x) \, e^{-x^* x b} = \frac{\pi}{b}.$$

Again, integrals over Grassmann and ordinary variables behave in some sense "inverse".

Blackboard video

Higher dimensional Gaussian integrals

For higher dimensional Gaussian integrals over Grassmann numbers we write

$$\int d^n\theta d^n\xi e^{-\theta_j a_{jk} \xi_k} = \int d\theta_n d\xi_n \cdots d\theta_1 d\xi_1 e^{-\theta_j a_{jk} \xi_k}.$$

One can now employ two unitary matrices with unit determinat to perform a change of variables

$$\theta_j = \theta'_l U_{lj}, \qquad \xi_k = V_{km}\xi'_m,$$

such that

$$U_{lj} a_{jk} V_{km} = \tilde{a}_l \delta_{lm},$$

is diagonal. This is always possible. The Gaussian integral becomes

$$d^n\theta d^n\xi\, e^{-\theta_j a_{jk}\xi_k} = \det(U)^{-1}\det(V)^{-1}\int d^n\theta' d^n\xi' e^{-\theta'_l\xi'_l\tilde{a}_l} = \prod_{l=1}^{n}\tilde{a}_l = \det(a_{jk}).$$

Again this is in contrast to a similar integral over commuting variables where the determinant would appear in the denominator.

Blackboard video

Gaussian integrals with sources

Finally let us consider a Gaussian integral with source forms,

$$\int d^n\bar{\psi}d^n\psi\ \exp\left[-\bar{\psi}M\psi + \bar{\eta}\psi + \bar{\psi}\eta\right] = Z(\bar{\eta},\eta).$$

We integrate here over independent Grassmann variables $\psi = (\psi_1,\ldots,\psi_n)$ and $\bar{\psi} = (\bar{\psi}_1,\ldots,\bar{\psi}_n)$ and we use the abbreviation

$$\bar{\psi}M\psi = \bar{\psi}_j M_{jk}\psi_k.$$

The source forms are also Grassmann variables $\eta = (\eta_1,\ldots,\eta_n)$ and $\bar{\eta} = (\bar{\eta}_1,\ldots,\bar{\eta}_n)$ with

$$\bar{\eta}\psi = \bar{\eta}_j\psi_j, \qquad \bar{\psi}\eta = \bar{\psi}_j\eta_j.$$

As usual, we can write

$$Z(\bar{\eta},\eta) = \int d^n\bar{\psi}d^n\psi\ \exp\left[-(\bar{\psi}-\eta M^{-1})M(\psi - M^{-1}\eta) + \bar{\eta}M^{-1}\eta\right].$$

A shift of integration variables does not change the result and thus we find

$$Z(\bar{\eta},\eta) = \det(M)\exp\left[\bar{\eta}M^{-1}\eta\right].$$

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.

Blackboard video

Functional integral over Grassmann fields

We can now take the limit $n\to\infty$ and write

$$\int d^n\bar{\psi}d^n\psi \to \int D\bar{\psi}D\psi, \qquad Z(\bar{\eta},\eta)\to Z[\bar{\eta},\eta],$$

with

$$Z[\bar{\eta},\eta] = \int D\bar{\psi}D\psi\ \exp[-\bar{\psi}M\psi + \bar{\eta}\psi + \bar{\psi}\eta] = \det(M)\exp\left[\bar{\eta}M^{-1}\eta\right].$$

In this way we obtain a formalism that can be used for fermionic or Grassmann fields.

Blackboard video

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,

$$S_2 = \int dt d^3x \left\{ -\bar{\psi} \left[\left(-i\partial_t - \frac{\boldsymbol{\nabla}^2}{2m} + V_0 \right) \mathbb{1} + \mu_B \boldsymbol{\sigma} \cdot \mathbf{B} \right] \psi \right\},$$

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 $\bar{\psi}_1(x)\psi_1(x) \neq 0$.

Blackboard video

LECTURE 17

9.3 Yukawa theory

Yukawa theory

Let us now investigate a theory for a non-relativistic fermion with spin 1/2 and a real, relativistic scalar boson

$$S = \int dt d^3x \left\{ -\bar{\psi} \left(-i\partial_t - \frac{\boldsymbol{\nabla}^2}{2m} + V_0 - i\epsilon \right) \psi - \tfrac{1}{2}\phi \left(\partial_t^2 - \boldsymbol{\nabla}^2 + M^2 - i\epsilon \right) \phi - g\phi\bar{\psi}\psi \right\}.$$

Blackboard video

Partition function for Yukawa theory

We will discuss this theory in terms of the partition function

$$Z[\bar{\eta}, \eta, J] = \int D\bar{\psi} D\psi D\phi \; e^{iS[\bar{\psi},\psi,\phi] + i\int_x \{\bar{\eta}\psi + \bar{\psi}\eta + J\phi\}}.$$

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 $\sim g$.

Quadratic action

Let us first concentrate on the quadratic theory and the corresponding partition function derived from the action

$$S_2 = \int dt d^3x \left\{ -\bar{\psi} \left(-i\partial_t - \frac{\boldsymbol{\nabla}^2}{2m} + V_0 - i\epsilon \right) \psi - \frac{1}{2}\phi(\partial_t^2 - \boldsymbol{\nabla}^2 + M^2 - i\epsilon)\phi \right\}.$$

By doing the Gaussian integration one finds

$$Z_2[\bar\eta,\eta,J] = \int D\bar\psi D\psi D\phi\ \exp\left[iS_2 + i\int_x \{\bar\eta\psi + \bar\psi\eta + J\phi\}\right]$$
$$= \exp\left[i\int d^4x d^4y \left\{\bar\eta(x)\Upsilon(x-y)\eta(y) + \frac{1}{2}J(x)\Delta(x-y)J(y)\right\}\right]$$

where $\Upsilon(x-y)$ is the Greens function for fermions in eq. (**??**). For the scalar bosons, the Green function is

$$\Delta(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{1}{-(p^0)^2+\mathbf{p}^2+M^2-i\epsilon}e^{ip(x-y)},$$

as discussed previously.

Blackboard video

Propagator for non-relativistic fermions

For the non-relativistic fermion, the propagator integral over p^0 has just a single pole at $p^0 = \frac{\mathbf{p}^2}{2m} + V_0 - i\epsilon$,

$$\Upsilon(x-y) = \mathbb{1}\int \frac{dp^0}{2\pi}\frac{d^3p}{(2\pi)^3}\frac{1}{-p^0+\frac{\mathbf{p}^2}{2m}+V_0-i\epsilon}e^{-ip^0(x^0-y^0)+i\mathbf{p}(\mathbf{x}-\vec y)}$$

When $x^0-y^0>0$ the contour can be closed below the real p^0-axis, leading to

$$\Upsilon(x-y) = i\ \mathbb{1}\int \frac{d^3p}{(2\pi)^3}\ e^{-i\left(\frac{\mathbf{p}^2}{2m}+V_0\right)(x^0-y^0)+i\mathbf{p}(\mathbf{x}-\vec y)} \qquad (x^0-y^0>0).$$

In contrast, for $x^0-y^0<0$, the contour can be closed above and there is no contribution at all. In summary

$$\Upsilon(x-y) = i\,\theta(x^0-y^0)\ \mathbb{1}\int \frac{d^3p}{(2\pi)^3}\ e^{-i\left(\frac{\mathbf{p}^2}{2m}+V_0\right)(x^0-y^0)+i\mathbf{p}(\mathbf{x}-\vec y)}.$$

As a consequence of the absence of anti-particle-type excitations, the time-ordered and retarded propagators agree here.

Blackboard video

Propagator and correlation functions

Let us also note the relation between propagators and correlation functions. For the free (quadratic) theory one has in the fermionic sector

$$\langle\psi_a(x)\bar\psi_b(y)\rangle = \left(\frac{1}{Z_2}\frac{\delta}{\delta\bar\eta_a(x)}\frac{\delta}{\delta\eta_b(y)}Z_2[\bar\eta,\eta,J]\right)_{\bar\eta=\eta=J=0}$$
$$= -i\Upsilon_{ab}(x-y),$$

Note that some care is needed with interchanges of Grassmann variables to obtain this expression. Similarly for the bosonic scalar field

$$\langle\phi(x)\phi(y)\rangle = \left(\frac{1}{Z_2}\frac{\delta}{\delta J(x)}\frac{\delta}{\delta J(y)}Z_2[\bar\eta,\eta,J]\right)_{\bar\eta=\eta=J=0}$$
$$= -i\Delta(x-y).$$

Blackboard video

Wick's theorem

As discussed previously, one finds for the free theory

$$\begin{aligned}\langle\phi(x_1)\dots\phi(x_n)\rangle &= \left(\frac{1}{Z_2}\left(-i\frac{\delta}{\delta J(x_1)}\right)\cdots\left(-i\frac{\delta}{\delta J(x_n)}\right)Z_2[\bar\eta,\eta,J]\right)_{\bar\eta=\eta=J=0}\\ &= \sum_{\text{pairings}}[-i\Delta(x_{j_1}-x_{j_2})]\cdots[-i\Delta(x_{j_{n-1}}-x_{j_n})]\,.\end{aligned}$$

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}), …, $(x_{j_{n-1}},x_{j_n})$. This result is known as *Wick's theorem*. It follows directly from the combinatorics of functional derivatives acting on Z_2.

For example,

$$\begin{aligned}\langle\phi(x_1)\,\phi(x_2)\,\phi(x_3)\phi(x_4)\rangle =&[-i\Delta(x_1-x_2)][-i\Delta(x_3-x_4)]\\ &+[-i\Delta(x_1-x_3)][-i\Delta(x_2-x_4)]\\ &+[-i\Delta(x_1-x_4)][-i\Delta(x_2-x_3)].\end{aligned}$$

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

$$\begin{aligned}\left\langle\psi_{a_1}(x_1)\psi_{a_2}(x_2)\bar\psi_{a_3}(x_3)\bar\psi_{a_4}(x_4)\right\rangle =&-\left\langle\psi_{a_1}(x_1)\bar\psi_{a_3}(x_3)\right\rangle\left\langle\psi_{a_2}(x_2)\bar\psi_{a_4}(x_4)\right\rangle\\ &+\left\langle\psi_{a_1}(x_1)\bar\psi_{a_4}(x_4)\right\rangle\left\langle\psi_{a_2}(x_2)\bar\psi_{a_3}(x_3)\right\rangle\\ =&-[-i\Upsilon_{a_1a_3}(x_1-x_3)][-i\Upsilon_{a_2a_4}(x_2-x_4)]\\ &+[-i\Upsilon_{a_1a_4}(x_1-x_4)][-i\Upsilon_{a_2a_3}(x_2-x_3)].\end{aligned}$$

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 from ψ and $\bar\psi$ at quadratic level, because $Z_2[\bar\eta,\eta,J]$ factorizes. For example,

$$\left\langle\phi(x_1)\,\psi_a(x_2)\,\phi(x_3)\bar\psi_b(x_4)\right\rangle = [-i\Delta(x_1-x_3)][-i\Upsilon_{ab}(x_2-x_4)]. \tag{9.1}$$

Graphical representation

It is useful to introduce also a graphical representation. We will represent the scalar propagator by a dashed line,

$$-i\Delta(x-y) = \quad x \text{ - - - - - - } y\ .$$

The Feynman propagator for the fermions will be represented by a solid line with arrow,

$$-i\Upsilon_{ab}(x-y) = \quad (x,a) \longleftarrow (y,b)\ .$$

We can then represent correlation functions graphically, for example, the mixed correlation function in eqn. (9.1) for the free theory would be

$$\begin{aligned}\langle\phi(x_1)\psi_a(x_2)\phi(x_3)\bar\psi_b(x_4)\rangle &= [-i\Delta(x_1-x_3)]\,[-i\Upsilon_{ab}(x_2-x_4)]\\ &= \quad x_1 \text{ - - - - - } x_3 \quad (x_2,a) \longleftarrow (x_4,b)\ .\end{aligned}$$

Blackboard video

Perturbation theory in g

Let us now also consider the interaction terms in the action. In the functional integral it contributes according to

$$e^{iS[\bar\psi,\psi,\phi]} = e^{iS_2[\bar\psi,\psi,\phi]}\ \exp\left[-ig\int d^4x\phi(x)\bar\psi_a(x)\psi_a(x)\right].$$

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 the spinor index a. The resulting expression involving correlation functions can then be evaluated as in the free theory. For example,

$$\begin{aligned}\big\langle\phi(x_1)\psi_b(x_2)\bar\psi_c(x_3)\big\rangle &= \big\langle\phi(x_1)\psi_b(x_2)\bar\psi_c(x_3)\big\rangle_0 \\ &\quad + \left\langle\phi(x_1)\psi_b(x_2)\bar\psi_c(x_3)\left[-ig\int_y\phi(y)\bar\psi_a(y)\psi_a(y)\right]\right\rangle_0 + \dots\end{aligned}$$

The index 0 indicates that the correlation functions get evaluated in the free theory. Graphically, we can represent the interaction term as a vertex,

$$-ig\int_y\sum_a = \quad\text{(y, a)}\quad .$$

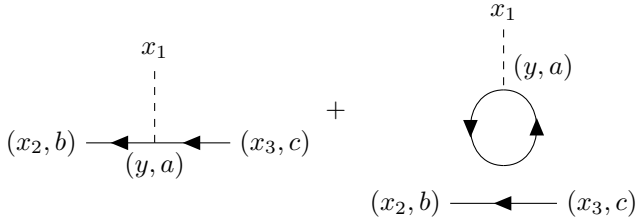
For each such vertex we need to include a factor $-ig$ as well as an integral over the space-time variable y and the spinor index a.

Blackboard video

Three point function

To order g, we find for the example above

$$\langle\phi(x_1)\psi_b(x_2)\bar\psi_c(x_3)\rangle =$$



$$\begin{aligned}&= -ig\int_y[-i\Delta(x_1-y)][-i\Upsilon_{ba}(x_2-y)][-i\Upsilon_{ac}(y-x_3)] \\ &\quad + ig\int_y[-i\Delta(x_1-y)][-i\Upsilon_{bc}(x_2-x_3)][-i\Upsilon_{aa}(y-y)].\end{aligned}$$

The sign in the last line is due to an interchange of Grassmann fields. The last expression involves the fermion propagator for vanishing argument

$$\Upsilon_{ab}(0) = \delta_{ab}\int\frac{d^4p}{(2\pi)^4}\frac{1}{-p^0+\frac{\mathbf{p}^2}{2m}+V_0-i\epsilon} = i\theta(0)\delta_{ab}\delta^{(3)}(0).$$

We will set here $\theta(0)=0$ so that the corresponding contribution vanishes. In other words, we will interpret

$$\Upsilon_{ab}(0) = \lim_{\Delta t\to 0}\Upsilon_{ab}(-\Delta t,\vec 0) = 0.$$

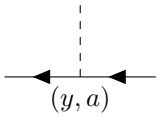
Although this is a little ambiguous at this point, it turns out that this is the right way to proceed.

Blackboard video

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 - - - - - - -
- have a fermion line ending on x for a factor $\psi_a(x)$: (x,a) ⟵
- have a fermion line starting on x for a factor $\bar{\psi}_a(x)$: (x,a) ⟶
- include a vertex $-ig\int_y$ with integral over y for every order g:



- 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 from amputated correlation functions

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

$$i\left[-i\partial_t - \frac{\boldsymbol{\nabla}^2}{2m} + V_0\right]\langle\ldots\psi_a(x)\ldots\rangle$$

and also go to momentum space by a Fourier transform

$$\int_x e^{+i\omega_p x^0 - i\mathbf{p}\mathbf{x}}.$$

The operator simply removes the propagator leading to x, because of

$$i\left[-i\partial_{x^0} - \tfrac{\boldsymbol{\nabla}_x^2}{2m} + V_0\right]\,[-i\Upsilon_{ab}(x-y)] = \delta_{ab}\int\frac{d^4p}{(2\pi)^4}e^{ip(x-y)}\frac{-p^0+\frac{\mathbf{p}^2}{2m}+V_0}{-p^0+\frac{\mathbf{p}^2}{2m}+V_0} = \delta_{ab}\delta^{(4)}(x-y).$$

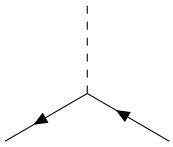
One says that the correlation function is "amputated" because the external propagator has been removed.

Feynman rules for S-matrix elements in momentum space

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 ⟵ $\mathbf{p}$ (to be read from right to left) associated with a momentum $\mathbf{p}$ and energy $\omega_{\mathbf{p}} = \frac{\mathbf{p}^2}{2m} + V_0$.
- Outgoing fermions are represented by an outgoing line $\mathbf{p}$ ⟵

- Incoming or outgoing bosons are represented by - - - - - - - $\mathbf{p}$ and $\mathbf{p}$ - - - - - - - respectively.
- Vertices,



  contribute a factor $-ig$.
- Internal lines that connect two vertices are represented by Feynmann propagators in momentum space, e. g.

$$\xleftarrow{\quad (p^0,\mathbf{p}) \quad} = \frac{-i\delta_{ab}}{-p^0+\frac{\mathbf{p}^2}{2m}+V_0}, \qquad \text{- - - - - - - -}\ (p^0,\mathbf{p}) = \frac{-i}{-(p^0)^2+\mathbf{p}^2+M^2}.$$

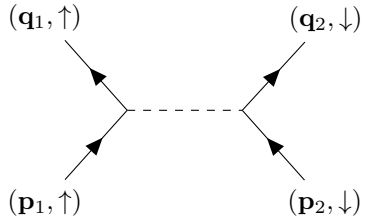
- Energy and momentum conservation are imposed on each vertex.
- For tree diagrams, all momenta are fixed by energy and momenta conservation. For loop diagrams one must include an integral over the loop momentum l_j with measure $d^4l_j/(2\pi)^4$.
- Some care is needed to fix overall signs for fermions.
- Some care is needed to fix overall combinatoric factors from possible interchanges of lines or 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.

Blackboard video

Fermion-fermion scattering

We will now discuss an example, the scattering of (spin polarized) fermions of each other. The tree-level diagram is



Because the interaction with the scalar field does not change the spin, the outgoing fermion with momentum $\mathbf{q}_1$ will have spin $\uparrow$, the one with momentum $\mathbf{q}_2$ will have spin $\downarrow$. By momentum conservation the scalar line carries the four momentum

$$(\omega_{\mathbf{p}_1}-\omega_{\mathbf{q}_1},\ \mathbf{p}_1-\mathbf{q}_1) = \left(\tfrac{\mathbf{p}_1^2}{2m}-\tfrac{\mathbf{q}_1^2}{2m},\mathbf{p}_1-\mathbf{q}_1\right) = (\omega_{\mathbf{q}_2}-\omega_{\mathbf{p}_2},\ \mathbf{q}_2-\mathbf{p}_2).$$

The last equality follows from overall momentum conservation, $p_1+p_2=q_1+q_2$. The Feynmann rules give

$$i\mathcal{T} = (-ig)^2\frac{-i}{-(\omega_{\mathbf{p}_1}-\omega_{\mathbf{q}_1})^2+(\mathbf{p}_1-\mathbf{q}_1)^2+M^2}.$$

In the center-of-mass frame, one has $\omega_{\mathbf{p}_1} = \omega_{\mathbf{p}_2} = \omega_{\mathbf{q}_1} = \omega_{\mathbf{q}_2}$ and thus

$$\mathcal{T} = \frac{g^2}{(\mathbf{p}_1 - \mathbf{q}_1)^2 + M^2}.$$

Blackboard video

Limits of large and small mass

Note that for $g^2 \to \infty$, $M^2 \to \infty$ 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

$$(\mathbf{p}_1 - \mathbf{q}_1)^2 = 2|\mathbf{p}_1|^2(1 - \cos(\vartheta)) = 4|\mathbf{p}_1|^2 \sin^2(\vartheta/2),$$

where we used $|\mathbf{p}_1| = |\mathbf{q}_1|$ in the center of mass frame and ϑ is the angle between $\mathbf{p}_1$ and $\mathbf{q}_1$ (incoming and outgoing momentum of the spin $\uparrow$ particle). For the differential cross-section

$$\frac{d\sigma}{d\Omega_{q_1}} = \frac{|\mathcal{T}|^2 m^2}{16(\pi)^2},$$

we find

$$\frac{d\sigma}{d\Omega_{q_1}} = \frac{g^4 m^2}{16\pi^2} \left[\frac{1}{4\mathbf{p}_1^2 \sin^2(\vartheta/2) + M^2}\right]^2.$$

Another interesting limit is $M^2 \to 0$. One has then

$$\frac{d\sigma}{d\Omega_{q_1}} = \frac{g^4 m^2}{64\pi^2 |\mathbf{p}_1|^4} \frac{1}{\sin^4(\vartheta/2)}.$$

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 $\vartheta \to 0$, and for $\mathbf{p}^2 \to 0$. These are known as colinear and soft singularities. Note that they are regulated by a small, nonvanishing mass $M > 0$.

Blackboard video

LECTURE 18

10 Relativistic fermions

To understand relativistic fermions we need to first understand the properties of the Lorentz group in more detail. Diracs description of relativistic fermions follows then very naturally.

Rotations and Lorentz transformations

We use here conventions where the metric in four dimensional Minkowski space is given by

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1).$$

Infinitesimal Lorentz transformations and rotations in Minkowski space are of the form

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \delta\omega^{\mu}{}_{\nu}, \tag{10.1}$$

with $\Lambda^{\mu}{}_{\nu} \in \mathbb{R}$ such that the metric $\eta_{\mu\nu}$ is invariant, $\eta_{\mu\nu} \to \eta_{\rho\sigma}\Lambda^{\rho}{}_{\mu}\Lambda^{\sigma}{}_{\nu} = \eta_{\mu\nu}$. This implies $(\Lambda^{-1})^{\mu}{}_{\nu} = \Lambda_{\nu}{}^{\mu}$ and, for the inifinitesimal transformation,

$$\delta\omega_{\mu\nu} = -\delta\omega_{\nu\mu}.$$

The spatial-spatial components describe rotations the three dimensional subspace and the spatial-temporal components Lorentz boost in Minkowski space or rotations around a particular three-dimensional direction in Euclidian space.

Blackboard video

Scalar, vector and tensor representations

Lorentz scalars are defined as objects that do not change at all under Lorentz transformations (including rotations). For scalar fields only the argument gets transformed,

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

Lorentz vectors are defined as quantities that get transformed by the matrix Λ. For example, the momentum of a particle transforms as

$$p^{\mu} \to p'^{\mu} = \Lambda^{\mu}{}_{\nu}p^{\nu}.$$

A vector field like for examle the velocity field of a relativistic fluid transforms as

$$u^{\mu}(x) \to u'^{\mu}(x) = \Lambda^{\mu}{}_{\nu}u^{\nu}(\Lambda^{-1}x).$$

In addition to the transformation of the space-time argument there is now an exlicit transformation matrix acting on the index of the field. In a similar way, a covector field like the electromagnetic gauge field transforms according to

$$A_{\mu}(x) \to A'_{\mu}(x) = (\Lambda^{-1})^{\nu}{}_{\mu}A_{\nu}(\Lambda^{-1}x) = \Lambda_{\mu}{}^{\nu}A_{\nu}(\Lambda^{-1}x).$$

One can go on in this way and define tensor field representations, for example a $(2,0)$-tensor field transforms like

$$T^{\rho\sigma}(x) \to T'^{\rho\sigma}(x) = \Lambda^{\rho}{}_{\mu}\Lambda^{\sigma}{}_{\nu}T^{\mu\nu}(\Lambda^{-1}x).$$

In the next step we generalize this concept even further.

Blackboard video

Lie groups and representaions

We consider representations of a group acting on a complex vector space. It can be seen as a map ρ

$$\rho : G \to \mathrm{GL}(n, \mathbb{C}),$$

where $\mathrm{GL}(n, \mathbb{C})$ is the general linear group in N complex dimensions or group of complex $n \times n$ matrices. The map must be such that

$$\rho(g_1)\rho(g_2) = \rho(g_1 g_2),$$

for all $g_1, g_2 \in G$. We are specifically interested in Lie groups where finite transformations can be written in terms of infinitesimal transformations through the exponential map,

$$g = \exp(i\xi^j T_j) = \lim_{N\to\infty} \left(\mathbb{1} + i\frac{\xi^j T_j}{N} \right)^N .$$

Here, T_j are the generators of the Lie algebra. The Lie algebra, and indirectly the Lie group, are characterized by the Lie bracket or commutation relation

$$[T_j, T_k] = i f_{jk}{}^l T_l,$$

where $f_{jk}{}^l$ are the structure constants.

A repesentation of a group element can similarly be written as an exponential map

$$\rho(g) = \exp\left(i\xi^j T_j^{(R)} \right)$$

where $T_j^{(R)}$ are now representations of the Lie algebra generators acting in some vector space. They must have the same Lie bracket relation as the original generators or fundamental representation,

$$\left[T_j^{(R)}, T_k^{(R)} \right] = i f_{jk}{}^l T_l^{(R)}.$$

In this sense one can construct representations of a Lie group by finding representations of the associated Lie algebra.

Complex conjugate representations

For Lie groups where the structure constants are real one can find for representations $T_j^{(R)}$ acting in a complex vector space also the complex conjugate representations

$$T_j^{(C)} = (T_j^{(R)})^\dagger.$$

Indeed this also fulfills the Lie bracket relation as follows by taking the hermitean conjugate on both sides. Sometimes the (representations of the) Lie algebra generators $T_j^{(R)}$ are hermitean already, and in this case the complex conjugate representation is equivalent to the original one, but that is not always the case.

Blackboard video

Lie algebra of Lorentz group

Representations of the Lorentz group with

$$L(\Lambda'\Lambda) = L(\Lambda')L(\Lambda),$$

can be written in infinitesimal form as

$$L(\Lambda) = \mathbb{1} + \frac{i}{2}\delta\omega_{\mu\nu}M^{\mu\nu},$$

where $M^{\mu\nu} = -M^{\nu\mu}$ are the generators of the Lorentz algebra (or Lie algebra associated to the Lorentz group) acting in some representation space with the commutation relation or Lie bracket

$$[M^{\mu\nu}, M^{\rho\sigma}] = i\left(\eta^{\mu\rho}M^{\nu\sigma} - \eta^{\mu\sigma}M^{\nu\rho} - \eta^{\nu\rho}M^{\mu\sigma} + \eta^{\nu\sigma}M^{\mu\rho}\right). \tag{10.2}$$

Blackboard video

Decomposition of Lie algebra

In general, one can decompose the generators into the spatial-spatial part

$$J_i = \frac{1}{2}\epsilon_{ijk}M^{jk}, \tag{10.3}$$

and a spatial-temporal part,

$$K_j = M^{j0}. \tag{10.4}$$

Equation (10.2) implies the commutation relations

$$\begin{aligned}
[J_i, J_j] &= +\,i\,\epsilon_{ijk}J_k,\\
[J_i, K_j] &= +\,i\,\epsilon_{ijk}K_k,\\
[K_i, K_j] &= -i\,\epsilon_{ijk}J_k.
\end{aligned}$$

One can define the linear combinations of generators

$$N_j = \frac{1}{2}(J_j - iK_j), \qquad \tilde{N}_j = \frac{1}{2}(J_j + iK_j),$$

for which the commutation relations become

$$\begin{aligned}
[N_i, N_j] &= i\epsilon_{ijk}N_k,\\
[\tilde{N}_i, \tilde{N}_j] &= i\epsilon_{ijk}\tilde{N}_k,\\
[N_i, \tilde{N}_j] &= 0.
\end{aligned}$$

This shows that the representations of the Lorentz algebra can be decomposed into two representations of SU(2) with generators N_j and $\tilde{N}_j$, respectively.

Blackboard video

Fundamental representation

In the fundamental representation (10.1) one has the generators

$$(M_F^{\mu\nu})^\alpha{}_\beta = -i(\eta^{\mu\alpha}\delta^\nu_\beta - \eta^{\nu\alpha}\delta^\mu_\beta).$$

It acts on the space of four-dimensional vectors p^α and the infinitesimal transformation in (10.1) induces the infinitesimal change

$$\delta p^\alpha = \frac{i}{2}\delta\omega_{\mu\nu}(M_F^{\mu\nu})^\alpha{}_\beta\, p^\beta = \delta\omega^\alpha{}_\beta p^\beta.$$

The generator of rotations in the fundamental representation is

$$(J_i^F)^j{}_k = -i\epsilon_{ijk},$$

where j, k are spatial indices. All other components vanish, $(J_i^F)^0{}_0 = (J_i^F)^0{}_j = (J_i^F)^j{}_0 = 0$. Note that J_i^F is hermitian, $(J_i^F)^\dagger = J_i^F$. The generator K_j has the fundamental representation

$$(K_j^F)^0{}_m = -i\delta_{jm}, \qquad (K_j^F)^m{}_0 = -i\delta_{jm},$$

and all other components vanish, $(K_j^F)^0{}_0 = (K_j^F)^m{}_n = 0$. From these expression one finds that the conjugate of the fundamental representation of the Lorentz algebra has the generators

$$J_j^C = (J_j^F)^\dagger = J_j^F, \qquad K_j^C = (K_j^F)^\dagger = -K_j^F. \tag{10.5}$$

This implies that K_j^F is anti-hermitian,

$$(K_j^F)^\dagger = -K_j^F.$$

Note that N_j and $\tilde{N}_j$ are hermitian and linearly independent in the fundamental reprsentation. There is however an interesting relation between them: Consider the hermitian conjugate representation of the Lorentz group as related to the fundamental one by eq. (10.5). The representation of the generators $N_j, \tilde{N}_j$ is

$$N_j^C = \frac{1}{2}(J_j^C - iK_j^C) = \frac{1}{2}(J_j^F + iK_j^F) = \tilde{N}_j^F,$$
$$\tilde{N}_j^C = \frac{1}{2}(J_j^C + iK_j^C) = \frac{1}{2}(J_j^F - iK_j^F) = N_j^F.$$

This implies that the role of N_j and $\tilde{N}_j$ is interchanged in the conjugate representation.

LECTURE 19

Classification of representaions

Representations of SU(2) are characterized by spin n of half integer or integer value. Accordingly, the representations of the Lorentz group can be classified as $(2n+1, 2\tilde{n}+1)$. For example

$(1,1)$ = scalar or singlet,
$(2,1)$ = left-handed spinor,
$(1,2)$ = right-handed spinor,
$(2,2)$ = vector.

One can also construct tensor product representations as done for a single copy of SU(2) and decompose the latter again in terms of irreducible representations.

Blackboard video

Pauli spinor representation

In the non-relativistic description of spin-1/2 particles due to Pauli, the generators of rotation are

$$J_i = \frac{1}{2}\sigma_i,$$

where the hermitian Pauli matrices are given by

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

and fulfill the algebraic relation

$$\sigma_i \sigma_j = \delta_{ij}\ \mathbb{1} + i\,\epsilon_{ijk}\ \sigma_k.$$

In other words, the Pauli matrices provide a mapping between the space of rotations SO(3) and the space of unitary matrices SU(2). More concrete, an infinitesimal rotation

$$\Lambda^i{}_j = \delta^i{}_j + \delta\omega^i{}_j,$$

corresponds to

$$L(\Lambda) = \mathbb{1} + \frac{i}{4}\delta\omega_{ij}\,\epsilon_{ijk}\,\sigma_k.$$

By exponentiating this one obtains the mapping. Note, however, that the group SU(2) covers SO(3) twice in the sense that a rotation by 360 degrees corresponds to $L(\Lambda) = -\mathbb{1}$.

Blackboard video

Left handed spinor representations

We now construct first the left handed spinor representation of the Lorentz group by using that it agrees with the Pauli representation for normal (spatial) rotations. When acting on the left-handed representation (2,1), the generator $\tilde{N}_j$ vanishes. Since $J_j = N_j + \tilde{N}_j$ and $K_j = i(N_j - \tilde{N}_j)$ one has

$$N_j = J_j = -iK_j = \frac{1}{2}\sigma_j, \qquad \tilde{N}_j = 0.$$

Using (10.3) and (10.4) this yields for the left handed spinor representation

$$\begin{aligned} (M_L^{jk}) &= \epsilon_{jkl} N_l = \frac{1}{2}\epsilon_{jkl}\,\sigma_l, \\ (M_L^{j0}) &= iN_j = i\frac{1}{2}\sigma_j. \end{aligned} \tag{10.6}$$

As the name suggests, this representation acts in the space of left-handed spinors which are two-components entities, for example

$$\psi_L = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}.$$

We also use a notation with explicit indices ψ_a with $a = 1, 2$. The infinitesimal transformation in (10.1) reads with the matrices (10.6)

$$\delta\psi_a = \frac{i}{2}\delta\omega_{\mu\nu}(M_L^{\mu\nu})_a{}^b\ \psi_b. \tag{10.7}$$

Blackboard video

Map to special linear group SL$(2,\mathbb{C})$

From the concrete matrix representation we find that

$$\frac{i}{2}\delta\omega_{\mu\nu}(M_L^{\mu\nu})_a{}^b = \frac{i}{2}(\delta a_k + i\delta b_k)(\sigma_k)_a{}^b$$

with some coefficients δa_l and δb_l. The right hand side is a general, complex but traceless 2×2 matrix and therefore an element of the Lie algebra of $\mathrm{SL}(2,\mathbb{C})$. Accordingly there is a map from the Lie algebra of the Lorentz group $\mathrm{SO}(1,3,\mathbb{R})$ to the Lie algebra of $\mathrm{SL}(2,\mathbb{C})$. In fact, there is an isomorphism between groups $\mathrm{SL}(2,\mathbb{C})/Z_2 = \mathrm{SO}(1,3,\mathbb{R})$. (Excercise: Construct the corresponding map between group elements explicitely.)

Blackboard video

Tensor representations

One can also construct spinors in tensor product representations. They have several indices and transform accordingly. For example, a spinor with two left-handed indices transforms according to

$$\delta\chi_{ab} = \frac{i}{2}\delta\omega_{\mu\nu}\left[(M_L^{\mu\nu})_a{}^c\delta_b{}^d + \delta_a{}^c(M_L^{\mu\nu})_b{}^d\right]\chi_{cd}.$$

One can decompose such a spinor with two indices into an anti-symmetric and a symmetric part, corresponding to the decomposition

$$(2,1)\otimes(2,1) = (1,1)_A \oplus (3,1)_S. \tag{10.8}$$

This is a decomposition into a spin singlet and a spin triplet or spin one representation.

Invariant symbol in left-handed singlet

From (10.8) it follows that there must be a Lorentz-singlet with two left-handed spinor indices and it has to be anti-symmetric. The corresponding invariant symbol can be taken as ε_{ab} with components $\varepsilon_{21}=1$, $\varepsilon_{12}=-1$ and $\varepsilon_{11}=\varepsilon_{22}=0$. Indeed one finds from the concrete presentation that

$$(M_L^{\mu\nu})_a{}^c\varepsilon_{cb} + (M_L^{\mu\nu})_b{}^c\varepsilon_{ac} = 0. \tag{10.9}$$

This is essentialy due to $\sigma_j\sigma_2 + \sigma_2\sigma_j^T = 0$ for $j=1,2,3$ as can be checked easily. For clarity the non-vanishing components are

$$\varepsilon^{12} = -\varepsilon^{21} = \varepsilon_{21} = -\varepsilon_{12} = 1. \tag{10.10}$$

As an invariant symbol, ε_{ab} plays for SU(2) a similar role as the metric $\eta_{\mu\nu}$ for the Lorentz group. It is natural to use ε_{ab} and its inverse ε^{ab} to pull the indices a,b,c up and down. We can write for example a left-handed spinor with upper spinor index

$$\psi^a = \varepsilon^{ab}\psi_b = -\varepsilon^{ba}\psi_b = -\psi_b\varepsilon^{ba} = \psi_b\varepsilon^{ab}.$$

(We wrote different equivalent expressions to show that care is needed here with minus signs.) From eq. (10.9) it follows also that

$$(M_L^{\mu\nu})_a{}^b = (M_L^{\mu\nu})^b{}_a,$$

as well as

$$(M_L^{\mu\nu})_{ab} = (M_L^{\mu\nu})_{ba},$$

so that

$$\varepsilon^{ab}(M_L^{\mu\nu})_{ab} = (M_L^{\mu\nu})_a{}^a = 0.$$

The symbol δ_b^a is also invariant when spinors with upper left-handed indices have the Lorentz-transformation behavior

$$\delta\psi^a = -\frac{i}{2}\delta\omega_{\mu\nu}\psi^b(M_L^{\mu\nu})_b{}^a = -\frac{i}{2}\delta\omega_{\mu\nu}(M_L^{\mu\nu})^a{}_b\psi^b.$$

This implies also that upper and lower indices can be contracted, for example

$$\psi^a\chi_a = \varepsilon^{ab}\psi_b\chi_a = -\varepsilon^{ba}\psi_b\chi_a = -\psi_b\chi^b,$$

is invariant. (Again we wrote several equivalent expressions to show that care is needed with signs.)

Right handed spinor representation

Similarly one finds for the right-handed spinor representation (1,2) using

$$N_j = 0, \qquad \tilde{N}_j = J_j = iK_j = \frac{1}{2}\sigma_j,$$

the relations

$$\begin{aligned}(M_R^{jk}) &= \epsilon_{jkl}\tilde{N}_l = \frac{1}{2}\epsilon_{jkl}\,\sigma_l \\ (M_R^{j0}) &= -i\tilde{N}_j = -i\frac{1}{2}\sigma_j.\end{aligned} \tag{10.11}$$

The representation (10.11) acts in the space of right handed spinors, for example

$$\bar{\psi} = \begin{pmatrix}\bar{\psi}^1 \\ \bar{\psi}^2\end{pmatrix}.$$

For right handed spinors we will also use a notation with an explicit index that has a dot in order to distinguish it from a left-handed index, i. e. $\bar{\psi}^{\dot{a}}$ with $\dot{a} = 1, 2$ denotes a right-handed spinor. The infinitesimal transformation in (10.1) reads with the matrices in (10.11)

$$\delta\bar{\psi}^{\dot{a}} = \frac{i}{2}\delta\omega_{\mu\nu}(M_R^{\mu\nu})^{\dot{a}}{}_{\dot{b}}\;\bar{\psi}^{\dot{b}}.$$

Blackboard video

Invariant symbol in right-handed singlet

In a completely analogous way the relation

$$(1,2)\times(1,2) = (1,1)_A + (1,3)_S$$

implies that there is a Lorentz singlet with two right-handed spinor indices. The corresponding symbol can be taken as $\varepsilon^{\dot{a}\dot{b}}$, with inverse $\varepsilon_{\dot{a}\dot{b}}$, with components as in (10.10). This symbol is used to lower and raise right-handed indices. Spinors with lower right handed index transform under Lorentz-transformations as

$$\delta\bar{\psi}_{\dot{a}} = -\frac{i}{2}\delta\omega_{\mu\nu}\bar{\psi}_{\dot{b}}(M_R^{\mu\nu})^{\dot{b}}{}_{\dot{a}} = -\frac{i}{2}\delta\omega_{\mu\nu}(M_R^{\mu\nu})_{\dot{a}}{}^{\dot{b}}\bar{\psi}_{\dot{b}}.$$

Invariant symbols for vectors

Consider now an object with a left-handed and a right-handed index. It is in the representation $(2,2)$ which should also contain the vector. There is therefore an invariant symbol which can be chosen as

$$(\sigma^\mu)_{a\dot{a}} = (\mathbb{1}, \boldsymbol{\sigma}),$$

and similarly

$$(\bar{\sigma}^\mu)^{\dot{a}a} = (\mathbb{1}, -\boldsymbol{\sigma}).$$

These symbols are invariant in the sense that they get mapped to themselves when all indices are transformed appropriately, e. g.

$$\Lambda^\mu{}_\nu L_a{}^b R_{\dot{a}}{}^{\dot{b}} (\sigma^\nu)_{b\dot{b}} = (\sigma^\mu)_{a\dot{a}}.$$

Here $L_a{}^b$ and $R_{\dot{a}}{}^{\dot{b}}$ are finite Lorentz transformation matrices in appropriate representations for left handed and right handed spinors respectively.

It turns out that the matrices for infinitesimal Lorentz transformations can be written as

$$\begin{aligned}(M_L^{\mu\nu})_a{}^b &= \frac{i}{4}(\sigma^\mu\bar{\sigma}^\nu - \sigma^\nu\bar{\sigma}^\mu)_a{}^b,\\ (M_R^{\mu\nu})^{\dot{a}}{}_{\dot{b}} &= \frac{i}{4}(\bar{\sigma}^\mu\sigma^\nu - \bar{\sigma}^\nu\sigma^\mu)^{\dot{a}}{}_{\dot{b}}.\end{aligned}$$

Some useful identities are

$$\begin{aligned}(\sigma^\mu)_{a\dot{a}}(\sigma_\mu)_{b\dot{b}} &= -2\,\varepsilon_{ab}\varepsilon_{\dot{a}\dot{b}}\,,\\ (\bar{\sigma}^\mu)^{\dot{a}a}(\bar{\sigma}_\mu)^{\dot{b}b} &= -2\varepsilon^{ab}\varepsilon^{\dot{a}\dot{b}},\\ \varepsilon^{ab}\varepsilon^{\dot{a}\dot{b}}(\sigma^\mu)_{a\dot{a}}(\sigma^\nu)_{b\dot{b}} &= -2\,\eta^{\mu\nu},\\ (\bar{\sigma}^\mu)^{\dot{a}a} &= \varepsilon^{ab}\varepsilon^{\dot{a}\dot{b}}(\sigma^\mu)_{b\dot{b}},\\ (\sigma^\mu\bar{\sigma}^\nu + \sigma^\nu\bar{\sigma}^\mu)_a{}^b &= -2\,\eta^{\mu\nu}\delta_a^b,\\ \mathrm{Tr}(\sigma^\mu\bar{\sigma}^\nu) &= \mathrm{Tr}(\bar{\sigma}^\mu\sigma^\nu) = -2\,\eta^{\mu\nu},\\ \bar{\sigma}^\mu\sigma^\nu\bar{\sigma}_\mu &= 2\,\bar{\sigma}^\nu,\\ \sigma^\mu\bar{\sigma}^\nu\sigma_\mu &= 2\,\sigma^\nu.\end{aligned}$$

We leave it as an excercise to prove these.

Complex conjugation

The matrices (10.6) and (10.11) are hermitian conjugate of each other, i. e.

$$(M_L^{\mu\nu})^\dagger = M_R^{\mu\nu}, \qquad (M_R^{\mu\nu})^\dagger = M_L^{\mu\nu}.$$

The hermitian conjugate of the Lorentz transformation (10.7) is given by

$$[\delta\psi_a]^\dagger = -\frac{i}{2}\delta\omega^*_{\mu\nu}\,[\psi_b]^\dagger \underbrace{\left[(M_L^{\mu\nu})_a{}^b\right]^\dagger}_{=(M_R^{\mu\nu})^b{}_{\dot{a}}}. \tag{10.12}$$

For $\delta\omega_{\mu\nu} \in \mathbb{R}$ this is of the same form as eq. (10). In Minkowski space it is therefore consistent to take $\psi^\dagger$ to be a right-handed spinor with lower dotted index, we write

$$[\psi_a]^\dagger = (\psi^\dagger)_{\dot{a}}, \qquad \text{(Minkowski)}$$

and in an analogous way one finds that it is consistent to write

$$\left[\bar{\psi}^{\dot{a}}\right]^{\dagger} = (\bar{\psi}^{\dagger})^{a}, \qquad \text{(Minkowski)}$$

which is a left-handed spinor. So far we have considered Minkowski space only. In Euclidean space or for more general complex $\delta\omega_{\mu\nu}$ the hermitian conjugation is more complicated. For complex $\delta\omega_{\mu\nu}$ eq. (10.12) constitutes a transformation behavior that is not of any already discussed type. For a consistent analytic continuation it is actually necessay to have all fields transforming such that the infinitesimal transformation law involves only $\delta\omega_{\mu\nu}$ (and not $\delta\omega^{*}_{\mu\nu}$).

Dirac spinors in chiral basis

Dirac spinors are composed of a left handed and a right handed spinor. In the chiral basis they read

$$\Psi = \begin{pmatrix} \psi_a \\ \bar{\xi}^{\dot{a}} \end{pmatrix}, \qquad \bar{\Psi} = \left(\xi^{a}, \bar{\psi}_{\dot{a}}\right). \tag{10.13}$$

Note that, as the notation suggests, ξ^{a} transforms as a left handed spinor and $\bar{\psi}_{\dot{a}}$ as a right-handed one. One should see Ψ and the Dirac conjugate $\bar{\Psi}$ as independent fields and they are in any case represented as independent Grassmann variables. In Minkowski space one can identify

$$\bar{\Psi} = \Psi^{\dagger}\beta, \qquad \Psi = \beta^{-1}\bar{\Psi}^{\dagger} \qquad \text{(Minkowski space)}.$$

with

$$\beta = \begin{pmatrix} & \delta^{\dot{a}}{}_{\dot{b}} \\ \delta_{a}{}^{b} & \end{pmatrix}, \qquad \beta^{-1} = \begin{pmatrix} & \delta_{a}{}^{b} \\ \delta^{\dot{a}}{}_{\dot{b}} & \end{pmatrix}.$$

This is useful to ckeck that actions are hermitean or real and lead to unitary time evolution.

Blackboard video

Blackboard video

Clifford algebra

The gamma matrices are in this representation given by[1]

$$\gamma^{\mu} = \begin{pmatrix} & -i(\sigma^{\mu})_{a\dot{b}} \\ -i(\bar{\sigma}^{\mu})^{\dot{a}b} & \end{pmatrix}.$$

They fulfill an anti-commutation relation

$$\{\gamma^{\mu}, \gamma^{\nu}\} = \gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu} = 2\eta^{\mu\nu}. \tag{10.14}$$

[1] This is the convention of Weinberg and Wetterich. There is an alternative definition of the gamma matrices often used in the literature (for example Peskin & Schroeder, Srednicki, Elvang & Huang) where one has

$$\gamma^{\mu} = \begin{pmatrix} & (\sigma^{\mu})_{a\dot{a}} \\ (\bar{\sigma}^{\mu})^{\dot{a}a} & \end{pmatrix}, \qquad \{\gamma^{\mu}, \gamma^{\nu}\} = -2\,\eta^{\mu\nu}.$$

An anti-commutation relation like (10.14) defines a *Clifford algebra* and can also be taken as the starting point for the coonstruction of spinor representations of the Lorentz group. Note that one may redefine the gamma matrices and spinors

$$\gamma^\mu \to U\gamma^\mu U^\dagger, \qquad \Psi \to U\Psi,$$

to obtain another representation of the Clifford algebra that works equally well.

Blackboard video

Antisymmetric matrices

Define the commutator of gamma-matrices as[2]

$$\sigma^{\mu\nu} = -\frac{i}{2}\left[\gamma^\mu, \gamma^\nu\right] = \begin{pmatrix} \frac{i}{2}(\sigma^\mu\bar\sigma^\nu - \sigma^\nu\bar\sigma^\mu)_a{}^b & \\ & \frac{i}{2}(\bar\sigma^\mu\sigma^\nu - \bar\sigma^\nu\sigma^\mu)^{\dot a}{}_{\dot b} \end{pmatrix}.$$

It is useful to define also

$$(\tau^{\mu\nu})_a{}^b = \frac{i}{2}(\sigma^\mu\bar\sigma^\nu - \sigma^\nu\bar\sigma^\mu)_a{}^b,$$
$$(\bar\tau^{\mu\nu})^{\dot a}{}_{\dot b} = \frac{i}{2}(\bar\sigma^\mu\sigma^\nu - \bar\sigma^\nu\sigma^\mu)^{\dot a}{}_{\dot b},$$

such that

$$\sigma^{\mu\nu} = \begin{pmatrix} \tau^{\mu\nu} & \\ & \bar\tau^{\mu\nu} \end{pmatrix}.$$

As matrix, $\sigma^{\mu\nu}$ can be written in terms of Pauli matrices,

$$\sigma^{ij} = \epsilon^{ijk}\begin{pmatrix} \sigma^k & \\ & \sigma^k \end{pmatrix},$$
$$\sigma^{j0} = -\sigma^{0j} = \begin{pmatrix} i\sigma^j & \\ & -i\sigma^j \end{pmatrix}.$$

Commutation relations with gamma matrices are

$$\begin{aligned}\frac{1}{2}\left[\sigma^{\mu\nu}, \gamma^\rho\right] &= -\frac{i}{4}\left[\left[\gamma^\mu, \gamma^\nu\right], \gamma^\rho\right] \\ &= -\frac{i}{4}\left(\{\gamma^\mu, \{\gamma^\nu, \gamma^\rho\}\} - \{\gamma^\nu, \{\gamma^\mu, \gamma^\rho\}\}\right) \\ &= -i\left(\gamma^\mu\eta^{\nu\rho} - \gamma^\nu\eta^{\mu\rho}\right).\end{aligned}$$

This can be understood as a kind of adjoint representation of the Lie algebra of the Lorentz group.

One can write the Lie algebra generators acting on Dirac spinors directly as

$$M^{\mu\nu} = \frac{1}{2}\sigma^{\mu\nu}.$$

This definition works also in other representations of the Clifford algebra because only the defining anticommutation relation (10.14) has been used. In other words, we could have started with some representation of (10.14) and would have obtained automatically a spinor representation of the Lorentz group. This would have been a reducible representation though, because it involves left-handed and right-handed spinors simultaneously. The representation can be decmposed again into its parts with gamma five, which we introduce now.

[2] These are the conventions of Peskin & Schroeder. Weinberg uses $\mathcal{J}^{\mu\nu} = \frac{1}{2}\sigma^{\mu\nu}$. Wetterich uses the opposite sign

Blackboard video

Gamma five

The matrix γ_5 is defined as[3]

$$\begin{aligned}\gamma_5 &= -i\gamma^0\gamma^1\gamma^2\gamma^3 \\ &= \frac{i}{4!}\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma .\end{aligned}$$

The four-dimensional Levi-Civita symbol is completely anti-symmetric with $\epsilon_{0123} = -1$. In chiral representation one has

$$\gamma^5 = \begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} = \begin{pmatrix} \delta_a{}^b & \\ & -\delta^{\dot a}{}_{\dot b} \end{pmatrix} .$$

The projectors to the left- and right-handed spinors are

$$P_L = \frac{1}{2}(\mathbb{1} + \gamma_5) = \begin{pmatrix} \delta_a{}^b & 0 \\ 0 & 0 \end{pmatrix} ,$$

and

$$P_R = \frac{1}{2}(\mathbb{1} - \gamma_5) = \begin{pmatrix} 0 & 0 \\ 0 & \delta^{\dot a}{}_{\dot b} \end{pmatrix} .$$

They can be used to project Dirac spinors with four components to Weyl spinors with only two nonvanishing components.

Due to the definitions one has

$$\begin{aligned}\{\gamma_5, \gamma^\mu\} &= 0 \\ [\gamma_5, \sigma^{\mu\nu}] &= 0 \\ (\gamma_5)^2 &= \mathbb{1}.\end{aligned}$$

Blackboard video

Charge conjugation

Consider a Dirac spinor and its conjugate as in eq. (10.13). The charge conjugate spinors are defined as

$$\Psi_C = \begin{pmatrix} \xi_a \\ \bar\psi^{\dot a} \end{pmatrix} , \qquad \bar\Psi_C = \left(\psi^a, \bar\xi_{\dot a}\right) .$$

A *Majorna spinor* is a Dirac spinor which obeys $\psi_C = \psi$ or, in terms of Weyl spinors, $\xi_a = \psi_a$ and $\bar\psi^{\dot a} = \bar\xi^{\dot a}$. In other words, we can write the four component Majorana spinor as

$$\Psi = \Psi_C = \begin{pmatrix} \psi_a \\ \bar\psi^{\dot a} \end{pmatrix} .$$

The relation between Dirac spinors and Majorana spinors is as the relation between a complex and a real scalar field. As one can construct complex scalar fields out of two real fields, one can construct

[3] These is the convention of Weinberg, Wetterich. Peskin & Schroeder, Srednicki define γ_5 with opposite sign

Dirac spinors out of Majorana spinors (exercise). Note that a Majorana spinor has formally as many real degrees of freedom as a Weyl spinor.

The charge conjugate fields can be written as

$$\Psi_{\mathsf{C}} = \mathscr{C}\bar{\Psi}^T, \qquad \bar{\Psi}_{\mathsf{C}} = -\Psi^T\mathscr{C}^{-1}, \tag{10.15}$$

with the transpose spinors

$$\bar{\Psi}^T = \begin{pmatrix} \xi^a \\ \bar{\psi}_{\dot{a}} \end{pmatrix}, \qquad \Psi^T = \left(\psi_a, \bar{\xi}^{\dot{a}}\right)$$

and the charge conjugation matrix[4]

$$\mathscr{C} = \begin{pmatrix} \varepsilon_{ab} & \\ & \varepsilon^{\dot{a}\dot{b}} \end{pmatrix}, \qquad \mathscr{C}^{-1} = \begin{pmatrix} \varepsilon^{ab} & \\ & \varepsilon_{\dot{a}\dot{b}} \end{pmatrix}.$$

As a matrix, $\mathscr{C}$ obeys

$$\mathscr{C} = -\mathscr{C}^{-1} = -\mathscr{C}^\dagger = -\mathscr{C}^T = \mathscr{C}^* = \begin{pmatrix} & -1 & & \\ 1 & & & \\ & & & 1 \\ & & -1 & \end{pmatrix}.$$

One has also

$$\begin{aligned} \mathscr{C}^{-1}\gamma^\mu\mathscr{C} &= \begin{pmatrix} \varepsilon^{ab} & \\ & \varepsilon_{\dot{a}\dot{b}} \end{pmatrix} \begin{pmatrix} & -i(\sigma^\mu)_{b\dot{c}} \\ -i(\bar{\sigma}^\mu)^{\dot{b}c} & \end{pmatrix} \begin{pmatrix} \varepsilon_{cd} & \\ & \varepsilon^{\dot{c}\dot{d}} \end{pmatrix} \\ &= \begin{pmatrix} & -i\varepsilon^{ab}(\sigma^\mu)_{b\dot{c}}\varepsilon^{\dot{c}\dot{d}} \\ -i\varepsilon_{\dot{a}\dot{b}}(\bar{\sigma}^\mu)^{\dot{b}c}\varepsilon_{cd} & \end{pmatrix} \\ &= \begin{pmatrix} & i(\bar{\sigma}^\mu)^{\dot{d}a} \\ i(\sigma^\mu)_{d\dot{a}} & \end{pmatrix}, \end{aligned}$$

or, in matrix notation

$$\mathscr{C}^{-1}\gamma^\mu\mathscr{C} = -(\gamma^\mu)^T. \tag{10.16}$$

Similarly,

$$\begin{aligned} \mathscr{C}^{-1}\sigma^{\mu\nu}\mathscr{C} &= -(\sigma^{\mu\nu})^T, \\ \mathscr{C}^{-1}\gamma_5\mathscr{C} &= (\gamma_5)^T, \\ \mathscr{C}^{-1}\gamma^5\gamma^\mu\mathscr{C} &= (\gamma^5\gamma^\mu)^T. \end{aligned} \tag{10.17}$$

The index structure in (10.16) and (10.17) is appropriate for transposed spinors.

Blackboard video

Parity

Parity transforms the coordinates as $(t, \vec{x}) \to (t, -\vec{x})$. For a Dirac spinor and its conjugate as in (10.13) one defines the parity transformed spinors

$$\Psi_{\mathsf{P}} = \begin{pmatrix} i\bar{\xi}^{\dot{a}} \\ i\psi_a \end{pmatrix}, \qquad \bar{\Psi}_{\mathsf{P}} = \left(i\bar{\psi}_{\dot{a}}, i\xi^a\right). \tag{10.18}$$

[4] This is the convention of Srednicki

Note that the role of right-handed and left-handed spinors (and their corresponding indices) is interchanged for the parity-transformed field. Eq. (10.18) can also be written as

$$\Psi_{\mathsf{P}} = i\beta\Psi \qquad \bar{\Psi}_{\mathsf{P}} = -i\bar{\Psi}\beta^{-1},$$

with the matrix

$$\beta = \begin{pmatrix} & \delta^{\dot{a}}{}_{\dot{b}} \\ \delta_a{}^b & \end{pmatrix}, \qquad \beta^{-1} = \begin{pmatrix} & \delta_a{}^b \\ \delta^{\dot{a}}{}_{\dot{b}} & \end{pmatrix}.$$

As a matrix, β obeys

$$\beta = \beta^{-1} = \beta^\dagger = \beta^T = \beta^* = \begin{pmatrix} & \mathbb{1} \\ \mathbb{1} & \end{pmatrix}.$$

Parity transformations of the gamma matrices are given by

$$\begin{aligned} \beta\gamma^j\beta^{-1} &= -\gamma^j, \\ \beta\gamma^0\beta^{-1} &= \gamma^0. \end{aligned} \tag{10.19}$$

The gamma matrices on the right hand side of (10.19) agree with (10) as matrices but have a different index structure such that they fit to the spinors in (10.18),

$$\beta\gamma^\mu\beta^{-1} = \begin{pmatrix} & -i(\bar{\sigma}^\mu)^{\dot{a}b} \\ -i(\sigma^\mu)_{a\dot{b}} & \end{pmatrix}.$$

When doing a parity transform of an expression (e.g. a Lagrangian) and replacing spinors with parity transformed spinors, one should also replace gamma matrices by the expressions in (10.19). Similarly for the antisymmetric matrices

$$\begin{aligned} \beta\sigma^{ij}\beta^{-1} &= \sigma^{ij}, \\ \beta\sigma^{j0}\beta^{-1} &= -\sigma^{j0}. \end{aligned}$$

The matrix γ_5 is a pseudoscalar in the sense

$$\beta\gamma_5\beta^{-1} = -\gamma_5.$$

Time reversal

Time reversal changes the time direction, $(t, \vec{x}) \to (-t, \vec{x})$. It is also a anti-unitary transformation that transforms all complex numbers to there complex conjugates. The time-reversed version of the Dirac spinor and its conjugate as in (10.13) is given by

$$\Psi_{\mathsf{T}} = \begin{pmatrix} \psi^a \\ -\bar{\xi}_{\dot{a}} \end{pmatrix}, \qquad \bar{\Psi}_{\mathsf{T}} = \left(-\xi_a, \bar{\psi}^{\dot{a}}\right). \tag{10.20}$$

Note that the role of upper and lower indices has been interchanged. With the matrices $\mathscr{C}$ and γ_5 one can write this as

$$\Psi_{\mathsf{T}} = \mathscr{C}^{-1}\gamma_5\Psi, \qquad \bar{\Psi}_{\mathsf{T}} = \bar{\Psi}\gamma_5\mathscr{C}.$$

When considering time-reversal transformations, the following identities are useful

$$\begin{aligned} \mathscr{C}^{-1}\gamma^0\mathscr{C}^{-1} &= (\gamma^0)^* \\ \mathscr{C}^{-1}\gamma^j\mathscr{C}^{-1} &= -(\gamma^j)^* \\ \mathscr{C}^{-1}\gamma_5\mathscr{C}^{-1} &= (\gamma_5)^* \\ \mathscr{C}^{-1}\sigma^{ij}\mathscr{C}^{-1} &= -(\sigma^{ij})^* \\ \mathscr{C}^{-1}\sigma^{j0}\mathscr{C}^{-1} &= (\sigma^{j0})^*. \end{aligned}$$

Similarly as in the case of parity, the index structure of these expressions is such that it fits to eq. (10.20). In other words, one should use these expressions for $(\gamma^\mu)^*$ ect. in connection with time-reversed spinors.

Grassmann property of fields and index-free notation

Define products of left-handed two-component Weyl spinors as

$$\chi\psi = \chi^a\psi_a, \tag{10.21}$$

and similarly for right handed spinors,

$$\bar{\chi}\bar{\psi} = \bar{\chi}_{\dot{a}}\bar{\psi}^{\dot{a}}. \tag{10.22}$$

In other words, contracted indices that are not written should be interpreted as ${}^{c}{}_{c}$ for left-handed spinors and ${}_{\dot{c}}{}^{\dot{c}}$ for right-handed spinors. Since χ and ψ are Grassmann valued their components anti-commute, e. g.

$$\chi\psi = \chi^a\psi_a = -\psi_a\chi^a = \psi^a\chi_a = \psi\chi. \tag{10.23}$$

So with this notation one has $\chi\psi = \psi\chi$ and similarly $\bar{\chi}\bar{\psi} = \bar{\psi}\bar{\chi}$ etc.

Hermitian conjugation includes also a commutation of Grassmann numbers, e. g.

$$(\chi\psi)^\dagger = (\chi^a\psi_a)^\dagger = [\psi_a]^\dagger\,[\chi^a]^\dagger = (\psi^\dagger)_{\dot{a}}(\chi^\dagger)^{\dot{a}} = \psi^\dagger\chi^\dagger.$$

Another example for manipulating spinor indices is

$$\begin{aligned}\bar{\psi}\bar{\sigma}^\mu\chi = \bar{\psi}_{\dot{a}}(\bar{\sigma}^\mu)^{\dot{a}b}\chi_b &= -\chi_b\left[\epsilon^{\dot{a}\dot{c}}\epsilon^{bd}(\sigma^\mu)_{d\dot{c}}\right]\bar{\psi}_{\dot{a}} \\ &= -\chi^d(\sigma^\mu)_{d\dot{c}}\bar{\psi}^{\dot{c}} = -\chi\sigma^\mu\bar{\psi}.\end{aligned} \tag{10.24}$$

The minus sign in the second equation is due to the Grassmann property and the interchange of spinors.

Blackboard video

LECTURE 20

Lagrangian for Weyl fermions

A Lagrangian for a left-handed, two-component Weyl fermion ψ_a and its right-handed hermitean conjugate $\bar{\psi}_{\dot{a}} = \psi^\dagger_{\dot{a}}$ in Minkowski space can be written as

$$\mathscr{L} = i\bar{\psi}_{\dot{a}}(\bar{\sigma}^\mu)^{\dot{a}b}\partial_\mu\psi_b - \frac{1}{2}m\psi^a\psi_a - \frac{1}{2}m\bar{\psi}_{\dot{a}}\bar{\psi}^{\dot{a}} = i\bar{\psi}\bar{\sigma}^\mu\partial_\mu\psi - \frac{1}{2}m\psi\psi - \frac{1}{2}m\bar{\psi}\bar{\psi}. \tag{10.25}$$

In the second equation we used the short hand notation introduced above, keeping in mind that we deal here with two-component spinors.

The two mass terms go into each other under hermitian conjugation assuming real m. More general one could allow them to have complex conjugate masses m and m^* respectively, but the complex phase of $m = e^{i\beta}|m|$ can be absorbed into a redefinition $\psi \to e^{-i\beta/2}\psi$, $\bar{\psi} \to e^{i\beta/2}\bar{\psi}$, so that real $m > 0$ can be assumed without loss of generality.

On the other side, if the theory is supposed to be invariant under the U(1) symmetry $\psi \to e^{i\alpha}\psi$, $\bar{\psi} \to e^{-i\alpha}\bar{\psi}$, a mass term as in (10.25) is actually excluded. In other words, such a symmetry would only be unbroken for $m = 0$.

Let us also consider the hermitian conjugate of the first, kinetic term in the Lagrangian,

$$\begin{aligned}\left[i\bar{\psi}_{\dot{a}}(\bar{\sigma}^\mu)^{\dot{a}b}\partial_\mu\psi_b\right]^\dagger &= -i\partial_\mu[\psi_b]^\dagger\left[(\bar{\sigma}^\mu)^{\dot{a}b}\right]^*[\bar{\psi}_{\dot{a}}]^\dagger = -i\partial_\mu[\psi_b]^\dagger(\bar{\sigma}^\mu)^{b\dot{a}}[\bar{\psi}_{\dot{a}}]^\dagger = -i\partial_\mu\bar{\psi}_{\dot{b}}(\bar{\sigma}^\mu)^{\dot{b}a}\psi_a \\ &= i\bar{\psi}_{\dot{b}}(\bar{\sigma}^\mu)^{\dot{b}a}\partial_\mu\psi_a - \partial_\mu\left[i\bar{\psi}_{\dot{b}}(\bar{\sigma}^\mu)^{\dot{b}a}\psi_a\right].\end{aligned}$$

The last term on the right is a total derivative and contributes only an irrelvant boundary term in the action. This shows that the kinetic term in also hermitian. Note also that using (10.24) with $\chi = \partial_\mu \psi$ and droping another boundary term from partial integration, one can also bring the kinetic term into the form

$$i\psi\sigma^\mu\partial_\mu\bar{\psi}.$$

This shows that (10.25) is as well a Lagrangian for the left-handed field ψ_a as for the right-handed field $\bar{\psi}^{\dot{a}}$.

Blackboard video

Lagrangian for Majorana fermions

We can now also write down directly a Lagrangian for four-component Majorana spionors,

$$\Psi = \Psi_C = \begin{pmatrix} \psi_a \\ \bar{\psi}^{\dot{a}} \end{pmatrix}, \qquad \bar{\Psi} = \bar{\Psi}_C = \left(\psi^a, \bar{\psi}_{\dot{a}}\right).$$

In fact, the Lagrangian (10.25) can be rewritten in terms of a Majorana spinor Ψ as

$$\mathscr{L} = -\frac{1}{2}\bar{\Psi}_C\gamma^\mu\partial_\mu\Psi - \frac{1}{2}m\bar{\Psi}_C\Psi = \frac{1}{2}\Psi^T\mathscr{C}^{-1}\gamma^\mu\partial_\mu\Psi + \frac{1}{2}m\Psi^T\mathscr{C}^{-1}\Psi. \tag{10.26}$$

In the second equation we wrote everything in terms of the spinor Ψ only, using (10.15), to make explicit that there is only one independant spionor field here.

Blackboard video

Lagrangian for Dirac fermions

Dirac fermions represented by the spinors

$$\Psi = \begin{pmatrix} \psi_b \\ \bar{\xi}^{\dot{b}} \end{pmatrix}, \qquad \bar{\Psi} = \left(\xi^b, \bar{\psi}_{\dot{b}}\right),$$

are charged fermions. This means one can do a U(1) transformation

$$\Psi \to e^{i\alpha}\Psi, \qquad \bar{\Psi} \to \bar{\Psi}e^{-i\alpha},$$

or

$$\psi_b \to e^{i\alpha}\psi_b, \qquad \bar{\xi}^{\dot{b}} \to e^{i\alpha}\bar{\xi}^{\dot{b}}, \qquad \xi^b \to e^{-i\alpha}\xi^b, \qquad \bar{\psi}_{\dot{b}} \to e^{-i\alpha}\bar{\psi}_{\dot{b}}. \tag{10.27}$$

This implies that a mass term as for the Majorana fermions is not allowed. However, for Dirac spinors, which have twice as many degrees of freedom as the Weyl or Majorana fermions, it is possible to include a differnt kind of mass term, involving the combination

$$-m\bar{\Psi}\Psi = -m\left[\xi^a\psi_a + \bar{\psi}_{\dot{a}}\bar{\xi}^{\dot{a}}\right]. \tag{10.28}$$

Note that this mixes the spinors ψ and ξ and is only allowed because they have opposite charges under the U(1) transformation in eq. (10.27). In turn a mass term as in (10.28) actually breaks another U(1) symmetry, the so-called chiral or axial symmetry

$$\psi_b \to e^{i\beta}\psi_b, \qquad \bar{\xi}^{\dot{b}} \to e^{-i\beta}\bar{\xi}^{\dot{b}}, \qquad \xi^b \to e^{i\beta}\xi^b, \qquad \bar{\psi}_{\dot{b}} \to e^{-i\beta}\bar{\psi}_{\dot{b}}.$$

In terms of gamma five this can be written as

$$\Psi \to e^{i\beta\gamma_5}\Psi, \qquad \bar{\Psi} \to \bar{\Psi}e^{i\beta\gamma_5}.$$

Indeed, the typical Lagrangian for charged massive Dirac fermions is given by

$$\begin{aligned}\mathscr{L} &= -\bar{\Psi}\gamma^\mu\partial_\mu\Psi - m\bar{\Psi}\Psi \\ &= i\bar{\psi}_{\dot{a}}(\bar{\sigma}^\mu)^{\dot{a}b}\partial_\mu\psi_b + i\xi^a(\sigma^\mu)_{a\dot{b}}\partial_\mu\bar{\xi}^{\dot{b}} - m[\xi^a\psi_a + \bar{\psi}_{\dot{a}}\bar{\xi}^{\dot{a}}].\end{aligned} \tag{10.29}$$

Blackboard video

Dirac equation

Variation with respect to $\bar{\Psi}$ yields the Dirac equation,

$$(\gamma^\mu\partial_\mu + m)\Psi = 0.$$

The equation of motion following from the Majorana Lagrangian (10.26) would actually be of the same form, but it would be for a constrained or "real" Majorana spinor and not for an unconstrained or "complex" Dirac spinor.

Blackboard video

Relation to Klein-Gordon equation

It is interesting to apply another derivative operator to the Dirac equation,

$$(-\gamma^\mu\partial_\mu + m)(\gamma^\nu\partial_\nu + m)\Psi = 0.$$

Here one can replace

$$\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu = \frac{1}{2}\{\gamma^\mu,\gamma^\nu\}\partial_\mu\partial_\nu = \eta^{\mu\nu}\partial_\mu\partial_\nu,$$

because the partial derivatives commute. This leads to

$$(-\eta^{\mu\nu}\partial_\mu\partial_\nu + m^2)\Psi = 0,$$

which shows that all components of the Dirac spionor that solves the free Dirac equation are also solutions to the Klein-Gordon equation. This implies also that one can solve the free Dirac equation in terms of plane waves.

Partition function for Dirac fermions

We can now also write down a partition function for free Dirac fermions in the form of a Grassmann functional integral,

$$Z_2[\bar{\eta},\eta] = \int D\bar{\Psi}D\Psi \exp\left(i\int d^4x\left\{-\bar{\Psi}\gamma^\mu\partial_\mu\Psi - m\bar{\Psi}\Psi\right\} + i\int d^4x\left\{\bar{\eta}\Psi + \bar{\Psi}\eta\right\}\right).$$

The fields Ψ and $\bar{\Psi}$, as well as the sources η and $\bar{\eta}$ are here Grassmann valued fields which also have the structure of four-component Dirac spinors. For example,

$$\bar{\eta}\Psi = \sum_{\alpha=1}^{4}\bar{\eta}_\alpha\Psi_\alpha.$$

As usual one can now determine correlation functions by taking functional derivatives, for example

$$\langle \Psi_\alpha(x)\bar{\Psi}_\beta(y)\rangle = \frac{1}{Z_2[\bar{\eta},\eta]}\left(\frac{1}{i}\frac{\delta}{\delta\bar{\eta}_\alpha(x)}\right)\left(i\frac{\delta}{\delta\eta_\beta(y)}\right)Z_2[\bar{\eta},\eta] = \frac{1}{i}S_{\alpha\beta}(x-y).$$

The signs take here the Grassmann properties into account.

Blackboard video

Feynmann propagator for Dirac fermions

As usual, it is possible to perform the Gaussian integral by completing the square,

$$Z_2[\bar{\eta},\eta] = \exp\left(i\int d^4x\,\{\bar{\eta}_\alpha(x)S_{\alpha\beta}(x-y)\eta_\beta(y)\}\right),$$

where the propagator or time-ordered Greens function $S_{\alpha\beta}(x-y)$ is defined such that

$$\left(\gamma^\mu\frac{\partial}{\partial x^\mu}+m\right)_{\alpha\kappa}S_{\kappa\beta}(x-y) = \delta_{\alpha\beta}\delta^{(4)}(x-y).$$

This opertor inversion can be done conveniently in Fourier space,

$$\begin{aligned} S_{\alpha\beta}(x-y) &= \int\frac{d^4p}{(2\pi)^4}\,e^{ip(x-y)}(i\not{p}+m)^{-1}_{\alpha\beta} \\ &= \int\frac{d^4p}{(2\pi)^4}\,e^{ip(x-y)}\frac{(-i\not{p}+m)_{\alpha\beta}}{p^2+m^2-i\epsilon}. \end{aligned}$$

We have used here that

$$(-i\not{p}+m)(i\not{p}+m) = \not{p}\not{p}+m^2 = \gamma^\mu\gamma^\nu p_\mu p_\nu + m^2 = \frac{1}{2}\{\gamma^\mu,\gamma^\nu\}p_\mu p_\nu + m^2 = p^2+m^2,$$

and have inserted the usual $i\epsilon$ term to ensure the right causality properties for a Feynmann propagator.

Blackboard video

Coupling to gauge fields

We can now also write down the Lagrangian for Dirac fermions coupled to the electromagnetic gauge field A_μ,

$$\mathscr{L} = -\bar{\Psi}\gamma^\mu(\partial_\mu - ieA_\mu)\Psi - m\bar{\Psi}\Psi.$$

This is invariant under the local U(1) gauge transformation

$$\Psi \to e^{i\alpha(x)}\Psi, \qquad \bar{\Psi}\to\bar{\Psi}e^{-i\alpha(x)}, \qquad A_\mu(x)\to A_\mu(x)+\frac{1}{e}\partial_\mu\alpha(x).$$

Blackboard video

Lagrangian for Quantum electrodynamics

We can now also write down a Lagrangian for quantum electrodynamics in complete form after adding a kinetic term for the gauge fields,

$$\mathscr{L} = -\bar{\Psi}\gamma^\mu(\partial_\mu - ieA_\mu)\Psi - m\bar{\Psi}\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}.$$

We use here the electromagnetic field strength tensor

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

It is obviously anti-symmetric and invariant under the U(1) gauge transformations introduced above. Accordingly entire Lagrangian for quantum electrodynamics (QED) is gauge invariant.

Blackboard video

LECTURE 21

11 Poincaré group, fields and particles

We have seen that quantum fields build representations of the Lorentz group. We have specifically investigated scalar and spinor fields, but will soon also turn to vector fields. On the other side, we have seen that excitations of fields in the asymptotic regimes, where they propagate over large distances, can be understood as particles. In this section we will investigate the relation further using spacetime symmetries.

Asymptotic states and the Poincaré group

Asymptotic states that describe freely propagating particles correspond fields correlated over large distances in spacetime and are independent of each other. Formally they can be associated to poles in propagators or correlation functions in momentum space. These asymptotic regions are governed by a set of symmetries, and it was shown by Eugene Wigner that one can use them to characterize the properties of particles.

Transformations of fields

So far we have discussed how the "internal" indices of a field transform under Lorentz transformations. However, a field depends on a space-time position x^μ which also transforms. This is already the case for a scalar field,

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

(A maximum at x^μ is moved to a maximum at $\Lambda^\mu{}_\nu x^\nu$.) In infinitesimal form

$$(\Lambda^{-1})^\mu{}_\nu = \delta^\mu{}_\nu - \delta\omega^\mu{}_\nu,$$

and thus

$$\phi(x) \to \phi'(x) = \phi(x) - x^\nu \delta\omega^\mu{}_\nu \partial_\mu \phi(x).$$

This can also be written as

$$\phi'(x) = \left(1 + \frac{i}{2}\delta\omega^{\mu\nu}\mathcal{M}_{\mu\nu}\right)\phi(x),$$

with generator

$$\mathcal{M}_{\mu\nu} = -i(x_\mu \partial_\nu - x_\nu \partial_\mu).$$

Indeed, these generators form a representation of the Lie algebra (10.2), i. e.

$$[\mathcal{M}_{\mu\nu}, \mathcal{M}_{\rho\sigma}] = i(\eta_{\mu\rho}\mathcal{M}_{\nu\sigma} - \eta_{\mu\sigma}\mathcal{M}_{\nu\rho} - \eta_{\nu\rho}\mathcal{M}_{\mu\sigma} + \eta_{\nu\sigma}\mathcal{M}_{\mu\rho}). \tag{11.1}$$

For fields with non-vanishing spin, the complete generator contains $\mathcal{M}_{\mu\nu}$ and the generator of "internal" transformations, for example for a left-handed spinor

$$\psi_a(x) \to \psi'_a(x) = \left(\delta_a{}^b + \frac{i}{2}\delta\omega^{\mu\nu}(M_{\mu\nu})_a{}^b\right)\psi_b(x),$$

with

$$(M_{\mu\nu})_a{}^b = (M^L_{\mu\nu})_a{}^b + \mathcal{M}_{\mu\nu}\,\delta_a{}^b.$$

This can now be extended to fields in arbitrary representations of the Lorentz group.

Pioncaré group

Poincaré transformations consist of Lorentz transformations plus translations,

$$x^\mu \to \Lambda^\mu{}_\nu x^\nu - b^\mu.$$

Translations only (without Lorentz transformations) form themselves an abelian Lie group, the additive group $\mathbb{R}^4$. It is clear that Poincaré transformations form a group. The composition law is

$$(\Lambda_2, b_2) \circ (\Lambda_1, b_1) = (\Lambda_2\Lambda_1, b_2 + \Lambda_2 b_1).$$

[Exercise: Show this.] The composition law is an example for a *semi-direct product*, namely of the Lorentz group $\mathrm{O}(1,3)$ and the additive group $\mathbb{R}^4$ of space and time translations,

$$\text{Poincaré group} \cong \mathrm{O}(1,3) \ltimes \mathbb{R}^4.$$

Lorentz transformations can be parametrized by six parameters, which are supplemented by four parameters for translations. The entire symmetry group of Minkowski space has therefore ten parameters.

Lie algebra of Poincaré group

Let us now find the Lie algebra associated with the Poincaré group. As transformations of fields, translations are generated by the momentum operator

$$P_\mu = -i\partial_\mu.$$

For example, as an infinitesimal transformation,

$$\begin{aligned}\phi(x) \to \phi'(x) &= \phi(\Lambda^{-1}(x+b)) \\ &= \phi(x^\mu - \delta\omega^\mu{}_\nu x^\nu + b^\mu) \\ &= \left(1 + \frac{i}{2}\delta\omega^{\mu\nu}\mathcal{M}_{\mu\nu} + ib^\mu P_\mu\right)\phi(x).\end{aligned}$$

One finds easily

$$[P_\mu, P_\nu] = 0, \tag{11.2}$$

and

$$[\mathcal{M}_{\mu\nu}, P_\rho] = i\left(\eta_{\mu\rho}P_\nu - \eta_{\nu\rho}P_\mu\right), \tag{11.3}$$

which together with (11.1) forms the Lie bracket relations of the Poincaré algebra. The commutator (11.2) tells that the different components of the energy-momentum operator can be diagonalized simultaneously, while (11.3) says that P_ρ transforms as a covector under Lorentz transformations.

Representations of the Poincaré group

Let us now discuss representations of the Poincaré algebra (and corresponding representations of the Poincaré group). We concentrate here on the part of the group that is connected to the identity transformations, i. e. $\mathrm{SO}^{\uparrow}(1,3) \ltimes \mathbb{R}^4$. It turns out that single-particle states can be understood as examples for such representations.

As we have done before, we will use a maximal number of commuting generators to label states. In particular, the different components of the momentum operator $P_\mu = -i\partial_\mu$ commute and we can work with corresponding eigenstates, namely plane waves $e^{ip_\mu x^\mu}$. The eigenvalues are then energy and momentum, $p_\mu = (-E, \vec{p})$.

Casimir operators

To classify representations, we first search for Casimir operators, i. e. operators that commute with all generators. One Casimir operator is

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

which obviously commutes with $\mathcal{M}_{\mu\nu}$ and P_μ. For single particle states of massive particles we have $p_\mu p^\mu + M^2 = 0$ so that $-P^2 = M^2$ gives the particle mass. The other Casimir operator follows from the *Pauli-Lubanski vector*

$$W^\mu = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\mathcal{M}_{\nu\rho}P_\sigma.$$

It is orthogonal to the momentum, $W^\mu P_\mu = 0$, and has the commutation relations

$$[W^\mu, P_\nu] = 0, \qquad [\mathcal{M}_{\rho\sigma}, W^\mu] = i\left(\delta^\mu_\rho W_\sigma - \delta^\mu_\sigma W_\rho\right),$$

as well as

$$[W^\mu, W^\nu] = -i\epsilon^{\mu\nu\rho\sigma}W_\rho P_\sigma.$$

The second Casimir of the Poincaré algebra is then given by

$$W^2 = W_\mu W^\mu.$$

The little group

When discussing representations of the Poincaré group it is convenient to first make a case separation in terms of the quadratic Casimir $P^2 = P_\mu P^\mu$. In each of the cases one can then fix a reference choice for p^μ_* and discuss remaining transformations that leave this reference invariant,

$$(\delta^\mu{}_\nu + \delta\omega^\mu{}_\nu)p^\nu_* = p^\mu_*. \tag{11.4}$$

This remaining symmetry group is then known as the *little group*. We will see this working in practise below.

Representations with vanishing momentum

The eigenvalue of the momentum operator P_μ may actually simply vanish, $p^\mu_* = (0,0,0,0)$. In that case the little group corresponds to the entire Lorentz group $\mathrm{SO}(1,3)$. An example for such a state is the vacuum.

Representations with positive mass squared

Let us now first consider situations with $-P^2 = M^2 > 0$. Examples for such representations are single particle states with positive mass.

We can fix a reference momentum $p^\mu_* = (M, 0, 0, 0)$ which corresponds to a particle momentum in its rest frame. The *little group* then consists of transformations that leave p^μ_* invariant. These are just rotations so the little group is here SO(3) or its double cover SU(2) which has the same Lie algebra. More explicitly, this follows from searching solutions to (11.4) which is here equivalent to $\delta\omega^{\mu 0} = 0$. Lorentz boosts are excluded; what is left are rotations.

In the particles rest frame, the Pauli-Lubanski vector evaluates to

$$W_0 = 0, \qquad W_j = \frac{M}{2}\epsilon_{jkl}\mathcal{M}_{kl} = MJ_j,$$

with angular momentum or spin operator J_j. The second Casimir of the Poincaré algebra is accordingly $W^2/M^2 = \vec{J}^2$. Single particle states $|p, j, m\rangle$ can be labeled by momentum p^μ, total spin $\vec{J}^2 = j(j+1)$ and eigenvalue m of the spin operator in z-direction J_3.

Symmetric spinor-tensor representation of SU(2)

Besides the standard representations of SU(2) discussed in the context of quantum mechanics, an alternativ representation is in terms of symmetric spinor-tensors. As we have seen, the spin-1/2 representation is a Pauli spionor ψ_s, where $s = 1, 2$, which transforms under rotations according to

$$\psi_s \to L_s{}^t\psi_t = \left(\delta_s{}^t + \frac{i}{4}\delta\omega_{ij}\epsilon_{ijk}(\sigma_k)_s{}^t\right)\psi_t, \tag{11.5}$$

where $L_s{}^t$ describes here the little group rotation matrix. One can now simply construct higher order representations as symmtric spinors with several SU(2) indices $\psi_{st\cdots u}$, and they transform accordingly under the little group transformations. In this way one obtains representations with higher spin, and a symmetric SU(2) tensor with n indices describes particles with spin $n/2$. The interesting feature about the little group is that every incoming or outgoing massive particle in a scattering experiment as its own little group and can be rotated independently in its respective rest frames.

Representations with negative mass squared

Here we have a situation with $-P^2 = M^2 < 0$. This corresponds to so-called tachyonic modes and if they appear they are usually associated to an instability.

We can fix a reference momentum as $p^\mu_* = (0, 0, 0, M)$. The *little group* consists now of transformations that leave p^μ_* invariant and these are Lorentz transformations in the remaining $1 + 2$ dimensional space, SO(1, 2). We will not discuss these representations in more detail.

Representations with vanishing mass

Let us now consider representations with $P^2 = 0$. This is again a rather interesting case. Examples are here single particle states with vanishing mass $M = 0$.

Massless particles do not have a restframe, so to discuss the little group one must pick another reference momentum, for example $p^\mu_* = (p, 0, 0, p) = p(\delta^\mu_0 + \delta^\mu_3)$. The little group consists of transformations that leave this invariant. Specifically, eq. (11.4) implies here $\delta\omega^{\mu 0} = \delta\omega^{\mu 3}$. One can write this as

$$\omega^{\mu\nu} = \begin{pmatrix} 0 & \alpha & \beta & 0 \\ -\alpha & 0 & \theta & -\alpha \\ -\beta & -\theta & 0 & -\beta \\ 0 & \alpha & \beta & 0 \end{pmatrix}.$$

Here, if only θ was non-vanishing, it would be the angle of a rotation in the 1-2-plane, i. e. around the propagation direction of the massless particle. Instead non-vanishing α would parametrize a combination of a boost in 1-direction together with a rotation in the 1-3 plane. Finally, β parametrizes a combination of a boost in the 2-direction with a rotation in 2-3-plane. An infinitesimal group transformation out of the litte group can be written as

$$\mathbb{1} + i\delta\theta J_3 + i\delta\alpha A + i\delta\beta B, \tag{11.6}$$

with

$$A = K_1 + J_2 = M_{10} + M_{31}, \qquad B = K_2 - J_1 = M_{20} + M_{32}.$$

The Lie algebra of the little group is

$$[J_3, A] = iB, \qquad [J_3, B] = -iA, \qquad [A, B] = 0. \tag{11.7}$$

This is in fact the Lie algebra of the so-called *special Euclidean group* $\mathrm{E}^+(2)$ consisting of translations and rotations in the two-dimensional Euclidean plane. It contains an $\mathrm{SO}(2)$ subgroup of rotations, as well as a subgroup of translations $\mathbb{R}^2$. The abelian subgroup of translations is in fact a *normal subgroup*. Similar to the Poincaré group itself, the Euclidean group $\mathrm{E}^+(2)$ has the structure of a direct product, $\mathrm{E}^+(2) = \mathrm{SO}(2) \ltimes \mathbb{R}^2$. [Exercise: Check all this!]

In the fundamental representation of the Lorentz algebra, the operators A and B are actually *nilpotent*. In fact, one has $A^3 = B^3 = AB = BA = 0$. However, there are also representations of (11.7) where A and B are hermitian such that the group has a unitary representation. However, as for any non-compact group, such unitary representations are necessarily infinite dimensional.

Physically, A and B can be related to gauge transformations. To see this consider polarization vectors for photons with momentum p^μ_*,

$$\epsilon^\mu_\pm = \frac{1}{\sqrt{2}}(0, 1, \pm i, 0).$$

These are eigenstates of J_3, namely in the fundamental or vector representation of the Lorentz group,

$$(J_3)^\mu{}_\nu \epsilon^\nu_\pm = \pm\epsilon^\mu_\pm.$$

The two polarizations $\epsilon^\mu_\pm$ describe therefore states with helicity ± 1, respectively. Now consider the action of (11.6) with $\delta\theta = 0$,

$$\epsilon^\mu_\pm \to \epsilon^\mu_\pm + \frac{(\delta\alpha \pm i\delta\beta)}{\sqrt{2}p} p^\mu_*.$$

Interestingly, the generators A and B are not realized trivially here, but they actually do change the polarization vector by a term proportional to the momentum. This is in fact a gauge transformation! To see this, consider the gauge transformation $A_\mu(x) \to A_\mu(x) + \partial_\mu\alpha(x)$, which becomes in momentum space

$$A_\mu(p) \to A_\mu(p) + ip_\mu\alpha(p).$$

The physical photon states are supposed to be independent of this gauge choice, and one takes them to be gauge equivalence classes. In other words, all states that differ by a gauge transformation are getting identified. This works similarly for massless particles of spin two, where the gauge symmetry is then the one of general relativity.

For such gauge equilvalence classes, or in a gauge fixed description, physical states of single massless particles can be characterized as having vanishing eigenvalues with respect to the operators A and B. We are then left with J_3 which generates rotations around the direction of propagation. This is in fact helicity, $J_3 = h$. The little group for massless particles is then $\mathrm{U}(1)$.

Fermionic massless particle states can change by a factor -1 under rotations of 2π around the propagation direction. This implies half-integer helicity h. In contrast, bosonic massless particle states should be invariant under 2π rotations, so that helicity h must be integer valued. These quantization conditions arise here from topological properties of the group, and not from properties of the Lie algebra.

Spinor helicity variables for massless momenta

Consider the $(2,2)$ representaion of a momentum p^μ,

$$p_\mu(\sigma^\mu)_{a\dot{a}} = \begin{pmatrix} -p^0+p^3 & p^1-ip^2 \\ p^1+ip^2 & -p^0-p^3 \end{pmatrix}.$$

The determinant of this matrix is $(p^0)^2 - \mathbf{p}^2 = -\eta_{\mu\nu}p^\mu p^\nu$. For momenta corresponding to massless particles this vanishes. This means that one eigenvalue is zero, or that the matrix has rank 1. As an example, take $p^\mu = p^\mu_*$, the reference momentum we have chosen above for massless particles. In that case,

$$p_\mu(\sigma^\mu)_{a\dot{a}} = \begin{pmatrix} 0 & 0 \\ 0 & -2p^0 \end{pmatrix}.$$

This shows that one can write for such momenta

$$p_\mu(\sigma^\mu)_{a\dot{a}} = -\lambda_a\tilde{\lambda}_{\dot{a}}. \tag{11.8}$$

The objects λ_a and $\tilde{\lambda}_{\dot{a}}$ are known as *spinor helicity variables* or *twistors*. They are formally left-handed and right-handed spinors, respectively, but are taken to be commuting, i. e. they are composed out of ordinary complex numbers and not Grassmann variables. For real momenta $p_\mu\sigma^\mu$ is hermiten and one has

$$\tilde{\lambda}_{\dot{a}} = \pm(\lambda_a)^*$$

where the positive sign must be choosen for $p^0 > 0$ and the negative sign for $p^0 < 0$. More generally one may also consider complex momenta and then λ_a and $\tilde{\lambda}_{\dot{a}}$ become independent. An explicit realization for $p^0 > 0$ and massless on-shell momenta is

$$\lambda_a = \frac{1}{\sqrt{p^0+p^3}}\begin{pmatrix} -p^1+ip^2 \\ p^0+p^3 \end{pmatrix}, \qquad \tilde{\lambda}_{\dot{a}} = \frac{1}{\sqrt{p^0+p^3}}\begin{pmatrix} -p^1-ip^2 \\ p^0+p^3 \end{pmatrix}.$$

This assignment is not unique, however. Specifically one could transform

$$\lambda_a \to e^{i\theta}\lambda_a, \qquad \tilde{\lambda}_{\dot{a}} \to e^{-i\theta}\tilde{\lambda}_{\dot{a}},$$

and obtain new twistors that also fulfill (11.8). One may use this to analytically continue for negative frequency $p^0 < 0$ such that for $p^\mu \to -p^\mu$

$$\lambda_a \to \lambda_a, \qquad \tilde{\lambda}_{\dot{a}} \to -\tilde{\lambda}_{\dot{a}}.$$

Pulling indices up and down

As we have seen previously one can sensibly pull indices up and down with ε_{ab} and its inverse ε^{ab}, for example a twistor with upper left-handed index is obtained by

$$\lambda^a = \varepsilon^{ab}\lambda_b.$$

Similarly this can be done for right-handed indices. One can then also write

$$p_\mu(\bar{\sigma}^\mu)^{\dot{a}a} = \varepsilon^{\dot{a}\dot{b}}\varepsilon^{ab}p_\mu(\sigma^\mu)_{b\dot{b}} = -\tilde{\lambda}^{\dot{a}}\lambda^a.$$

Bracket notation and scalar products

It is customary to introduce the following notation for the left handed twistor associated with the momentum p,

$$|p\rangle_a = \lambda_a, \qquad \langle p|^a = \lambda^a = \varepsilon^{ab}|p\rangle_a,$$

and similarly for the right handed twistor,

$$|p]^{\dot a} = \tilde\lambda^{\dot a}, \qquad [p|_{\dot a} = \tilde\lambda_{\dot a} = \varepsilon_{\dot a\dot b}|p]^{\dot b}.$$

The analytic continuation properties can be choosen such that

$$|-p\rangle_a = |p\rangle_a, \qquad |-p]^{\dot a} = -|p]^{\dot a}.$$

The bracket notation (not to be confused with the bra-ket notation of quantum mechanics) allows also to work nicely in situations where differnt momenta are involved, for example $|p\rangle_a = \lambda_a$ and $|q\rangle_a = \mu_a$. One defines the SU(2) invariant scalar product between left-handed twistors

$$\langle pq\rangle = \langle p|^a|q\rangle_a = \lambda^a\mu_a = \varepsilon^{ab}\lambda_b\mu_a = -\mu^a\lambda_b = -\langle p|^a|q\rangle_a = -\langle qp\rangle,$$

and between right-handed twistors,

$$[pq] = [p|_{\dot a}q]^{\dot a} = \tilde\lambda_{\dot a}\tilde\mu^{\dot a} = \varepsilon_{\dot a\dot b}\tilde\lambda^{\dot b}\tilde\mu^{\dot a} = -\tilde\mu_{\dot a}\tilde\lambda^{\dot a} = -[q|^{\dot a}|p]_{\dot a} = -[qp].$$

Note that it is not possible to sensibly define a scalar product between a left-handed and a right-handed twistor. The anti-symmetry implies also that for massless momenta

$$\langle pp\rangle = [pp] = 0.$$

This will be useful in the following. In the bracket notation we can also write

$$p_\mu(\sigma^\mu)_{a\dot b} = -|p\rangle_a[p|_{\dot b}, \qquad p_\mu(\bar\sigma^\mu)^{\dot a b} = -|p]^{\dot a}\langle p|^b.$$

This also implies

$$2p_\mu q^\mu = -\mathrm{tr}\{p_\mu\sigma^\mu q_\mu\bar\sigma^\nu\} = -|p\rangle_a[p|_{\dot b}|q]^{\dot b}\langle q|^a = -\langle qp\rangle[pq] = \langle pq\rangle[pq].$$

Weyl equations

The equation of motion $i(\bar\sigma^\mu)^{\dot a b}\partial_\mu\psi_b(x)$ for free massless left-handed Weyl spinor in plane wave form $\psi_b(x) = \psi_b(p)e^{ipx}$ reads now

$$-p_\mu(\bar\sigma^\mu)^{\dot a b}\psi_b(p) = |p]^{\dot a}\langle p|^b\psi_b(p) = 0.$$

One can see that $\psi_b(p) \sim |p\rangle_b$ is actually a solution,

$$|p]^{\dot a}\langle p|^b|p\rangle_b = |p]^{\dot a}\langle pp\rangle = 0.$$

(Some care is needed here, however, because $\psi_b(p)$ is a spinor with Grassmann property while $|p\rangle_b$ is a twistor with no Grassmann property.) The complex conjugate

$$\bar\psi^{\dot a}(p) = \varepsilon^{\dot a\dot b}(\psi_b(-p))^* = -|-p]^{\dot a} = |p]^{\dot a}$$

is similarly a solution of the right-handed Weyl equation,

$$-p_\mu(\sigma^\mu)_{a\dot b}|p]^{\dot b} = |p\rangle_a[p|_{\dot b}|p]^{\dot b} = .|p\rangle_a[pp] = 0.$$

In other words, we can identify $\psi_a(p) \sim |p\rangle_a$ and $\bar\psi^{\dot a}(p) \sim |p]^{\dot a}$ is the corresponding conjugate spinor (of course some care is needed with Grassmann properties).

A solution of the free Majorana equation at vanishing mass, $\gamma^\mu p_\mu\Psi(p) = 0$, can accordingly be written in the form

$$\Psi(p) = \begin{pmatrix}\psi_a(p)\\ \bar\psi^{\dot a}(p)\end{pmatrix} \sim \begin{pmatrix}|p\rangle_a\\ |p]^{\dot a}\end{pmatrix}.$$

Little group representation for massless twistors

To construct $|p\rangle_b = \lambda_b$ explicitely it is useful to start with a specific reference frame, for example the one where $p^\mu = p^\mu_* = (p, 0, 0, p)$ points in the three-direction. One has then

$$|p\rangle_a = \lambda_a = \begin{pmatrix} 0 \\ \sqrt{2p} \end{pmatrix}.$$

One can then apply a Lorentz transformation in the left-handed representation $L_a{}^b$ to go to other momenta.

At this point it is also interesting how the little group acts here. For the left-handed representation one finds

$$\mathbb{1} + i\delta\theta J_3 + i\delta\alpha A + i\delta\beta B = \begin{pmatrix} 1 + \frac{i}{2}\delta\theta & 0 \\ -\delta\alpha - i\delta\beta & 1 - \frac{i}{2}\delta\theta \end{pmatrix}.$$

Interestingly this implies that under such a transformation

$$|p\rangle_a \to e^{-i\theta/2}|p\rangle_a.$$

For the left handed twistor the generators A and B are represented trivially and have no effect. The little group is here just a complex phase $e^{-i\theta/2}$. The negative sign and the factor $1/2$ tell that we are dealing with states of helicity $h = -1/2$ as appropriate for left-handed Weyl fermions. For right-handed twistors oe has instead

$$|p]^{\dot a} \to e^{i\theta/2}|p]^{\dot a},$$

which tells that they have helicity $h = 1/2$. For practical caluclations it is useful to know that the little group transformations can be done independently for each incoming or outgoing particle and a scattering amplitude must transform accordingly with the respective phase factors.

Extension to nonzero mass

Consider now momenta p^μ such that $\eta_{\mu\nu}p^\mu p^\nu + m^2 = 0$ with nonzero mass m. In that case one can introduce two twistors λ_{sa} with $s = 1, 2$ and write

$$p^\mu(\sigma_\mu)_{a\dot b} = -\lambda_{sa}\tilde\lambda^s_{\dot b} = -\lambda_{sa}\varepsilon^{st}\tilde\lambda_{t\dot b}.$$

By taking the determinant of this matrix one obtains $-(p^0)^2 + \mathbf{p}^2 = -m^2$ and accordingly one may, as a matrix with indices a and s, take

$$\det(\lambda_{sa}) = \det(\tilde\lambda_{t\dot b}) = m.$$

The index s is here a kind of internal SU(2) index associated with a particle of momentum p^μ where the group SU(2) corresponds to the appropriate little group of roations for massive particles with spin 1/2. Accordingly, these indices can also be pulled up and down with ε_{st} and its inverse ε^{st}, for example

$$\tilde\lambda^s_{\dot a} = \varepsilon^{st}\lambda_{t\dot a}.$$

Bracket notation for massive twistors

Also for the massive twistors one can work with the bracket notation

$$|p_s\rangle_a = \lambda_{sa}, \qquad \langle p_s|^a = \lambda^a_s = \varepsilon^{ab}|p_s\rangle_b.$$

We can follow conventions where the little group SU(2) index is taken in an upper position. Similarly for right-handed massive twistors (where $\tilde\lambda_{s\dot a} = (\lambda_{sa})^*$ in Minkowski space for positive energy)

$$|p_s]^{\dot a} = \tilde\lambda^{\dot a}_s, \qquad [p_s|_{\dot a} = \tilde\lambda_{s\dot a} = \varepsilon_{\dot a\dot b}|p_s]^{\dot b}.$$

Sometimes one uses also a notation where the little group SU(2) or spin indices are suppressed but then uses bold face letters to indicate that these are massive twistors, e. g.

$$|\boldsymbol{p}\rangle_a = |p_s\rangle_a.$$

Little group transformations for massive twistors

The little group of SU(2) rotations acts on these twistors according to

$$|p_s\rangle_a \to L_s{}^t |p_t\rangle_a,$$

which is precisely the transformation in (11.5). This is actually the same transformation for any twistor with an upper little group index.

Dirac equation for Majorana spionors

The Majorana spinor in momentum space can be written as

$$\Psi_s(p) = \begin{pmatrix} \psi_{sa}(p) \\ \bar{\psi}_s^{\dot{a}}(p) \end{pmatrix} \sim \begin{pmatrix} |p_s\rangle_a \\ |p_s]^{\dot{a}} \end{pmatrix}.$$

The index s labels now the little group or, in other words, parametrizes the two independent spin states in the particles rest frame. The Dirac equation $(i\gamma^\mu p_\mu + m)\Psi_s(p) = 0$ reads

$$\begin{pmatrix} m\,\delta_a{}^b & p_\mu(\sigma^\mu)_{a\dot{b}} \\ p_\mu(\bar{\sigma}^\mu)^{\dot{a}b} & m\,\delta^{\dot{a}}{}_{\dot{b}} \end{pmatrix} \begin{pmatrix} |p_s\rangle_b \\ |p_s]^{\dot{b}} \end{pmatrix} = \begin{pmatrix} m\,\delta_a{}^b & -|p_t\rangle_a[p^t|_{\dot{b}} \\ -|p_t]^{\dot{a}}\langle p^t|^b & m\,\delta^{\dot{a}}{}_{\dot{b}} \end{pmatrix} \begin{pmatrix} |p_s\rangle_b \\ |p_s]^{\dot{b}} \end{pmatrix} = 0.$$

One infers from this that

$$m|p_s\rangle_a + p_\mu(\sigma^\mu)_{a\dot{b}}|p_s]^{\dot{b}} = m|p_s\rangle_a - |p_t\rangle_a[p^t p_s] = 0,$$

and similarly

$$m|p_s]^{\dot{a}} + p_\mu(\bar{\sigma}^\mu)^{\dot{a}b}|p_s\rangle_b = m|p_s]^{\dot{a}} - |p_s]^{\dot{a}}\langle p^t p_s\rangle = 0.$$

These are just other ways to write the free Dirac equation for free massive Majorana particles in Minkowski space. Here we are using the scalar products $\langle p_s q_t\rangle = -\langle q_t p_s\rangle$ and $[p_s q_t] = -[q_t p_s]$, which generalize the corresponding definitions for massless twistors. Because of the additional index, such scalar products with the same momentum $\langle p_s p_t\rangle$ and $[p_s p_t]$ are now non-zero, albeit they are antisymmetric and proportional to ε_{st}.

We leave the generalization of the formalism from Majorana to Dirac spinors for the future.

Mode expansion for Dirac fields

We will also need a mode expansion for free Dirac fields in order to describe asymptotic (incoming and outgoing) fermion states. We write the fields as

$$\Psi(x) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left\{ b^s_{\mathbf{p}}\, u_s(p)\, e^{ipx} + d^{s\dagger}_{\mathbf{p}}\, v_s(p)\, e^{-ipx} \right\},$$

$$\bar{\Psi}(x) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left\{ d^s_{\mathbf{p}}\, \bar{v}_s(p)\, e^{ipx} + b^{s\dagger}_{\mathbf{p}}\, \bar{u}_s(p)\, e^{-ipx} \right\}.$$

Here, $b^s_{\mathbf{p}}$, $d^s_{\mathbf{p}}$ etc. can be seen as expansion coefficients and become operators in the operator picture. Because of the fermionic or Grassmann exchange symmetry, they fulfill actually anti-commutation relations,

$$\{b^r_{\mathbf{p}}, b^{s\dagger}_{\mathbf{q}}\} = \{d^r_{\mathbf{p}}, d^{s\dagger}_{\mathbf{q}}\} = (2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p} - \mathbf{q}),$$

with all other anti-commutators vanishing. The index s sums over independent spin states. One can interpret $b^{s\dagger}_{\mathbf{p}}$ as a creation opertors for fermionoc particles (such as electrons) with momentum $\mathbf{p}$ and spin s, and similarly $d^{s\dagger}_{\mathbf{p}}$ as the corresponding creation operator for anti-particles (such as positrons). The fact that there are particles and anti-particles arises from the "complex" nature of Dirac spionors; for Majorana fermions there would be only one kind of creation operator.

Blackboard video

Solutions of Dirac equation

The Dirac equation

$$(\gamma^\mu \partial_\mu + m)\psi(x) = 0,$$

becomes for the plane waves

$$(i\not{p} + m)\ u_s(\mathbf{p}) = 0,$$
$$(-i\not{p} + m)\ v_s(\mathbf{p}) = 0,$$

with the Dirac slash notation $\not{p} = \gamma^\mu p_\mu$. We consider this first in the frame where the spatial momentum vanishes, $\mathbf{p} = 0$, such that $p^\mu = (m, 0, 0, 0)$,

$$\not{p} = -\gamma^0 m = im \begin{pmatrix} & \mathbb{1} \\ \mathbb{1} & \end{pmatrix}.$$

The last equation holds in the chiral basis where

$$\gamma^\mu = \begin{pmatrix} 0 & -i\sigma^\mu \\ -i\bar{\sigma}^\mu & 0 \end{pmatrix}.$$

with $\sigma^\mu = (\mathbb{1}, \vec{\sigma})$ and $\bar{\sigma}^\mu = (\mathbb{1}, -\vec{\sigma})$. For the spinor $u_s(\mathbf{0})$ one has the equation

$$(i\not{p} + m)u_s(\mathbf{0}) = m \begin{pmatrix} +\mathbb{1} & -\mathbb{1} \\ -\mathbb{1} & +\mathbb{1} \end{pmatrix} u_s(\mathbf{0}) = 0.$$

The two independent solutions are

$$u_1(\mathbf{0}) = \sqrt{m} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad u_2(\mathbf{0}) = \sqrt{m} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

The normalization has been chosen for later convenience. Similarly

$$(-i\not{p} + m)v_s(\mathbf{0}) = m \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix} v_s(\mathbf{0}) = 0,$$

has the two independent solutions

$$v_1(\mathbf{0}) = \sqrt{m} \begin{pmatrix} 0 \\ +1 \\ 0 \\ -1 \end{pmatrix}, \qquad v_2(\mathbf{0}) = \sqrt{m} \begin{pmatrix} -1 \\ 0 \\ +1 \\ 0 \end{pmatrix}.$$

We see here that the Dirac equation has two independent solutions (for spin up and 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 to obtain $u_s(\mathbf{p})$, $v_s(\mathbf{p})$ and their conjugates $= u_s^\dagger(\mathbf{p})\beta$ as well as $\bar{v}_s(\mathbf{p}) = v_s^\dagger(\mathbf{p})\beta$. We derive more identities about these objects when we need them later on.

Blackboard video

LECTURE 22

11.1 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

$$\langle \mathbf{p}|\mathbf{q}\rangle = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q}).$$

For relativistic physics this has the drawback that it is not Lorentz invariant. To see this let us consider a boost in z-direction

$$\begin{aligned} E' =& \gamma(E+\beta p^3),\\ p^{1\prime} =& p^1,\\ p^{2\prime} =& p^2,\\ p^{3\prime} =& \gamma(p^3+\beta E). \end{aligned}$$

Using the identity

$$\delta\left(f(x)-f(x_0)\right) = \frac{1}{|f'(x_0)|}\delta(x-x_0),$$

one finds

$$\begin{aligned} \delta^{(3)}(\mathbf{p}-\mathbf{q}) &= \delta^{(3)}(\mathbf{p}'-\mathbf{q}')\frac{dp^{3\prime}}{dp^3} = \delta^{(3)}(\mathbf{p}-\mathbf{q})\gamma\left(1+\beta\frac{dE}{dp^3}\right)\\ &= \delta^{(3)}(\mathbf{p}'-\mathbf{q}')\frac{1}{E}\gamma\left(E+\beta p^3\right)\\ &= \frac{E'}{E}\delta^{(3)}(\mathbf{p}'-\mathbf{q}'). \end{aligned}$$

This shows, however, that $E\,\delta^{(3)}(\mathbf{p}-\mathbf{q})$ is in fact Lorentz invariant.

This motivates to change the normalization such that

$$|p;\text{in}\rangle = \sqrt{2E_p}a_{\mathbf{p}}^\dagger(-\infty)|0\rangle = \sqrt{2E_{\mathbf{p}}}\,|\mathbf{p};\text{in}\rangle.$$

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

$$\langle p;\text{in}|q;\text{in}\rangle = 2E_p(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q}).$$

With this normalization we must divide by $2E_p$ at the same places. In particular the completeness relation for single particle incoming states is

$$\mathbb{1}_{1-\text{particle}} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} |p;\text{in}\rangle\langle p;\text{in}|.$$

In fact, what appears here is a Lorentz invariant momentum measure. To see this consider

$$\int \frac{d^4p}{(2\pi)^4} \, (2\pi) \, \delta(p^2 + m^2) \, \theta(p^0) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}}.$$

The left hand side is explicitly Lorentz invariant and so is the right hand side.

Blackboard video

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

$$S_{\beta\alpha} = \langle \beta;\text{out}|\alpha;\text{in}\rangle = \delta_{\beta\alpha} + i\,\mathcal{T}_{\beta\alpha}(2\pi)^4\delta^{(4)}(p^{\text{in}} - p^{\text{out}}).$$

But now we take elements with relativistic normalization, e.g. for $2 \to 2$ scattering

$$S_{q_1q_2,p_1p_2} = \langle q_1, q_2;\text{out}|p_1, p_2;\text{in}\rangle.$$

We can calculate these matrix elements as before using the LSZ reduction formula to replace $\sqrt{2E_p}a^\dagger_{\mathbf{p}}(-\infty)$ by fields. For example, for relativistic scalar fields

$$\sqrt{2E_{\mathbf{p}}}\; a^\dagger_{\mathbf{p}}(-\infty) = \sqrt{2E_{\mathbf{p}}}\; a^\dagger_{\vec{p}}(\infty) + i\left[-(p^0)^2 + \mathbf{p}^2 + m^2\right]\phi^*(p).$$

This allows to calculate S-matrix elements through correlation functions.

Blackboard video

Cross sections for $2 \to n$ scattering

Let us now generalize our discussion of $2 \to 2$ scattering of non-relativistic particles to a scattering $2 \to n$ of relativistic particles. The transition probability is as before

$$P = \frac{|\langle \beta;\text{out}|\alpha;\text{in}\rangle|^2}{\langle \beta;\text{out}|\beta;\text{out}\rangle\langle \alpha;\text{in}|\alpha;\text{in}\rangle}.$$

Rewriting the numerator in terms of $\mathcal{T}_{\beta\alpha}$ and going over to the transition rate we obtain as before

$$\dot{P} = \frac{V(2\pi)^4\delta^{(4)}(p^{\text{out}} - p^{\text{in}})|\mathcal{T}|^2}{\langle \beta;\text{out}|\beta;\text{out}\rangle\langle \alpha;\text{in}|\alpha;\text{in}\rangle}. \tag{11.9}$$

But now states are normalized in a covariant way

$$\begin{aligned}
\langle p|p\rangle &= \lim_{q\to p}\langle p|q\rangle \\
&= \lim_{q\to p} 2E_p(2\pi)^3\delta^{(3)}(\mathbf{p} - \mathbf{q}) \\
&= 2E_p(2\pi)^3\delta^{(3)}(0) \\
&= 2E_pV
\end{aligned}$$

One has thus for the incoming state of two particles

$$\langle \alpha; \text{in} | \alpha; \text{in} \rangle = 4E_1 E_2 V^2.$$

For the outgoing state of n particles one has instead

$$\langle \beta; \text{out} | \beta; \text{out} \rangle = \prod_{j=1}^{n} \{2q_j^0 V\}.$$

The product goes over final state particles which have the four-momentum q_j^n. So, far we have thus

$$\dot{P} = \frac{V\ (2\pi)^4\ \delta^{(4)}(p^{\text{in}} - p^{\text{out}}) |\mathcal{T}|^2}{4E_1 E_2 V^2\ \prod_{j=1}^{n} \{2q_j^0 V\}}.$$

Blackboard video

Lorentz invariant phase space

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

$$\sum_{\vec{n}_j} \to V \int \frac{d^3 q}{(2\pi)^3}$$

for each final state particle. The transition rate becomes

$$\dot{P} = \frac{|\mathcal{T}|^2}{4E_1 E_2 V} \left[(2\pi)^4\ \delta^{(4)}\Big(p^{\text{in}} - \sum_j q_j\Big) \prod_{j=1}^{n} \left\{ \frac{d^3 q_j}{(2\pi)^3 2q_j^0} \right\} \right]$$

The expression in square brackets is known as the *Lorentz-invariant phase space* measure (sometimes "LIPS").

Flux and differential cross section

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

$$\mathcal{F} = \frac{|v|}{V} = \frac{|v_1 - v_2|}{V} = \frac{\left| \frac{p_1^3}{p_1^0} - \frac{p_2^3}{p_2^0} \right|}{V}.$$

In the last equality we chose the beam axis to coincide with the z-axis. For the differential cross section we obtain

$$d\sigma = \frac{|\mathcal{T}|^2}{4E_1 E_2 |v_1 - v_2|}\ [\text{LIPS}].$$

The expression in the prefactor can be rewritten like

$$\frac{1}{E_1 E_2 |v_1 - v_2|} = \frac{1}{p_1^0 p_2^0 \left| \frac{p_1^3}{p_1^0} - \frac{p_2^3}{p_2^0} \right|} = \frac{1}{|p_2^0 p_1^3 - p_1^0 p_2^3|} = \frac{1}{|\epsilon_{\mu x y \nu} p_2^\mu p_1^\nu|}.$$

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.

Blackboard video

Differential cross section in the centre of mass frame

In the center of mass frame one has $p_2^3 = -p_1^3 = \pm|\mathbf{p}_1|$ and

$$\frac{1}{|p_2^0 p_1^3 - p_1^0 p_2^0|} = \frac{1}{|\mathbf{p}_1|(p_1^0 + p_2^0)} = \frac{1}{|\mathbf{p}_1|_{\mathrm{COM}}\sqrt{s}}$$

This leads finally to the result for the differential cross section

$$d\sigma = \frac{|\mathcal{T}|^2}{4|\mathbf{p}_1|_{\mathrm{COM}}\sqrt{s}} \left[(2\pi)^4 \delta^{(4)}\Big(p^{\mathrm{in}} - \sum_j q_j\Big) \prod_{j=1}^{n} \left\{ \frac{d^3 q_j}{(2\pi)^3 2q_j^0} \right\} \right].$$

Blackboard video

$2 \to 2$ scattering

For the case of $n = 2$ one can write the Lorentz invariant differential phase space element in the center of mass frame (exercise)

$$\left[(2\pi)^4 \, \delta^{(4)}(p^{\mathrm{in}} - q_1 - q_2) \frac{d^3 q_1}{(2\pi)^3 2q_1^0} \, \frac{d^3 q_2}{(2\pi)^3 q_2^0} \right] = \frac{|\mathbf{q}_1|}{16\pi^2 \sqrt{s}} d\Omega$$

such that

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s} \frac{|\mathbf{q}_1|}{|\mathbf{p}_1|} |\mathcal{T}|^2.$$

Decay rate

Let us now consider the decay rate of a single particle, i. e. a process $1 \to n$. We can still use equation (11.9), but now the initial state is normalized like

$$\langle \alpha; \mathrm{in} | \alpha; \mathrm{in} \rangle = 2E_1 V.$$

We find then for the differential transition or decay rate $d\Gamma = \dot{P}$

$$d\Gamma = \frac{|\mathcal{T}|^2}{2E_1} \left[(2\pi)^4 \delta^{(4)}\Big(p^{\mathrm{in}} - \sum_j q_j\Big) \prod_{j=1}^{n} \left\{ \frac{d^3 q_j}{(2\pi)^3 2q_j^0} \right\} \right]$$

In the center of mass frame one has $E_1 = m_1$. For the special case of $1 \to 2$ decay one finds in the center of mass frame or rest frame of the initial particle

$$d\Gamma = \frac{|\mathcal{T}|^2 |\mathbf{q}_1|}{32\pi^2 m_1^2} d\Omega.$$

Blackboard video

LECTURE 23

12 Quantum electrodynamics

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

$$\begin{aligned} Z_2[J] &= \int DA \ \exp\left[iS_2[A] + i\int J^\mu A_\mu\right] \\ &= \int DA \ \exp\left[i\int d^4x \ \left\{-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + J^\mu A_\mu\right\}\right]. \end{aligned}$$

One can go to momentum space as usual

$$A_\mu(x) = \int \frac{d^4p}{(2\pi)^4} \ e^{ipx} A_\mu(p),$$

and finds for the term in the exponential

$$\begin{aligned} &\int_x \left\{-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + J^\mu A_\mu\right\} \\ &= \frac{1}{2}\int \frac{d^4p}{(2\pi)^4} \left\{-A_\mu(-p)\left(p^2\eta^{\mu\nu} - p^\mu p^\nu\right) \ A_\nu(p) + J^\mu(-p)A_\mu(p) + A_\mu(-p)J^\mu(p)\right\}. \end{aligned}$$

Blackboard video

Attempt to invert the inverse propagator and gauge fixing

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

$$p^2\eta^{\mu\nu} - p^\mu p^\nu = p^2 \mathscr{P}^{\mu\nu}(p),$$

is not invertible. We wrote it here in terms of

$$\mathscr{P}_\mu{}^\nu(p) = \delta_\mu{}^\nu - \frac{p_\mu p^\nu}{p^2},$$

which is in fact a projector to the space orthogonal to p_ν

$$\mathscr{P}_\mu{}^\nu(p)\mathscr{P}_\nu{}^\rho(p) = \mathscr{P}_\mu{}^\rho(p).$$

As a projector matrix it has eigenvalues 0 and 1, only. However,

$$\mathscr{P}_\mu{}^\nu(p) \ p_\nu = 0.$$

The field $A_\nu(p)$ can be decomposed into two parts,

$$A_\nu(p) = \frac{i}{e}p_\nu\beta(p) + \hat{A}_\nu(p),$$

with

$$\hat{A}_\nu(p) = \mathscr{P}_\nu{}^\rho(p)A_\rho(p),$$

such that $p^\nu \hat{A}_\nu(p) = 0$. Moreover

$$\beta(p) = \frac{e}{ip^2} p^\nu A_\nu(p).$$

When acting on $\hat{A}_\nu(p)$, the projector $\mathscr{P}_\mu{}^\nu(p)$ is simply the unit matrix.

Recall that gauge transformations shift the field according to

$$A_\mu(x) \to A_\mu(x) + \frac{1}{e}\partial_\mu \alpha(x),$$

or in momentum space

$$A_\mu(p) \to A_\mu(p) + \frac{i}{e} p_\mu \alpha(p).$$

One can therefore always perform a gauge transformation such that $\beta(p) = 0$ or

$$\partial^\mu A_\mu(x) = 0.$$

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.

Blackboard video

Quadratic partition function

Now we can easily perform the Gaussian integral,

$$\begin{aligned} Z_2[J] = \int DA\ \exp\left[\frac{i}{2}\int_p \left\{ -\left(A_\mu(-p) - J_\rho(-p)\frac{\mathscr{P}^\rho{}_\mu}{p^2}\right) p^2 \mathscr{P}^{\mu\nu} \left(A_\nu(p) - \frac{\mathscr{P}_\nu{}^\sigma}{p^2} J_\sigma(p)\right)\right\}\right] \\ \times \exp\left[\frac{i}{2}\int_p J^\mu(-p)\frac{\mathscr{P}_{\mu\nu}(p)}{p^2} J^\nu(p)\right] \\ = \text{const} \times \exp\left[\frac{i}{2}\int_{x,y} J^\mu(x)\Delta_{\mu\nu}(x-y)J^\nu(y)\right]. \end{aligned}$$

In the last line we used the photon propagator in position space (in Landau gauge)

$$\Delta_{\mu\nu}(x-y) = \int \frac{d^4p}{(2\pi)^4}\, e^{ip(x-y)} \frac{\mathscr{P}_{\mu\nu}(p)}{p^2 - i\epsilon}.$$

In the last step we have inserted the $i\epsilon$ term as usual.

Blackboard video

Photon propagator in position space

In the free theory one has

$$\langle A_\mu(x) A_\nu(x)\rangle = \frac{1}{i^2}\left(\frac{1}{Z[J]}\frac{\delta^2}{\delta J^\mu(x)\delta J^\nu(y)} Z[J]\right)_{J=0} = \frac{1}{i}\Delta_{\mu\nu}(x-y).$$

We use the following graphical notation

$$(x,\mu) \sim\sim\sim\sim\sim\sim\sim\sim (y,\nu) = \frac{1}{i}\Delta_{\mu\nu}(x-y),$$

or with sources $iJ^\mu(x)$ at the end points

$$\bullet\sim\sim\sim\sim\sim\sim\sim\sim\bullet = \frac{1}{2}\int_{x,y} iJ^\mu(x)\,\frac{1}{i}\Delta_{\mu\nu}(x-y)\,iJ^\nu(y).$$

Free solutions

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

$$p^2 \mathscr{P}_\mu{}^\nu(p) \hat{A}_\nu(p) = p^2 \hat{A}_\mu(p) = 0.$$

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.

Polarizations

Quite generally, a four-vector can be written as

$$\hat{A}_\nu(p) = \left(b, \frac{a_1 + a_2}{\sqrt{2}}, \frac{-ia_1 + ia_2}{\sqrt{2}}, c \right).$$

From the Landau gauge condition $p^\nu \hat{A}_\nu = 0$ it follows that $b = -c$, so that one can write

$$\hat{A}_\nu(p) = \tilde{c} \times (-E, 0, 0, E) + a_1 \epsilon_\nu^{(1)} + a_2 \epsilon_\nu^{(2)},$$

with

$$\epsilon_\nu^{(1)} = \left(0, \frac{1}{\sqrt{2}}, \frac{-i}{\sqrt{2}}, 0 \right), \qquad \epsilon_\nu^{(2)} = \left(0, \frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0 \right).$$

However, the term $\sim \tilde{c}$ is in fact proportional to $p_\nu = (-E, 0, 0, E)$. We can do another gauge transformation such that $\tilde{c} = 0$. 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.

Blackboard video

Mode expansion

In summary, we can expand free solutions of the photon field like

$$A_\mu(x) = \sum_{\lambda=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left\{ a_{\mathbf{p},\lambda}\, \epsilon_\mu^{(\lambda)}(p)\, e^{ipx} + a_{\mathbf{p},\lambda}^\dagger\, \epsilon_\mu^{(\lambda)*}(p)\, e^{-ipx} \right\},$$

where $E_p = |\mathbf{p}|$ is the energy of a photon. The index λ labels the two polarization states.

In the current setup, $a_{\mathbf{p},\lambda}$ and $a_{\mathbf{p},\lambda}^\dagger$ are simply expansion coefficients, while they become annihilation and creation operators in the operator picture. The non-trivial commutation relation becomes then

$$\left[a_{\mathbf{p},\lambda}, a_{\mathbf{p}',\lambda'}^\dagger \right] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}') \delta_{\lambda\lambda'}.$$

Blackboard video

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

$$a_{\mathbf{q}}(\infty) \to i\left[-q^0 + \tfrac{\mathbf{q}^2}{2m} + V_0\right]\varphi(q),$$
$$a^\dagger_{\mathbf{q}}(-\infty) \to i\left[-q^0 + \tfrac{\mathbf{q}^2}{2m} + V_0\right]\varphi^*(q).$$

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

$$\sqrt{2E_p}\; a_{\mathbf{p},\lambda}(\infty) \to i\epsilon^{\nu *}_{(\lambda)}(p)\int d^4x\; e^{-ipx}[-\partial_\mu\partial^\mu]A_\nu(x)$$
$$\sqrt{2E_p}\; a^\dagger_{\mathbf{p},\lambda}(-\infty) \to i\epsilon^{\nu}_{(\lambda)}(p)\int d^4x\; e^{ipx}[-\partial_\mu\partial^\mu]A_\nu(x).$$

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.

LSZ reduction for Dirac fermions

Finally, let us give the LSZ reduction formulas for Dirac fermions (again neglecting renormalization effects)

$$\sqrt{2E_p}b_{\mathbf{p},s}(\infty) \to i\int d^4x\, e^{-ipx}\bar{u}_s(\mathbf{p})(\gamma^\mu\partial_\mu + m)\psi(x),$$
$$\sqrt{2E_p}d^\dagger_{\mathbf{p},s}(-\infty) \to -i\int d^4x\, e^{ipx}\bar{v}_s(\mathbf{p})(\gamma^\mu\partial_\mu + m)\psi(x),$$
$$\sqrt{2E_p}d_{\mathbf{p},s}(\infty) \to -i\int d^4x\, \bar{\psi}_s(x)(-\gamma^\mu\overleftarrow{\partial}_\mu + m)v_s(\mathbf{p})\, e^{-ipx},$$
$$\sqrt{2E_p}b^\dagger_{\mathbf{p},s}(-\infty) \to i\int d^4x\, y\bar{\psi}_s(x)(-\gamma^\mu\overleftarrow{\partial}_\mu + m)u_s(\mathbf{p})\, e^{ipx}.$$

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.

The factors of i will typically cancel out in claculations of S-matrix elements with similar factors that come with propagators. For anti-fermion lines there is an additional factor (-1) in the propagator because the momentum is there counted opposite to the fermion flow direction.

Blackboard video

Action and partition function

We are now ready to formulate the Feynman rules for a perturbative treatment of quantum electrodynamics. The microscopic action is

$$S = \int d^4x\left\{-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \bar{\psi}\gamma^\mu(\partial_\mu - ieA_\mu)\psi - m\bar{\psi}\psi\right\}$$
$$= S_2[\bar{\psi},\psi,A] + ie\int d^4x\; \bar{\psi}\gamma^\mu A_\mu\psi.$$

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

$$Z[\bar{\eta},\eta,J] = \int D\bar{\psi}D\psi DA\ \exp\left[iS[\bar{\psi},\psi,A] + i\int\left\{\bar{\eta}\psi + \bar{\psi}\eta + J^\mu A_\mu\right\}\right]$$

with $\bar{\eta}\psi = \bar{\eta}_\alpha\psi_\alpha$ where $\alpha = 1,\ldots,4$ sums over spinor components. Formally, one can write

$$Z[\bar{\eta},\eta,J] = \exp\left[-e\int d^4x \left(\frac{1}{i}\frac{\delta}{\delta J^\mu(x)}\right)\left(i\frac{\delta}{\delta\eta_\alpha(x)}\right)(\gamma^\mu)_{\alpha\beta}\left(\frac{1}{i}\frac{\delta}{\delta\bar{\eta}_\beta(x)}\right)\right] Z_2[\bar{\eta},\eta,J],$$

with quadratic partition function

$$\begin{aligned} Z_2 &= \int D\bar{\psi}D\psi DA\ \exp\left[iS_2[\bar{\psi},\psi,A] + i\int\left\{\bar{\eta}\psi + \bar{\psi}\eta + J^\mu A_\mu\right\}\right] \\ &= \exp\left[i\int d^4x d^4y\ \bar{\eta}(x)S(x-y)\eta(y)\right] \times \exp\left[\frac{i}{2}\int d^4x d^4y\ J^\mu(x)\Delta_{\mu\nu}(x-y)J^\nu(y)\right]. \end{aligned}$$

Propagator for Dirac fermions

We have used here also the propagator for Dirac fermions $S_{\alpha\beta}(x-y)$ introduced previously. 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.

The correlation function of two Dirac fields can also be expressed in terms of the Dirac propagator,

$$\langle\psi_\alpha(x)\bar{\psi}_\beta(y)\rangle = \frac{1}{Z_2}\left(\frac{1}{i}\frac{\delta}{\delta\bar{\eta}_\alpha(x)}\right)\left(i\frac{\delta}{\delta\eta_\beta(y)}\right) Z_2\Big|_{\bar{\eta}=\eta=J=0} = \frac{1}{i}S_{\alpha\beta}(x-y).$$

We introduce a graphical representation for thus, as well,

$$(x,\alpha)\longleftarrow(y,\beta) = \frac{1}{i}S_{\alpha\beta}(x-y).$$

With sources $i\bar{\eta}_\alpha(x)$ and $i\eta_\beta(y)$ at the end this would be

$$\bullet\longleftarrow\bullet = \int_{x,y} i\bar{\eta}_\alpha(x)\ \frac{1}{i}S_{\alpha\beta}(x-y)\ i\eta_\beta(y) = i\int_{x,y}\bar{\eta}(x)S(x-y)\eta(y).$$

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.

LECTURE 24

Expanding out exponentials

We now consider the full partition function and expand out the exponentials,

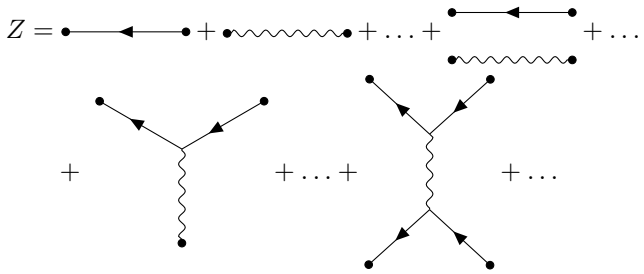
$$\begin{aligned} Z[\bar{\eta},\eta,J] = &\sum_{V=0}^\infty \frac{1}{V!}\left[\int_x\left(\frac{1}{i}\frac{\delta}{\delta J^\mu(x)}\right)\left(i\frac{\delta}{\delta\eta_\alpha(x)}\right)\left(-e\gamma^\mu_{\alpha\beta}\right)\left(\frac{1}{i}\frac{\delta}{\delta\bar{\eta}_{\beta(x)}}\right)\right]^V \\ &\times\sum_{F=0}^\infty\frac{1}{F!}\left[\int_{x',y'} i\bar{\eta}_\alpha(x')\left(\frac{1}{i}S_{\alpha\beta}(x'-y')\right)i\eta_\beta(y')\right]^F \\ &\times\sum_{p=0}^\infty\frac{1}{P!}\left[\frac{1}{2}\int_{x'',y''} iJ^\mu(x'')\left(\frac{1}{i}\Delta_{\mu\nu}(x''-y'')\right)iJ^\nu(y'')\right]^P. \end{aligned} \tag{12.1}$$

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 $-e\gamma^{\mu}_{\alpha\beta}$ to connect the lines in the graphical representation. The organization where factors of $1/i$ come together with the fermion and boson propagators has also the advantage that factors of i that appear in the LSZ reduction rules nicely cancel out against them.

Blackboard video

Graphical representation for partition function

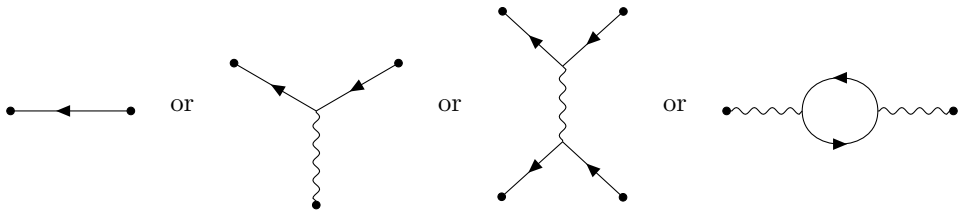
In the full expression for $Z[\bar{\eta}, \eta, J]$ many terms are present, in fact all graphs one can construct with fermion lines, photon lines and the vertex. For example



Note that these graphs need not be connected and the last diagram in the first line consists of two disconnected pieces.

Connected and disconnected diagrams

One distinguishes connected diagrams where all endpoints are connected with lines to each other, for example

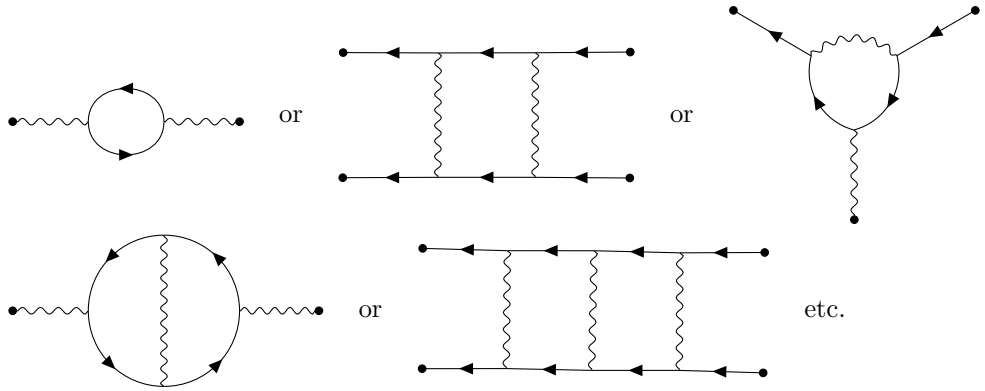


Disconnected diagrams can be decomposed into several connected diagrams.

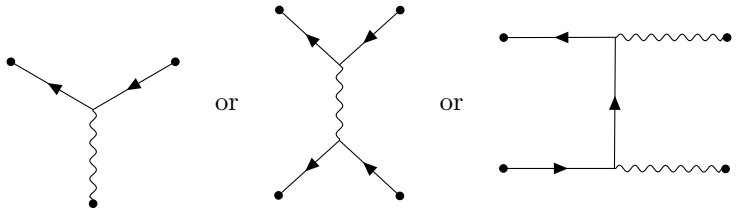
Blackboard video

Tree and loop diagrams

One also distinguishes tree diagrams and loop diagrams. Loop diagrams have closed loops of particle flow, for example



Tree diagrams have no closed loop, for example



Blackboard video

Counting the number of loops

Consider the partition function in (12.1). We have in each term

- V vertices, each with 3 functional derivatives
- F fermion propagators, each with 2 sources
- P photon propagators, each with 2 sources

At the end this will lead to a term with the following number of sources left

$$E = 2(F + P) - 3V.$$

It is also useful to count the number of internal lines (those not connected to one of the E sources) and it must be such that

$$2I = 3V - E.$$

Another useful formula relates the number of loops L to the number of internal lines and vertices

$$I = L + V - 1.$$

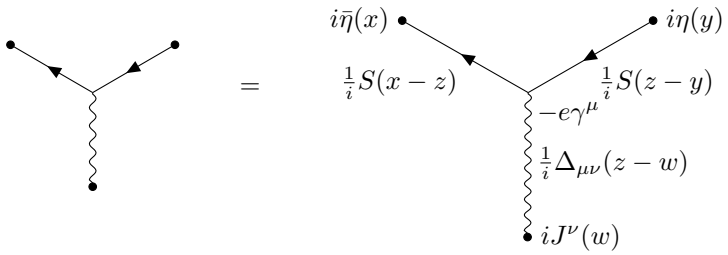
To see this one may start drawing each loop in a simple topology with just one vertex and one line and then to modify it by adding more vertices. One can combine these formulas to give the number of loops as

$$L = \frac{1}{2}V - \frac{1}{2}E + 1 = 2V - (F + P) + 1.$$

It is reassuring to check this formula on a few examples.

Corresponding algebraic expressions

To each of these diagrams with sources corresponds of course an algebraic expression, for example



$$= \int_{x,y,z,w} i\bar{\eta}(x) \left[\frac{1}{i}S(x-z)\right] (-e\gamma^\mu) \left[\frac{1}{i}S(z-y)\right] i\eta(y) \left[\frac{1}{i}\Delta_{\mu\nu}(z-w)\right] iJ^\nu(w).$$

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.

Blackboard video

S-matrix elements

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

$$\sqrt{2E_p}\; a_{\mathbf{p},\lambda}(\infty) \to i\epsilon^{\nu *}_{(\lambda)}(p) \int d^4x\; e^{-ipx} \left[-\partial_\mu \partial^\mu\right] A_\nu(x).$$

To obtain the field $A_\nu(x)$ under the functional integral we can use

$$A_\nu(x) \to \frac{1}{i} \frac{\delta}{\delta J^\nu(x)},$$

acting on $Z[\bar{\eta}, \eta, J]$. Moreover, $i[-\partial_\mu \partial^\mu]$ will remove one propagator line for the outgoing photon,

$$i\left[-\partial_\mu \partial^\mu\right] \frac{1}{i} \Delta_{\rho\sigma}(x-y) = \left[-\partial_\mu \partial^\mu\right] \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \frac{\mathscr{P}_{\rho\sigma}(p)}{p^2 - i\epsilon}$$
$$= \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \mathscr{P}_{\rho\sigma}(p) \to \eta_{\rho\sigma} \delta^{(4)}(x-y).$$

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

$$\epsilon^*_{(\lambda)\mu}(p)$$

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

$$\epsilon_{(\lambda)\mu}(p),$$

instead.

For out-going electrons we need to remove the external fermion propagator and multiply with $\bar{u}_s(\mathbf{p})$ 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 $u_s(p)$. For out-going positrons we need to contract with $v_s(p)$. For an incoming positron the corresponding external spinor is $\bar{v}_s(p)$.

Blackboard video

Propagators in momentum space

Working now directly in momentum space, the photon line is represented by

$$-i\Delta_{\mu\nu}(p) = -i\frac{\mathscr{P}_{\mu\nu}(p)}{p^2-i\epsilon} = -i\frac{\eta_{\mu\nu}-\frac{p_\mu\, p_\nu}{p^2}}{p^2-i\epsilon}.$$

The fermion line corresponds to

$$-iS(p) = -i\frac{-i\not{p}+m}{p^2+m^2-i\epsilon}.$$

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.

LECTURE 25

12.1 Elementary scattering processes

We are now ready to use the formalism of quantum field theory, specifically quantum electrodynamics, to determine actually scattering amplitudes and cross section. The incoming and outgoing states can consist of photons, electrons and positrons but also muons or anti-muons and more generally any charged particles. When the charged particles are scalar bosons, one would use a variant of the theory called scalar electrodynamics, but we are here concerned with charged spin-1/2 particles which are described by standard spinor electrodynamics.

In the following we will bring together several of the elements we have discussed before, such as

- the Lagrangian of spinor quantum electrodynamics,
- the idea of perturbation theory as an expansion in the coupling constant e,
- the graphical representation in terms of Feynman diagrams,
- solutions to the free Dirac equation for incoming or outgoing electrons and positrons (or muons and anti-muons),
- the propagators for Dirac femions and for photons,

It might be a good idea to go back and revise these topics if you feel uncertain about them. We will see on the way that we need some additional technical knowledge, specifically

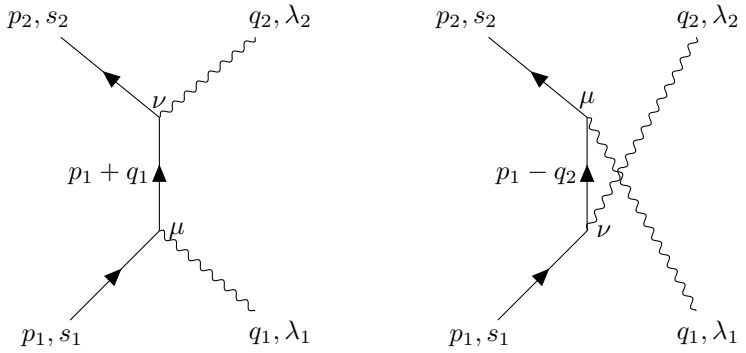
- how to do spin sums,
- how to calculate traces of gamma matrices
- how Mandelstam variables are defined and how one can work with them.

These points will also be discussed in the exercises.

We will then start to look at the elastic scattering of a photon and an electron, a process known as Compton scattering. We will write down the Feynman diagrams and the corresponding algebraic expressions. For another process, namely the scattering of an electron-positron pair to a muon-anti-muon pair we will do this, as well, but then also go on and evaluate the expressions further until we arrive at a nice and compact result for the scattering cross-section.

Compton Scattering

As a first example let us consider Compton scattering $e^-\gamma \to e^-\gamma$



These are two diagrams at order e^2, as shown above. The first diagram corresponds to the expression

$$\bar{u}_{s_2}(p_2)(-e\gamma^\nu)\left(-i\frac{-i(\not{p}_1+\not{q}_1)+m}{(p_1+q_1)^2+m^2}\right)(-e\gamma^\mu)\; u_{s_1}(p_1)\;\epsilon_{(\lambda_1)\mu}(q_1)\;\epsilon^*_{(\lambda_2)\nu}(q_2).$$

Similarly, the second diagram gives

$$\bar{u}_{s_2}(p_2)(-e\gamma^\mu)\left(-i\frac{-i(\not{p}_1-\not{q}_1)+m}{(p_1-q_2)^2+m^2}\right)(-e\gamma^\nu)\; u_{s_1}(p_1)\;\epsilon_{(\lambda_1)_\mu}(q_1)\;\epsilon^*_{(\lambda_2)\nu}(q_2).$$

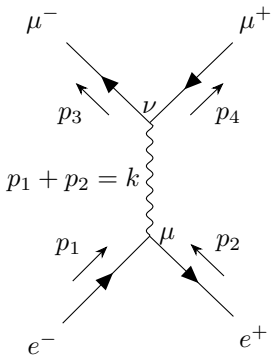
Combining terms and simplifying a bit leads to

$$i\mathcal{T} = -ie^2\epsilon_{(\lambda_1)\mu}(q_1)\;\epsilon^*_{(\lambda_2)\nu}(q_2)\;\bar{u}_{s_2}(p_2)\left[\gamma^\nu\;\frac{-i(\not{p}_1+\not{q}_1)+m}{(p_1+q_1)^2+m^2}\gamma^\mu+\gamma^\mu\frac{-i(\not{p}_1-\not{q}_2)+m}{(p_1-q_2)^2+m^2}\gamma^\nu\right]u_{s_1}(p_1).$$

Electron-positron to muon-anti-muon scattering

As another example for an interesting process in QED we consider $e^-e^+ \to \mu^-\mu^+$. 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

$$i\mathcal{T} = \bar{v}(p_2)(-e\gamma^\mu)\, u(p_1) \left(-i\frac{\eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}}{(k^2)}\right) \bar{u}(p_3)\, (-e\gamma^\nu)\, v(p_4),$$

with $k = p_1 + p_2 = p_3 + p_4$.

Blackboard video

On-shell conditions

The external momenta are on-shell and the spinors $u(p_1)$ etc. satisfy the Dirac equation,

$$(i\not{p}_1 + m_e)u(p_1) = 0, \qquad (-i\not{p}_4 + m_\mu)v(p_4) = 0,$$
$$\bar{u}(p_3)(i\not{p}_3 + m_\mu) = 0, \qquad \bar{v}(p_2)(-i\not{p}_2 + m_e) = 0.$$

This allows to write

$$\bar{v}(p_2)\, i\gamma^\mu k_\mu\, u(p_1) = \bar{v}(p_2)\, i(\not{p}_1 + \not{p}_2)\, u(p_1) = \bar{v}(p_2)\, (-m_e + m_e)\, u(p_1) = 0,$$
$$\bar{u}(p_3)\, i\gamma^\nu k_\nu\, v(p_4) = \bar{u}(p_3)\, i(\not{p}_3 + \not{p}_4)\, v(p_4) = \bar{u}(p_3)\, (-m_\mu + m_\mu)\, v(p_4) = 0.$$

These arguments show that the term $\sim k_\mu k_\nu$ can be dropped. This is essentially a result of gauge invariance.

Complex conjugate and squared amplitudes

We are left with

$$\mathcal{T} = -\frac{e^2}{k^2}\bar{v}(p_2)\gamma^\mu u(p_1)\, \bar{u}(p_3)\gamma_\mu v(p_4).$$

To calculate $|\mathcal{T}|^2$ we also need $\mathcal{T}^*$ which follows from hermitian conjugation

$$\mathcal{T}^* = -\frac{e^2}{k^2} v^\dagger(p_4)\gamma_\mu^\dagger \bar{u}^\dagger(p_3)\; u^\dagger(p_1)\gamma^{\mu\dagger}\bar{v}^\dagger(p_2).$$

Recall that $\bar{u}(p) = u(p)^\dagger\beta$ with $\beta = i\gamma^0$. With the explicit representation

$$\gamma^\mu = \begin{pmatrix} & -i\bar{\sigma}^\mu \\ -i\sigma^\mu & \end{pmatrix},$$

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

$$\mathcal{T}^* = -\frac{e^2}{k^2}\bar{v}(p_4)\gamma_\mu u(p_3)\ \bar{u}(p_1)\gamma^\mu v(p_2)$$

Putting together and using $s = -k^2 = -(p_1+p_2)^2$ we obtain

$$|\mathcal{T}|^2 = \frac{e^4}{s^2}\bar{u}(p_1)\gamma^\mu v(p_2)\ \bar{v}(p_2)\gamma^\nu u(p_1)\ \bar{u}(p_3)\gamma_\nu v(p_1)\ \bar{v}(p_4)\gamma_\mu u(p_3).$$

Blackboard video

Spin sums and averages

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)

$$\sum_{s=1}^{2} v_s(p_4)\bar{v}_s(p_4) = -i\not{p}_4 - m_\mu,$$

and similarly for μ^-

$$\sum_{s=1}^{2} u_s(p_3)\bar{u}_s(p_3) = -i\not{p}_3 + m_\mu.$$

We can therefore write

$$\sum_{\text{spins}} \bar{u}(p_3)\gamma_\nu v(p_4)\,\bar{v}(p_4)\gamma_\mu u(p_3) = \text{tr}\left\{(-i\not{p}_3 + m_\mu)\gamma_\nu(-i\not{p}_4 - m_\mu)\gamma_\mu\right\}.$$

Spins of the electron and positron must be averaged instead,

$$\frac{1}{2}\sum_{s=1}^{2} u(p_1)\bar{u}(p_1) = \frac{1}{2}(-i\not{p}_1 + m_e),$$
$$\frac{1}{2}\sum_{s=1}^{2} v(p_2)\bar{v}(p_2) = \frac{1}{2}(-i\not{p}_2 - m_e).$$

This leads to

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{T}|^2 = \frac{e^4}{4s^2}\text{tr}\left\{(-i\not{p}_1 + m_e)\gamma^\mu(-i\not{p}_2 - m_e)\gamma^\nu\right\} \times \text{tr}\left\{(-i\not{p}_3 + m_\mu)\gamma_\nu(-i\not{p}_4 - m_\mu)\gamma_\mu\right\}.$$

In order to proceed further, we need to know how to evaluate traces of up to four gamma matrices.

Blackboard video

Traces of gamma matrices

We need to understand how to evaluate traces of the form

$$\mathrm{tr}\{\gamma^{\mu_1}\cdots\gamma^{\mu_n}\}.$$

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, $\mathrm{tr}\{\mathbb{1}\}=4$. First we prove that traces of an odd number of gamma matrices must vanish,

$$\begin{aligned}\mathrm{tr}\{\gamma^{\mu_1}\cdots\gamma^{\mu_n}\} &= \mathrm{tr}\{\gamma_5^2\,\gamma^{\mu_1}\gamma_5^2\cdots\gamma_5^2\gamma^{\mu_n}\}\\ &= \mathrm{tr}\{(\gamma_5\gamma^{\mu_1}\gamma_5)\cdots(\gamma_5\gamma^{\mu_1}\gamma_5)\}\\ &= \mathrm{tr}\{(-\gamma_5^2\gamma_1^\mu)\cdots(-\gamma_5^2\gamma^{\mu_n})\}\\ &= (-1)^n\mathrm{tr}\{\gamma^{\mu_1}\cdots\gamma^{\mu_n}\}.\end{aligned}$$

This implies what we claimed.

Now for even numbers

$$\mathrm{tr}\{\gamma^\mu\gamma^\nu\}=\mathrm{tr}\{\gamma^\nu\gamma^\mu\}=\tfrac{1}{2}\mathrm{tr}\{\gamma^\mu\gamma^\nu+\gamma^\nu\gamma^\mu\}=\eta^{\mu\nu}\mathrm{tr}\{\mathbb{1}\}=4\eta^{\mu\nu}.$$

From this it also follows that

$$\mathrm{tr}\{\not{p}\not{q}\}=4p\cdot q.$$

Now consider $\mathrm{tr}\{\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\}$. This idea is to commute γ^μ to the right using $\{\gamma^\mu,\gamma^\nu\}=2\eta^{\mu\nu}$. Thus

$$\begin{aligned}\mathrm{tr}\{\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\} &= -\mathrm{tr}\{\gamma^\nu\gamma^\mu\gamma^\rho\gamma^\sigma\}+2\eta^{\mu\nu}\,\mathrm{tr}\{\gamma^\rho\gamma^\sigma\}\\ &= \mathrm{tr}\{\gamma^\nu\gamma^\rho\gamma^\mu\gamma^\sigma\}-2\eta^{\rho\mu}\mathrm{tr}\{\gamma^\nu\gamma^\sigma\}+2\eta^{\mu\nu}\,\mathrm{tr}\{\gamma^\rho\gamma^\sigma\}\\ &= -\mathrm{tr}\{\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^\mu\}+2\eta^{\sigma\mu}\,\mathrm{tr}\{\gamma^\nu\gamma^\rho\}-2\eta^{\rho\mu}\,\mathrm{tr}\{\gamma^\nu\gamma^\sigma\}+2\eta^{\mu\nu}\,\mathrm{tr}\{\gamma^\rho\gamma^\sigma\}.\end{aligned}$$

But by the cyclic property of the trace

$$\mathrm{tr}\{\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^\mu\}=\mathrm{tr}\{\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\}$$

which is also on the left hand side. Bringing it to the left and dividing by 2 gives

$$\begin{aligned}\mathrm{tr}\{\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\} &= \eta^{\sigma\mu}\,\mathrm{tr}\{\gamma^\nu\gamma^\rho\}-\eta^{\rho\mu}\,\mathrm{tr}\{\gamma^\nu\gamma^\sigma\}+\eta^{\mu\nu}\mathrm{tr}\{\gamma^\rho\gamma^\sigma\}\\ &= 4(\eta^{\sigma\mu}\eta^{\nu\rho}-\eta^{\rho\mu}\eta^{\nu\sigma}+\eta^{\mu\nu}\eta^{\rho\sigma}).\end{aligned}$$

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.

Blackboard video

Result so far

Coming back to $e^-e^+\to\mu^-\mu^+$ we find

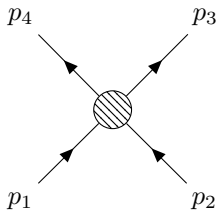
$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}}|\mathcal{T}|^2 =& \frac{4e^4}{s^2}\left[-p_1^\mu p_2^\nu-p_1^\nu p_2^\mu+(p_1\cdot p_2-m_e^2)\eta^{\mu\nu}\right]\\ &\times\left[-(p_3)_\nu(p_4)_\mu-(p_3)_\mu(p_4)_\nu+(p_3\cdot p_4-m_\mu^2)\eta^{\mu\nu}\right]\\ =& \frac{8e^4}{s^2}\left[(p_1\cdot p_4)(p_2\cdot p_3)+(p_1\cdot p_3)(p_2\cdot p_4)-m_\mu^2(p_1\cdot p_2)-m_e^2(p_3\cdot p_4)+2m_e^2m_\mu^2\right].\end{aligned}$$

This looks already quite decent but it can be simplified even further in terms of Mandelstam variables.

Blackboard video

Mandelstam Variables

The Mandelstam variables for a $2 \to 2$ process



are given by

$$\begin{aligned} s &= -(p_1 + p_2)^2 = -(p_3 + p_4)^2, \\ t &= -(p_1 - p_3)^2 = -(p_2 - p_4)^2, \\ u &= -(p_1 - p_4)^2 = -(p_2 - p_3)^2. \end{aligned}$$

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

$$s + t + u = -(p_1^2 + p_2^2 + p_3^2 + p_4^2) = m_1^2 + m_2^2 + m_3^2 + m_4^2.$$

Using these variables for example through

$$p_1 \cdot p_4 = -\frac{1}{2}\left[(p_1 - p_4)^2 - p_1^2 - p_4^2\right] = \frac{1}{2}\left[u - m_e^2 + m_\mu^2\right],$$

one finds for $e^- e^+ \to \mu^- \mu^+$

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{T}|^2 = \frac{2e^4}{s^2}\left[t^2 + u^2 + 4s(m_e^2 + m_\mu^2) - 2(m_e^2 + m_\mu^2)^2\right].$$

Blackboard video

Differential cross section

From the squared matrix element we can calculate the differential cross section in the center of mass frame. For relativistic kinematics of $2 \to 2$ scattering and the normalization conventions we employ here one has in the center of mass frame

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s}\frac{|\mathbf{p}_3|}{|\mathbf{p}_1|}\frac{1}{4}\sum_{\text{spins}} |\mathcal{T}|^2.$$

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.

$$\begin{aligned} |\mathbf{p}_1| &= \sqrt{E^2 - m_e^2}, & s &= 4E^2, \\ |\mathbf{p}_3| &= \sqrt{E^2 - m_\mu^2}, & t &= m_e^2 + m_\mu^2 - 2E^2 + 2\mathbf{p}_1 \cdot \mathbf{p}_3, \\ \mathbf{p}_1 \cdot \mathbf{p}_3 &= |\mathbf{p}_1||\mathbf{p}_3| \cos\theta, & u &= m_e^2 + m_\mu^2 - 2E^2 - 2\mathbf{p}_1 \cdot \mathbf{p}_3. \end{aligned}$$

With these relations we can express $d\sigma/d\Omega$ in terms of E and θ only.

Blackboard video

Ultrarelativistic limit

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 $|\mathbf{p}_1| = |\mathbf{p}_3|$ and

$$t^2 + u^2 = 8E^4(1 + \cos^2\theta), \qquad \frac{2(t^2+u^2)}{s^2} = 1 + \cos^2\theta,$$

which leads to

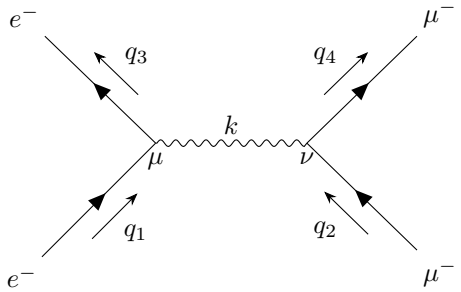
$$\frac{d\sigma}{d\Omega} = \frac{e^4}{64\pi^2 s}(1 + \cos^2\theta) = \frac{\alpha^2}{4s}(1 + \cos^2\theta).$$

In the last equation we used $\alpha = e^2/(4\pi)$.

Blackboard video

Electron-Muon Scattering

We can also consider the scattering process $e^-\mu^- \to e^-\mu^-$,



$$i\mathcal{T} = \bar{u}(q_3)(-e\gamma^\mu)u(q_1)\left(-i\frac{\eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}}{k^2}\right)\bar{u}(q_4)(-e\gamma^\nu)u(q_2).$$

By a similar argument as before the term $\sim k_\mu k_\nu$ drops out,

$$\mathcal{T} = -\frac{e^2}{(q_1 - q_3)^2}\bar{u}(q_3)\gamma^\mu u(q_1)\bar{u}(q_4)\gamma_\mu u(q_2) \qquad (e^-\mu^- \to e^-\mu^-).$$

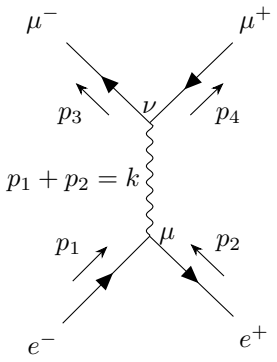
Blackboard video

Comparison to electron to muon scattering

Compare this to what we have found for $e^-e^+ \to \mu^-\mu^+$

$$\mathcal{T} = -\frac{e^2}{(p_1 + p_2)^2}\bar{v}(p_2)\gamma^\mu u(p_1)\bar{u}(p_3)\gamma_\mu v(p_4),$$

where the conventions were according to



There is a close relation and the expressions agree if we put

$$\begin{aligned} q_1 &= +p_1, & u(q_1) &= u(p_1), \\ q_2 &= -p_4, & u(q_2) &= u(-p_4) \to v(p_4), \\ q_3 &= -p_2, & \bar{u}(q_3) &= \bar{u}(-p_2) \to \bar{v}(p_2), \\ q_4 &= +p_3, & \bar{u}(q_4) &= \bar{u}(p_3). \end{aligned}$$

Crossing symmetry

Recall that

$$(i\not{p} + m)\, u(p) = 0 \qquad \text{but} \qquad (-i\not{p} + m)\, v(p) = 0.$$

However one sign arises from the spin sums

$$\begin{aligned} \sum_{s=1}^{2} u_s(p)\bar{u}_s(p) &= -i\not{p} + m, \\ \sum_{s=1}^{2} v_s(p)\, \bar{v}_s(p) &= -i\not{p} - m = -\sum_{s} u_s(-p)\, \bar{u}_s(-p). \end{aligned}$$

Because it appears twice, the additional sign cancels for $|\mathcal{T}|^2$ after spin averaging and one finds indeed the same result as for $e^-e^+ \to \mu^-\mu^+$ but with

$$\begin{aligned} s_q &= -\,(q_1 + q_2)^2 = -(p_1 - p_4)^2 = u_p, \\ t_q &= -\,(q_1 - q_3)^2 = -(p_1 + p_2)^2 = s_p, \\ u_q &= -\,(q_1 - q_4)^2 = -(p_1 - p_3)^2 = t_p. \end{aligned}$$

We can take what we had calculated but must change the role of s, $t$and u ! This is an example of *crossing symmetries*.

Blackboard video

Electron-muon scattering in the massless limit

Recall that we found for $e^-e^+ \to \mu^-\mu^+$ in the massless limit $m_e = m_\mu = 0$ simply

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{T}|^2 = \frac{2e^4}{s^2} \left[t^2 + u^2\right].$$

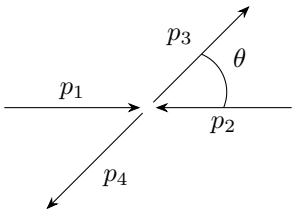
For $e^-\mu^- \to e^-\mu^-$ we find after the replacements $u \to s$, $s \to t$, $t \to u$,

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{T}|^2 = \frac{2e^4}{t^2}\left[u^2 + s^2\right].$$

Blackboard video

More on Mandelstam variables

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.



$$p_1^\mu = (E, \mathbf{p}), \qquad p_2^\mu = (E, -\mathbf{p}),$$
$$p_3^\mu = (E, \mathbf{p}'), \qquad p_4^\mu = (E, -\mathbf{p}').$$

While

$$s = -(p_1 + p_2)^2 = (2E)^2$$

measures the center of mass energy,

$$t = -(p_1 - p_3) = -2\mathbf{p}^2[1 - \cos(\theta)]$$

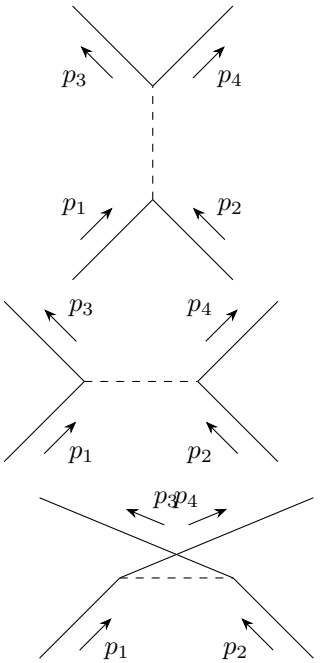
is a momentum transfer that vanishes in the soft limit $\mathbf{p}^2 \to 0$ and in the colinear limit $\theta \to 0$. Similarly,

$$u = -(p_1 - p_4)^2 = -2\mathbf{p}^2[1 + \cos(\theta)]$$

vanishes for $\mathbf{p}^2 \to 0$ and for backward scattering $\theta \to \pi$.

s-, t- and u-channels

One speaks of interactions in different channels for tree diagrams of the following generic types,



$$\propto \frac{1}{-s+m^2} \qquad s\text{-channel}$$

$$\propto \frac{1}{-t+m^2} \qquad t\text{-channel}$$

$$\propto \frac{1}{-u+m^2} \qquad u\text{-channel}$$

Electron-muon scattering

For the cross section we find for $e^-\mu^- \to e^-\mu^-$ in the massless limit

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s}\frac{1}{4}\sum_{\text{spins}} |\mathcal{T}|^2 = \frac{\alpha^2[4+(1+\cos\theta)^2]}{2s(1-\cos\theta)^2}$$

This diverges in the colinear limit $\theta \to 0$ 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.

LECTURE 26

12.2 Higgs/Yukawa theory

In the following two lectures we will discuss a quantum field theoretic model that extends somewhat beyond quantum electrodynamics. We add to the theory a neutral massive scalar field that couples to the fermions through a Yukawa interaction. One may see that additional massive scalar particle as an analog of the Higgs boson, even though our model reflects only a few of the properties of the real electroweak standard model.

We discuss the model as a further example for an interesting quantum field theory and because we can nicely study there decay processes.

- A massive Higgs boson can decay into two fermions through the Yukawa interaction. This is a tree level process and rather easy to calculate.

- Interestingly a neutral and massive Higgs boson can also decay into two photons. This process is not allowed at tree level (because the Higgs boson is neutral), but it is induced by loop diagrams. This will be the first loop diagram we will calculate in detail.

In the second part of the lecture course loop diagrams and their physical consequences will be studied in much more detail. For the Higgs decay into photons we do not need renormalization yet, which simplifies the discussion. Nevertheless there will be some new elements to be discussed.

Action for Higgs/Yukawa theory and fermion mass

Let us consider the following extension of QED by a neutral scalar field (with $m = gv$)

$$S[\bar{\psi}, \psi, A, \phi] = \int_x \left\{ -\bar{\psi}\gamma^\mu \left(\partial_\mu - ieA_\mu\right)\psi - m\bar{\psi}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}\phi\left(-\partial_\mu\partial^\mu + M^2\right)\phi - g\phi\bar{\psi}\psi \right\}.$$

Note that a constant (homogeneous) scalar field ϕ modifies the fermion mass according to

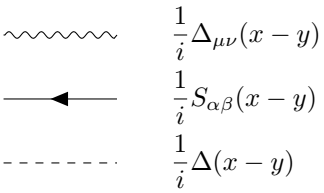
$$m_{\text{eff}} = m + g\phi = g(v + \phi)$$

In fact, one can understand the massses 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.

Blackboard video

Propagators and vertices

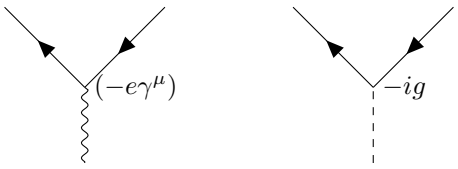
In the theory above we have now different propagators



with scalar propagator

$$\Delta(x-y) = \int_p e^{ip(x-y)} \frac{1}{p^2 + M^2}.$$

The vertices are

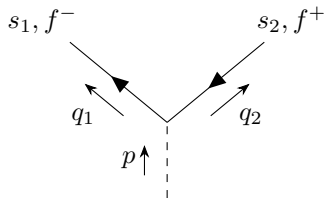


Blackboard video

12.3 Higgs decay into fermions

Higgs decay to fermions

Let us discuss first the process $\phi \to f^- f^+$. The fermions could be leptons (e, μ, τ) or quarks (u, d, s, c, b, t). The Feynman diagram for the decay is simply



According to the Feynman rules we obtain

$$\mathcal{T} = g\, \bar{u}_{s_1}(q_1) v_{s_2}(q_2).$$

For the absolute square one finds

$$|\mathcal{T}|^2 = g^2\, \bar{u}_{s_1}(q_1) v_{s_2}(q_2)\, \bar{v}_{s_2}(q_2) u_{s_1}(q_1).$$

Blackboard video

Spin sums and Dirac traces

We will assume that the final spins are not observed and sum them

$$\sum_{\text{spins}} |\mathcal{T}|^2 = g^2\, \text{tr}\left\{(-i\not{q}_2 - m)(-i\not{q}_1 + m)\right\}$$

We used here again the spin sum formula

$$\sum_s v_s(p)\bar{v}_s(p) = -i\not{p} - m, \qquad \sum_s u_s(p)\bar{u}_s(p) = -i\not{p} + m.$$

Performing also the Dirac traces gives

$$\sum_{\text{spins}} |\mathcal{T}|^2 = g^2 \left(-4q_1 \cdot q_2 - 4m^2\right).$$

Kinematics in the Higgs boson rest frame

Let us now go into the rest frame of the decaying particle where

$$p = (M, 0, 0, 0), \qquad q_1 = \left(\tfrac{M}{2}, \mathbf{q}\right), \qquad q_2 = \left(\tfrac{M}{2}, -\mathbf{q}\right),$$

with

$$\mathbf{q}^2 = -m^2 + \tfrac{M^2}{4}, \qquad q_1 \cdot q_2 = -\frac{M^2}{4} - \mathbf{q}^2 = -\frac{M^2}{2} + m^2,$$

and

$$\sum_{\text{spins}} |\mathcal{T}|^2 = 2\, g^2 M^2 \left(1 - 4\frac{m^2}{M^2}\right).$$

Note that the decay is kinematically possible only for $M > 2m$ so that the bracket is always positive.

Blackboard video

Decay rate

For the particle decay rate we get

$$\frac{d\Gamma}{d\Omega} = \frac{|\mathbf{q}_1|}{32\pi^2 M^2} \sum_{\text{spins}} |\mathcal{T}|^2 = \frac{g^2 M}{32\pi^2} \left(1 - 4\frac{m^2}{M^2}\right)^{3/2}.$$

Because this is independent of the solid angle Ω one can easily integrate to obtain the decay rate

$$\Gamma = \frac{g^2 M}{8\pi} \left(1 - 4\frac{m^2}{M^2}\right)^{3/2}.$$

Blackboard video

Dependence on fermion mass

If the scalar boson ϕ is the Higgs boson, the Yukawa coupling is in fact proportional to the fermion mass m,

$$g = \frac{m}{v}.$$

One has then

$$\Gamma = \frac{M^3}{32\pi v^2} f\left(\frac{2m}{M}\right)$$

where

$$f(x) = x^2(1 - x^2)^{3/2}$$

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.)

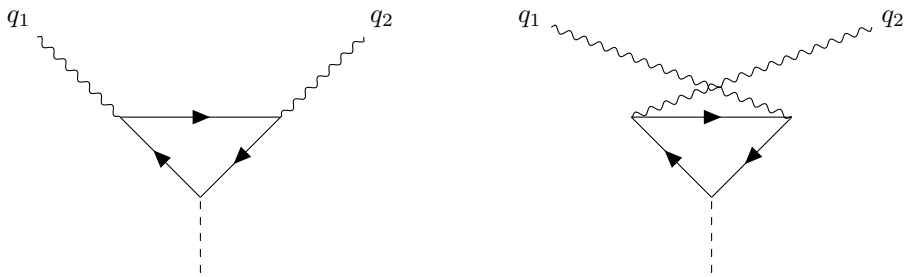
LECTURE 27

12.4 Higgs decay into photons

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 the diagrams



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.

Signs in fermion loops

Schematically, the vertices are derivatives

$$\left[(-e\gamma^\mu)\left(\frac{1}{i}\frac{\delta}{\delta J^\mu}\right)\left(i\frac{\delta}{\delta\eta}\right)\left(\frac{1}{i}\frac{\delta}{\delta\bar{\eta}}\right)\right] \qquad \text{or} \qquad \left[(-ig)\left(\frac{1}{i}\frac{\delta}{\delta J}\right)\left(i\frac{\delta}{\delta\eta}\right)\left(\frac{1}{i}\frac{\delta}{\bar{\eta}}\right)\right]$$

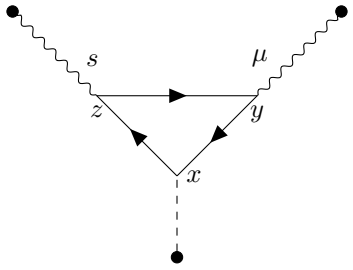
and they act here on a chain like

$$\left[(i\bar{\eta})\left(\frac{1}{i}S\right)(i\eta)\right]\left[(i\bar{\eta})\left(\frac{1}{i}S\right)(i\eta)\right]\left[(i\bar{\eta})\left(\frac{1}{i}S\right)(i\eta)\right].$$

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.*

Position space representation

In position space and including sources, the first diagram is

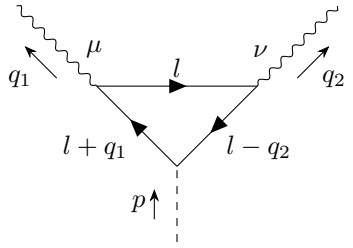


$$(-1)(-ig)\int_{x,y,z}\mathrm{tr}\left\{\left[\frac{1}{i}S(x-y)\right](-e\gamma^\mu)\left[\frac{1}{i}S(y-z)\right](-e\gamma^\nu)\left[\frac{1}{i}S(z-x)\right]\right\}$$
$$\times\int_{u,v,w}\left[\frac{1}{i}\Delta_{\mu\alpha}(y-u)\right](iJ^\alpha(u)\left[\frac{1}{i}\Delta_{\nu\beta}(z-v)\right](iJ^\beta(v)\left[\frac{1}{i}\Delta(x-w)\right](iJ(w))$$

The trace is for the Dirac matrix indices.

Momentum space representation for first diagram

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 (taking a factor (-1) from the closed fermion loop, $-ig$ from the Yukawa vertex and i^3 from the LSZ reduction into account)



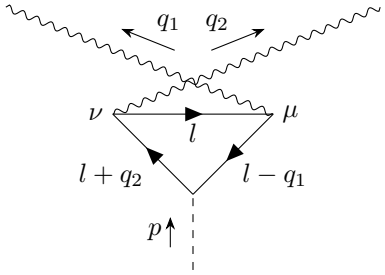
$$ge^2\,\epsilon^*_\mu(q_1)\epsilon^*_\nu(q_2)\int_l \frac{1}{[l+q_1)^2+m^2-i\epsilon][l^2+m^2-i\epsilon][l-q_2)^2+m^2+i\epsilon]}$$
$$\times\,\mathrm{tr}\left\{\left[-i(\not{l}+\not{q}_1)+m\right]\gamma^\mu\left[-i\not{l}+m\right]\gamma^n u\left[-i(\not{l}-\not{q}_2)+m\right]\right\}$$

Here we use here the abbreviation

$$\int_l = \int \frac{d^4l}{(2\pi)^4}.$$

Momentum space representation for second diagram

For the second diagram we can write



$$ge^2\,\epsilon^*_\mu(q_1)\,\epsilon^*_\nu(q_2)\int_l \ldots$$

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.

Analytic continuation and Dirac traces

The Feynman $i\epsilon$ terms allow to perform a Wick rotation to Euclidean space $l^0 = i\tilde{l}^0_E$ so that l^2 is then positive. 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

$$\begin{aligned}
&\mathrm{tr}\left\{\left[-i(\not{l}+\not{q}_1)+m\right]\gamma^\mu\left[-i\not{l}+m\right]\gamma^\nu\left[-i(\not{l}-\not{q}_2)+m\right]\right\}\\
&= -m\,\mathrm{tr}\left\{(\not{l}+\not{q}_1)\gamma^\mu\not{l}\gamma^\nu+(\not{l}+\not{q}_1)\gamma^\mu\gamma^\nu(\not{l}-\not{q}_2)+\gamma^\mu\not{l}\gamma^\nu(\not{l}-\not{q}_2)\right\}+m^3\,\mathrm{tr}\{\gamma^\mu\gamma^\nu\}\\
&= -4m\Big[(l+q_1)^\mu l^\nu+(l+q_1)^\nu l^\mu-(l+q_1)\cdot l\,\eta^{\mu\nu}\\
&\qquad+(l+q_1)^\mu(l-q_2)^\nu+(l+q_1)\cdot(l-q_2)\eta^{\mu\nu}-(l+q_1)^\nu(l-q_2)^\mu\\
&\qquad+l^\mu(l-q_2)^\nu+(l-q_2)^\mu l^\nu-\eta^{\mu\nu}l\cdot(l-q_2)\Big]+4\eta^{\mu\nu}m^3\\
&= -4m\left[4l^\mu l^\nu-l^2\eta^{\mu\nu}+2q_1^\mu l^\nu-2q_2^\nu l^\mu-q_1^\mu q_2^\nu+q_1^\nu q_2^\mu-(q_1\cdot q_2)\eta^{\mu\nu}\right]+4\eta^{\mu\nu}m^3.
\end{aligned}$$

Feynman parameters

Let us now consider the denominator. One can introduce so-called Feynman parameters to write

$$\begin{aligned}
&\frac{1}{[(l+q_1)^2+m^2][l^2+m^2][(l-q_2)^2+m^2]}\\
&= 2!\int_0^1 du_1\cdots du_3\,\delta(u_1+u_2+u_3-1)\frac{1}{\left[u_1[(l+q_1)^2+m^2]+u_2[l^2+m^2]+u_3[(l-q_2)^2+m^2]\right]^3}\\
&= 2\int_0^1 du_1\cdots du_3\frac{\delta(u_1+u_2+u_3-1)}{\left[l^2+2l(u_1q_1-u_3q_2)+u_1q_1^2+u_3q_2^2+m^2\right]^3}.
\end{aligned}$$

We have used here the identity (will be proven in the exercise classes)

$$\frac{1}{p_1 \cdots p_n} = (n-1)! \int_0^1 du_1 \dots du_n \frac{\delta(u_1 + \dots + u_n - 1)}{[u_1 A_1 + \dots + u_n A_n]^n}.$$

In a next step one commutes the integral over $u_1 \dots u_3$ with the integral over l.

Shifting momenta

It is useful to change integration variables according to

$$l + u_1 q_1 - u_3 q_2 \to k,$$
$$l = k - u_1 q_1 + u_3 q_2.$$

Collecting terms we find for the first diagram

$$ge^2\, \epsilon_\mu^*(q_1)\, \epsilon^*(q_2)\, 2 \int_0^1 du_1 \cdots du_3\, \delta(u_1{+}u_2{+}u_3{-}1) \int \frac{d^4k}{(2\pi)^4} \frac{A^{\mu\nu}}{\left[k^2 + u_1 q_1^2 + u_3 q_2^2 - (u_1 q_1 - u_3 q_2)^2 + m^2\right]^3},$$

where the numerator contains the combination

$$\begin{aligned} A^{\mu\nu} = -4m\Big[& 4k^\mu k^\nu - k^2 \eta^{\mu\nu} + \text{terms linear in } k \\ & + 4(u_1 q_1 - u_3 q_2)^\mu (u_1 q_1 - u_3 q_2)^\nu - (u_1 q_1 - u_3 q_2)^2 \eta^{\mu\nu} \\ & - q_1^\mu q_2^\nu + q_1^\nu q_2^\mu - (q_1 \cdot q_2)\eta^{\mu\nu} - \eta^{\mu\nu} m^2 \Big]. \end{aligned}$$

The integral over k is now symmetric around the origin. Accordingly, there is no contribution from linear terms in k

$$\int \frac{d^dk}{(2\pi)^d} k^\mu f(k^2) = 0.$$

Some simplifications are also possible due to $\epsilon_\mu^*(q_1) q_1^\mu = \epsilon_\nu^*(q_2) q_2^\nu = 0$ and $q_1^2 = q_2^2 = 0$, and we can replace

$$A^{\mu\nu} = -4m\Big[4k^\mu k^\nu - k^2 \eta^{\mu\nu} + (1 - 4u_1 u_3) q_2^\mu q_1^\nu + (2u_1 u_3 - 1) q_1 \cdot q_2 \eta^{\mu\nu} - \eta^{\mu\nu} m^2\Big].$$

Dimensional regularization

In manipulating the remaining integral over k we need to be careful because a first analysis based on power counting suggests that the integral of the terms quadratic in k in the numerator might not converge (in the numerator we have $d^4k \times k^2$ and in the denominator $(k^2 + A)^3$ which allows a logarithmic divergence in the UV regime). We need to first introduce some regularization and then do the analysis carefully. An often used method is to extend the integrals from four spacetime dimensions to d spacetime dimensions, where d can actually be taken a complex number, see below. For d slightly different from 4 one finds convergent expressions and one can take the limit $d \to 4$ at the end. Let us now proceed using this *dimensional regularization*.

For the quadratic terms we use the identity

$$\int \frac{d^dk}{(2\pi)^d} k^\mu k^\nu f(k^2) = \frac{1}{d} \eta^{\mu\nu} \int \frac{d^dk}{(2\pi)^d} k^2 f(k^2). \tag{12.2}$$

The overall structure follows from Lorentz invariance and the prefactor from taking the trace on both sides.

With this, all remaining integrals are of the form

$$\int \frac{d^dp}{(2\pi)^d} \frac{p^{2a}}{(p^2 + A)^b} = \frac{\Omega_d}{(2\pi)^d} \int_0^\infty dp \frac{p^{2a+d-1}}{(p^2 + A)^b}. \tag{12.3}$$

Here we have used the surface of the unit sphere in d dimensions Ω_d. To derive a formula for the latter we perform a d-dimensional Gaussian integral twice, first in cartesian then in polar coordinates,

$$\int d^d x\, e^{-x^2} = \pi^{d/2} = \Omega_d \int_0^\infty dx\, x^{d-1} e^{-x^2} = \frac{\Omega_d}{2} \int_0^\infty dt\, t^{\frac{d}{2}-1} e^{-t} = \frac{\Omega_d}{2}\Gamma(d/2).$$

This yields the formula

$$\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}.$$

We used here the Euler Gamma function $\Gamma(z)$ with the notable properties $\Gamma(z+1) = z\Gamma(z)$ for $\mathrm{Re}(z) > 0$ and

$$\Gamma(n+1) = n!, \qquad \Gamma(n+1/2) = \frac{(2n)!}{n!4^n}\sqrt{\pi},$$

for non-negative integers n. Indeed one finds with this the known special cases

$$\Omega_1 = 2, \qquad \Omega_2 = 2\pi, \qquad \Omega_3 = 4\pi, \qquad \Omega_4 = 2\pi^2.$$

The remaining integral over the magnitude of momentum can actually also be evaluated in terms of Gamma functions,

$$\int_0^\infty dp \frac{p^{2a+d-1}}{(p^2+A)^b} = \frac{\Gamma(b-a-d/2)\Gamma(a+d/2)}{2\Gamma(b)} A^{-(b-a-d/2)}.$$

Taken together, this leads to the useful result

$$\int \frac{d^d p}{(2\pi)^d} \frac{p^{2a}}{(p^2+A)^b} = \frac{\Gamma(b-a-d/2)\Gamma(a+d/2)}{(4\pi)^{d/2}\Gamma(b)\Gamma(d/2)} A^{-(b-a-d/2)}. \tag{12.4}$$

In this context, the following properties of the gamma function is very useful for integer $n \geq 0$ and small x,

$$\Gamma(-n+x) = \frac{(-1)^n}{n!}\left[\frac{1}{x} - \gamma + \sum_{k=1}^n \frac{1}{k} + \mathcal{O}(x)\right],$$

where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant. In particular $\Gamma(z)$ has a simple pole at the origin.

Result so far

Using dimensional regularization as outlined above, it is now a straight-forward excercise to prove that

$$\lim_{d\to 4} \int \frac{d^d k}{(2\pi)^d} \frac{4k^\mu k^\nu - (k^2+A)\eta^{\mu\nu}}{(k^2+A)^3} = 0.$$

Taking this into account leads to

$$A^{\mu\nu} = -4m\left[1 - 4u_1 u_2\right]\left[q_1^\nu q_2^\mu - (q_1\cdot q_2)\eta^{\mu\nu}\right].$$

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

$$\begin{aligned}\mathcal{T} = &-8ge^2 m\, \epsilon_\mu^*(q_1)\epsilon_\nu^*(q_2)\left[q_1^\nu q_2^\mu - (q_1\cdot q_2)\eta^{\mu\nu}\right] \\ &\times 2\int_0^1 du_1 \cdots du_3\, \delta(u_1+u_2+u_3-1)[1-4u_1u_3] \int \frac{d^4k}{(2\pi)^4} \frac{1}{\left[k^2 + 2u_1u_3 q_1\cdot q_2 + m^2\right]^3}\end{aligned}$$

Momentum integral

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

$$\int \frac{d^4k}{(2\pi)^4} \frac{1}{[k^2 + m^2]^3} = i \frac{1}{(4\pi)^2} \frac{1}{2m^2}.$$

This i is due to the Wick rotation $k^0 = ik_E^0$.

Integral over Feynman parameters

Also the integral over Feynman parameters can now easily be performed

$$\begin{aligned}
&2 \int_0^1 du_1 \dots du_3 \, \delta(u_1 + u_2 + u_3 - 1)[1 - 4u_1 u_3] \\
&= 2 \int_0^1 du_1 du_3 \, \theta(1 - u_1 - u_3) \, [1 - 4u_1 u_3] \\
&= 2 \int_0^1 du_1 \int_0^{1-u_1} du_3 \, [1 - 4u_1 u_3] \\
&= 2 \int_0^1 du_1 [(1 - u_1) - 4u_1 \tfrac{1}{2}(1 - u_1)^2] \\
&= 2 - 3 + \frac{8}{3} - 1 = \frac{2}{3}.
\end{aligned}$$

Collecting terms we find

$$i\mathcal{T} = i \frac{8ge^2}{3(4\pi)^2 m} \, \epsilon_\mu^*(q_1) \, \epsilon_\nu^*(q_2) \, [q_1^\nu q_2^\mu - (q_1 \cdot q_2)\eta^{\mu\nu}] \, .$$

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

$$\sum_{\text{polarizations}} |\mathcal{T}|^2 = \sum_{\text{polarizations}} \epsilon_\mu^*(q) \epsilon_\nu(q) \mathcal{M}^\mu(q) \mathcal{M}^{\nu *}(q).$$

We have extended here the polarization vector of a photon from the amplitude by decomposing

$$\mathcal{T} = \epsilon_\mu^*(q) \mathcal{M}^\mu(q).$$

Let us choose without loss of generality $q^\mu = (E, 0, 0, E)$ and use the polarization vector introduced previously,

$$\begin{aligned}
\epsilon_\mu^{(1)} &= \left(0, \frac{1}{\sqrt{2}}, -\frac{i}{\sqrt{2}}, 0\right), \\
\epsilon_\mu^{(2)} &= \left(0, \frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0\right),
\end{aligned}$$

such that

$$\epsilon_\mu^{*(1)} \, \epsilon_\nu^{(1)} + \epsilon_\mu^{*(2)} \, \epsilon_\nu^{(2)} = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}.$$

This would give

$$\sum_{j=1}^{2} \epsilon_\mu^{*(j)} \, \epsilon_\nu^{(j)} \, \mathcal{M}^\mu \, \mathcal{M}^{*\nu} = |\mathcal{M}^1|^2 + |\mathcal{M}^2|^2.$$

Ward identity

To simplify this one can use an identity known as the *Ward identity*,

$$q_\mu \mathcal{M}^\mu(q) = 0.$$

This is a direct consequence of gauge symmetry. For the above choice of q^μ it follows

$$-\mathcal{M}^0 + \mathcal{M}^3 = 0.$$

Accordingly, one can add $0 = -|\mathcal{M}^0|^2 + |\mathcal{M}^3|^2$ to the spin sum

$$\sum_{j=1}^{2} \epsilon_\mu^{*(j)}\, \epsilon_\nu^{(j)}\, \mathcal{M}^\mu\, \mathcal{M}^{*\nu} = -|\mathcal{M}^0|^2 + |\mathcal{M}^1|^2 + |\mathcal{M}^2|^2 + |\mathcal{M}^3|^2 = \eta_{\mu\nu}\mathcal{M}^\mu \mathcal{M}^{*\nu}.$$

In this sense we can use for external photons

$$\sum_{j=1}^{2} \epsilon_\mu^{*(j)}\, \epsilon_\nu^{(j)} \to \eta_{\mu\nu}.$$

Squared amplitude

With this we can now calculate the sums over final state photon polarizations

$$\begin{aligned}\sum_{\text{pol.}} |\mathcal{T}|^2 &= \left(\frac{8ge^2}{3\,(4\pi)^2\,m}\right)^2 \Big[q_1^\nu q_2^\mu - (q_1\cdot q_2)\eta^{\mu\nu}\Big]\Big[q_1^\beta q_2^\alpha - (q_1\cdot q_2)\eta^{\alpha\beta}\Big] \\ &\quad \times \sum_{\text{pol.}} \epsilon_\mu^*(q_1)\,\epsilon_\alpha(q_1) \sum_{\text{pol.}} \epsilon_\nu^*(q_2)\,\epsilon_\beta(q_2) \\ &= \left(\frac{8ge^2}{3\,(4\pi)^2\,m}\right)^2 2(q_1\cdot q_2)^2 = \frac{2g^2\alpha^2}{9\pi^2 m^2} M^4.\end{aligned}$$

In the last step we have used that the momentum of the incoming Higgs particle is $p = q_1 + q_2$. The square is given by the rest mass, $p^2 = -M^2 = 2(q_1 \cdot q_2)$. Here we also used that the photons are massless, $q_1^2 = q_2^2 = 0$. We also used the fine structure constant $\alpha = e^2/(4\pi)$.

Decay rate

For the differential particle decay rate $\varphi \to \gamma\gamma$ this gives in the rest frame of the Higgs particle with $|\mathbf{q}_1| = M/2$,

$$\frac{d\Gamma}{d\Omega} = \frac{|\mathbf{q}_1|}{32\pi^2 M^2} \sum_{\text{pol.}} |\mathcal{T}|^2 = \frac{g^2\alpha^2}{9\times 32\pi^4 m^2} M^3.$$

Finally, we integrate over solid angle $\Omega = (1/2)4\pi$ where the factor $(1/2)$ is due to the fact that the photons in the final state are indistinguishable. The decay rate for $\varphi \to \gamma\gamma$ through a heavy fermion loop is finally

$$\Gamma = \frac{g^2\alpha^2}{144\pi^3 m^2}\, M^3$$

Note that because of $g = m/v$ this is in fact independant of the heavy fermion mass m.

LECTURE 28

12.5 One-loop corrections in quantum electrodynamics

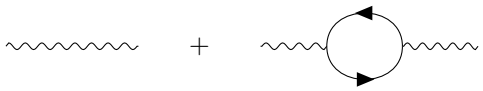
In this section we consider systematically one loop corrections in perturbative quantum electrodynamics. We have calculated previously a one-loop expression for a process that has no tree-level contribution and found a finite result. In contrast to this, if one finds one-loop corrections to processes that actually do have a tree-level contribution, one finds formally infinite results. To deal with these we need to first carefully introduce a *regularization* of the theory (we can use dimensional regularization), and in a second step we have to reinterpret what we are doing from a different persepctive. This leads to *renormalisation*. Specifically, it turns out that the terms appearing in the Lagrangian are in fact themselves subject to corrections from quantum fluctuations from loop diagrams.

Corrected propagators and vertices

Consider again the Lagrangian of QED as written before

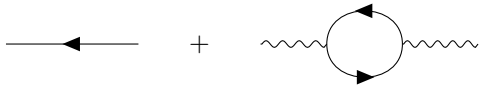
$$\mathscr{L} = -\bar{\Psi}\gamma^\mu(\partial_\mu - ieA_\mu)\Psi - m\bar{\Psi}\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}.$$

This Lagrangian enters the functional integral but it is in fact not directly observable. Corrections from quantum fluctuations are always present and one can in fact only observe asymptotic particles described by the full propagators with fluctuation corrections taken into account, and similarly vertices with corrections taken into account. Specifically, at one loop order we have for the effective photon propagator the tree-level term plus a one-loop term,

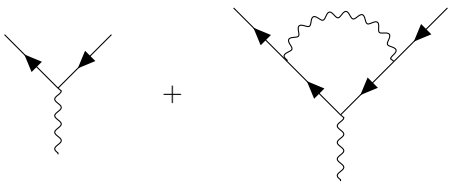


and only the sum (including in fact further loop corrections at higher order) can actually describe a propagating photon.

Similarly, the fermion propagator at on-lopp order contains the tree-level term plus a one loop term,



and more terms at higher order. Finally, the vertex is a combination of the tree-level expression and a loop expression,



and more terms would contribute at higher orders.

Modified Lagrangian

Motivated by these considerations we modify the Lagrangian by introducing *wave function renormalization factors* Z_Ψ and Z_A for the fields through the replacements

$$\Psi \to \sqrt{Z_\Psi}\Psi, \qquad \bar{\Psi} \to \sqrt{Z_\Psi}\bar{\Psi}, \qquad A_\mu \to \sqrt{Z_A}A_\mu,$$

and we replace the fermion mass by

$$m \to m + \Delta m,$$

and the electric charge by

$$e \to e + \Delta e.$$

The Lagrangian becomes thus

$$\mathscr{L} = -Z_\Psi \bar{\Psi}\gamma^\mu(\partial_\mu - i(e+\Delta e)\sqrt{Z_A}A_\mu)\Psi - (m+\Delta m)Z_\Psi \bar{\Psi}\Psi - \frac{1}{4}Z_A F_{\mu\nu}F^{\mu\nu}.$$

This is supposed to enter the functional integral of the quantum field theory which has been regularized, for example by putting it on a spacetime lattice, by introducing a UV cutoff, or through dimensional regularization.

The idea is now to determine the parameters Z_Ψ, Z_A and Δm through physics conditions on the full theory (or some approximation at a given order of perturbation theory to it). For example, Z_A will be chose such that the corrected photon propagator has the standard residue at the points with

$$p^2 = -(p^0)^2 + \mathbf{p}^2 = 0.$$

Similarly, Z_Ψ and Δm can be determined such that the fermion propagator including quantum corrections has poles at the frequencies p^0 such that

$$p^2 + m^2 = -(p^0)^2 + \mathbf{p}^2 + m^2 = 0,$$

and that these poles have the standard residue. This is in fact a bit intricate because of infrared divergences and the fact that a charged particle can experimentally not distinguished from a charged particle plus a very soft photon. (These two issues are closely connected and need to be dealt with simultaneously.)

Finally, the correction to the charge Δe will be determined such that the corrected vertex function represents the physical charge e.

Perturbative expansion

So far we have worked at leading order or tree-level and saw there how the QED Lagrangian lead to consistent physics results when loop diagrams could be neglected. In the spirit of perturbation theory we therefore expect that to the leading order in the coupling constant, or in the fine structure constant

$$\alpha = \frac{e^2}{4\pi}$$

we have simply $Z_\Psi = Z_A = 1$ and $\Delta m = \Delta e = 0$. In fact, from inspection of the diagrams we infer

$$Z_A - 1 = \mathcal{O}(\alpha), \qquad Z_\Psi - 1 = \mathcal{O}(\alpha), \qquad \Delta m = \mathcal{O}(\alpha), \qquad \Delta e/e = \mathcal{O}(\alpha).$$

The idea is now to allow so-called counterterms for these additional contributions to the QED Lagrangian and to determine them through the physics conditions mentioned above. This must be done consistently in perturbation theory, i. e. when one wants to calculate including the $\mathcal{O}(\alpha)$ corrections one must take counterterms at that order into account.

Photon self energy

We start with the effective photon propagator. Quantum corrections as the one-loop fermion diagram introduced above can be taken into account through a geometric series by writing the corrected photon propagator in momentum space as (known as Dyson equation)

$$G_{\mu\nu}(p) = \Delta_{\mu\nu}(p) + \Delta_{\mu\rho}(p)\Pi^{\rho\sigma}(p)\Delta_{\sigma\nu}(p) + \dots$$

Here we use the free photon propagator (in Lorenz or Landau gauge)

$$\Delta_{\mu\nu}(p) = \frac{\mathscr{P}_{\mu\nu}(p)}{p^2 - i\epsilon} = \frac{\eta_{\mu\nu} - p_\mu p_\nu/p^2}{p^2 - i\epsilon},$$

and $\Pi^{\rho\sigma}(p)$ is known as the photon self-energy. The Dyson equation can be better understood in terms of the inverse corrected propagator obtained by suming the geometric series,

$$p^2\eta^{\mu\nu} - p^\mu p^\nu - \Pi^{\mu\nu}(p) = p^2\,\mathscr{P}^{\mu\nu}(p) - \Pi^{\mu\nu}(p)$$

In other words, $-\Pi^{\mu\nu}(p)$ is the quantum correction to the inverse photon propagator! The first term is the microscopic or bare inverse propagator with which we have started our discussion of QED. For the corrected inverse propagator to be gauge invariant, we expect the structure

$$\Pi^{\mu\nu}(p) = \Pi(p^2)\left(p^2\eta^{\mu\nu} - p^\mu p^\nu\right),$$

where $\Pi(p^2)$ is now a scalar function of momentum. In perturbation theory at one-loop order, the photon self-energy, and therefore $\Pi(p)$ has contributions from the fermion loop diagram, as well as from the counter term $Z_A - 1$. The latter is in fact easy to determine and

$$\Pi(p^2) = \Pi_{\text{1-loop}}(p^2) - (Z_A - 1).$$

It remains to determine the loop contribution. Using the Feynman rules and dropping terms of higher order in α, we find

$$\Pi^{\mu\nu}_{\text{1-loop}}(p) = (-1)ie^2\int\frac{d^4l}{(2\pi)^4}\frac{\mathrm{Tr}\left\{[-i(\not{l}+\not{p})+m]\gamma^\mu[-i\not{l}+m]\gamma^\nu\right\}}{[(p+l)^2+m^2-i\epsilon][l^2+m^2-i\epsilon]}.$$

The minus sign comes from the closed fermion loop.

Evaluation of loop integral

To perform the Dirac trace one decomposes it into one term with four gamma matrices,

$$\mathrm{Tr}\left\{-(\not{l}+\not{p})\gamma^\mu\not{l}\gamma^\nu\right\} = 4\left[-(l+p)^\nu l^\mu + (l+p)\cdot l\,\eta^{\mu\nu} - (l+p)^\mu l^\nu\right],$$

and a term with two gamma matrices,

$$\mathrm{Tr}\left\{m^2\gamma^\mu\gamma^\nu\right\} = 4m^2\eta^{\mu\nu}.$$

For the denominator we introduce an integral over a Feynman propagator,

$$\frac{1}{[(p+l)^2+m^2-i\epsilon][l^2+m^2-i\epsilon]} = \int_0^1 du\frac{1}{[l^2+2ul\cdot p+up^2+m^2-i\epsilon]^2}$$

Now we can shift the integration variable setting $k = l + up$. We obtain, dropping linear terms in k and going to d dimensions,

$$\Pi^{\mu\nu}_{\text{1-loop}}(p) = -i4e^2\int\frac{d^dk}{(2\pi)^d}\int_0^1 du\frac{k^2\eta^{\mu\nu} - 2k^\mu k^\nu - u(1-u)p^2\eta^{\mu\nu} + 2u(1-u)p^\mu p^\nu + m^2}{\left[k^2+u(1-u)p^2+m^2-i\epsilon\right]^2}.$$

In the next step we replace $2k^\mu k^\nu \to (2/d)k^2\eta^{\mu\nu}$ and use the following identity valid in dimensional regularization (exercise)

$$\int\frac{d^dk}{(2\pi)^d}\frac{\left(\frac{2}{d}-1\right)k^2 - A}{[k^2+A]^2} = 0.$$

This yields

$$\Pi^{\mu\nu}_{\text{1-loop}}(p) = \left(p^2\eta^{\mu\nu} - p^\mu p^\nu\right)ie^2\int_0^1 du\int\frac{d^dk}{(2\pi)^d}\frac{8u(1-u)}{[k^2+u(1-u)p^2+m^2-i\epsilon]^2}.$$

Note that this has indeed the tensor structure of the classical inverse propagator for photons, as expected.

Scalar integral

We continue to evaluate the integral over k. We do the Wick rotation to Euclidean frequency, $k^0 = ik_{\mathrm{E}}^0$ which gives a factor i, and yields thus

$$\Pi_{\text{1-loop}}(p) = -8e^2 \int_0^1 du\, u(1-u) \int \frac{d^d k_{\mathrm{E}}}{(2\pi)^d} \frac{1}{[k_{\mathrm{E}}^2 + u(1-u)p^2 + m^2]^2}.$$

For the remaining integral we can now use the formula in (12.4). This gives

$$\int \frac{d^d k_{\mathrm{E}}}{(2\pi)^d} \frac{1}{[k_{\mathrm{E}}^2 + A]^2} = \frac{\Gamma(2 - d/2)}{(4\pi)^{d/2}} A^{-(2-d/2)}.$$

We need to evaluate this in the vicinity of $d = 4$ and set $d = 4 - \varepsilon$. We use

$$\Gamma(2 - d/2) = \Gamma(\varepsilon/2) = \frac{2}{\varepsilon} - \gamma + \mathcal{O}(\varepsilon),$$

and

$$(4\pi)^{d/2} = (4\pi)^{2-\varepsilon/2} = 16\pi^2 e^{-\frac{\varepsilon}{2}\ln(4\pi)} = 16\pi^2 \left[1 - \frac{\varepsilon}{2}\ln(4\pi) + \ldots\right],$$

as well as

$$A^{-(2-d/2)} = A^{-\varepsilon/2} = 1 - \frac{\varepsilon}{2}\ln(A).$$

Finally we need to take into account that electric charge is not dimensionless away from $d = 4$. We replace therefore e^2 with

$$e^2 \mu^\varepsilon = e^2 \left[1 + \frac{\varepsilon}{2}\ln(\mu^2) + \ldots\right],$$

where μ is some parameter with dimension of mass and e remains dimensionless.

Fixing the counterterm

Combining everything and expanding to leading order in ε gives

$$\Pi(p) = -\frac{e^2}{\pi^2} \int_0^1 du\, u(1-u) \left[\frac{1}{\varepsilon} - \frac{1}{2}\ln\left(\frac{u(1-u)p^2 + m^2}{4\pi e^{-\gamma}\mu^2}\right)\right] - (Z_A - 1).$$

One now fixes the counterterm $Z_A - 1$ such that $\Pi(0) = 0$, such that for on-shell photons the propagator is not modified. This leads to

$$Z_A - 1 = -\frac{e^2}{\pi^2} \int_0^1 du\, u(1-u) \left[\frac{1}{\varepsilon} - \frac{1}{2}\ln\left(\frac{m^2}{4\pi e^{-\gamma}\mu^2}\right)\right].$$

The integral over u can easily be performed here and gives a finite result. The important thing is that the counterterm is formally infinite due to the UV divergence of the involved momentum integrals.

Result for photon self-energy at one loop

Once the counterterm is subtracted from the self energy, we are actually left with a finite photon self-energy

$$\Pi(p) = \frac{e^2}{2\pi^2} \int_0^1 du\, u(1-u) \ln\left(\frac{u(1-u)p^2 + m^2}{m^2}\right).$$

Indeed this satisfies $\Pi(0)$ and is also independent of the arbitrary scale μ we had to intrduce in an intermediate step. For negative $p^2 = -(p^0)^2 + \mathbf{p}^2$ one finds that $\Pi(p)$ has a brach cut starting from the point where

$$(p^0)^2 - \mathbf{p}^2 > 4m^2.$$

This corresponds to the physical process where a virtual photon dissociates into an electron-positron pair. Energetically this is only possible when the virtual photon has an energy of at least twice the electron mass.

Final remark

One can now go on and analyse other loop diagrams in this spirit. The story is in principle similar to what we have seen for the photon self-energy. Counterterms in the form of the terms in the Lagrangian are needed to counterbalace UV divergences, and they can be fixed through physical renormalization conditions. Within perturbation theory one obtains then finite results, except when additional infrared divergences appear. These need a somewhat differnt treatment and have a different physical significance. After proper renormalization one can calculate finite terms, such as the electron magnetic moment. We leave a more detailed discussion of renormalization for a continuation of this lecture course in the next term.